



Op zoek naar Vaste Waarden

Vooronderzoek Digitale Archivering aan het Vlaams Parlement, 2005-2006

... digitaal: een reeks gehele getallen voorgesteld als vaste waarden ...

... analoog: werken met waarden in een continuüm zonder stappen ...

Vincent Sterken
Eindverhandeling
Master in de Archivistiek: Erfgoed- en Hedendaags Documentbeheer
Academiejaar 2005-2006
Promotor: Prof. Dr. F. Scheelings
Stagebegeleider: Mevr. H. Van Ongevalle

Dankwoord

In de eerste plaats zou ik mijn erkentelijkheid willen uiten ten opzichte van de medewerkers van de opleiding archivalie aan de Vrije Universiteit Brussel en het Vlaams Parlement, om mij de kans te geven dit onderzoek te kunnen uitvoeren.

Voor bijkomende ondersteuning op het vlak van bepaalde concepten bedank ik prof. Dr. Frank Scheelings en Stephanie Waeyenbergh.

Voor een fijne periode van samenwerking bedank ik Luc Vanpaemel van het Archief en de Biografische Dienst (Vlaams Parlement) en mijn collega student Soizik Van Huele.

Ik bedank Katleen Sledz voor het *last-minute* nalezen van deze verhandeling en het verbeteren van die vervelende taalfouten.

En niet te vergeten: dank u Hilde Van Ongevalle, archivaris van het Vlaams Parlement, voor de morele ondersteuning en het uitgevoerde onderzoek in goede banen te leiden.

Inhoudsopgave

Dankwoord

Inleiding

p. 1

Deel I: Het theoretische verhaal in kaart gebracht

1. Digitale problematiek

1.1 *A fine romance ...: computerbestanden en archiefstukken* p. 5

1.2 *Het digitale archiefdocument*

2. Wetenschappelijk onderzoek

2.1 *Preserveringsstrategieën* p. 10

A. Hard Copy p. 10

B. Computermuseum p. 11

C. Conversie p. 11

D. Migratie p. 12

E. Emulatie p. 14

2.2 *Het digitaal depot*

A. Institutioneel depot p. 17

B. Betrouwbaar Digitaal Depot (Trusted Digital Repositories) p. 18

C. Open Archival Information System (OAIS) p. 20

2.3 *Standaarden* p. 25

A. Electronic Document and Records Management Systems p. 26
ISO 15489

Model Requirements for the Management of Electronic Records (MoReq)

DoD 5015.2-STD

The National Archives (UK)

The National Archives of Australia

B. Internationale Archiefraad p. 31

C. Encoded Archival Description (EAD) p. 32

D. Dublin Core Metadata Initiative p. 33

E. Technische standaarden p. 32

3. Bevindingen

p. 36

Deel II: De eerste stappen in de praktijk

| | |
|---|--------|
| 1. Het archief van het Vlaams Parlement als Betrouwbaar Digitaal Depot uitbouwen | p. 41 |
| 1.1 <i>Criteria voor een Betrouwbaar Digitaal Depot</i> | |
| A. Organisatie | p. 42 |
| B. Functies, processen en procedures | p. 44 |
| C. De <i>designated community</i> en de bruikbaarheid van informatie | p. 47 |
| D. Technologieën en technische infrastructuur | p. 48 |
| 1.2 <i>Implicaties voor het archief van het Vlaams Parlement</i> | |
| A. Organisatie | p. 50 |
| B. Functies, processen en procedures | p. 52 |
| C. De <i>designated community</i> en de bruikbaarheid van informatie | p. 57 |
| D. Technologieën en technische infrastructuur | p. 60 |
| 2. De technologische omgeving van het Vlaams Parlement | |
| 2.1 <i>Een korte introductie</i> | p. 63 |
| 2.2 <i>De gebruikte software en formaten</i> | |
| A. Tekstverwerking | p. 65 |
| B. Spreadsheets | p. 66 |
| C. Databanken (databases) | p. 66 |
| D. E-mail | p. 69 |
| E. Websites | p. 71 |
| F. Audio | p. 74 |
| G. Video | p. 77 |
| 3. Inventarisatie van software en formaten, met aanbevelingen voor bewaring | p. 80 |
| 3.1 <i>De bevraging (ICT enquête)</i> | p. 81 |
| 3.2 <i>Analyse</i> | |
| A. Algemeen | p. 89 |
| B. Kantoordocumenten | p. 90 |
| C. E-mail | p. 94 |
| D. Databanken | p. 98 |
| E. Website | p. 102 |
| F. Audio | p. 108 |
| G. Video | p. 110 |
| 3.3 <i>Dragers</i> | p. 112 |
| 3.4 <i>Rol van de digitale handtekening</i> | p. 114 |
| Besluit | p. 118 |



Inleiding

*The mutation in technology changes not simply the archiving process,
but what is archivable –
that is, the content of what has to be archived is changed by the technology.¹*

De digitale revolutie heeft de laatste twintig jaar voor ingrijpende veranderingen gezorgd in onze manier van leven en werken. Meer en meer informatie wordt verwerkt via computers, waardoor vele routinematige handelingen efficiënter verlopen, bepaalde processen goedkoper worden en beslissingen sneller kunnen worden genomen. Hierbij lijkt het belang van papier stilaan aan kracht in te boeten. Er worden minder brieven geschreven, bepaalde handboeken die aan regelmatige herziening toe zijn kunnen integraal online geraadpleegd worden en voor sommige juridische documenten zijn nu ook al elektronische handtekeningen mogelijk (om maar enkele voorbeelden te geven).

Nietteminste staande zitten we in een overgangsfase. De computer is nog steeds een relatief nieuw product (historisch gesproken) en moet zijn waarde in de huidige vorm nog bewijzen. De ontwikkeling van het digitale instrumentarium is nog in volle evolutie. Het gevolg hiervan is de ontwikkeling van een samenleving die, i.p.v. vast te roesten in de bestaande tradities, zich richt op de toekomst. Helaas, voor de archivistische professie, wordt niet genoeg tijd genomen om naar het verleden te kijken – het toekomstige verleden dan toch. Archiefbestanden die op de “klassieke” manier worden aangemaakt, kunnen door onze zintuigen worden gelezen en geanalyseerd. Digitale bestanden zijn per definitie virtueel en vereisen daarom een intermediair middel om ze te kunnen visualiseren, d.i. de computer zelf met de bijbehorende software. De snelle opeenvolging van technologieën, bestandsformaten en vernieuwde applicaties zorgen ervoor dat informatie die nu wordt opgeslagen binnen vijf tot maximum tien jaar niet meer zal kunnen opgeroepen worden binnen een zinvolle context. De digitale omgeving waarin wij dagelijks werken zal nog voor geruime tijd aan verandering onderhevig zijn. Het is niet gezegd dat de oplossing van vandaag nog steeds operationeel zal zijn binnen honderd jaar. De archivariissen zijn ondertussen gelukkig niet bij de pakken blijven zitten. Een aantal oplossingen werd al voorgesteld², maar dit blijft een doekje voor het bloeden. Er is nog niet genoeg tijd verstreken om na te gaan of deze mogelijkheden op

¹ Derrida, J. *Refiguring the archive*, 2001, zoals geciteerd in Ketelaar, E. *Tacit narratives: the meanings of archives*. In: Ketelaar, E. *Tacit narratives: the meanings of archives*. In: Thomassen, T., Looper, B., Kloosterman, J. (eds.), *Toegang. Ontwikkelingen in de ontsluiting van archieven*, 's-Gravenhage: Stichting Archiefpublicaties, Jaarboek 2001, p.71

² oa. migratie, emulatie, en het invoeren van bepaalde standaarden en EDRMS. Boudrez, F. & Dekeyzer, H. *Digitaal archiefbeheer in de praktijk. Handboek*. Stadsarchief Antwerpen, 2004, p. 51 - 70

de lange termijn de bestaande documenten veilig zullen stellen. Maar, op zijn minst kunnen documenten op de middellange termijn op een duurzame manier bewaard worden, terwijl verder gezocht wordt naar andere, of verbetering van de bestaande, methodes.

Voor dat naar afdoende manieren kon gezocht worden om digitale archiefbestanden veilig te stellen, was eerst nood aan een nieuw paradigma binnen de archivistiek. Zo lang pas gearchiveerd kan worden nadat de bestanden worden over gegeven, lopen we eigenlijk achter de feiten aan. Op dat moment wordt het moeilijk om alles in een juiste context te plaatsen. Dit is nog mogelijk bij een dynamische administratie waar de archiefvormer nog in directe relatie staat met het geproduceerde document. Maar hoe meer tijd tussen de overdracht en de beëindiging van de activiteiten ligt, hoe moeilijker het wordt. Daarom dat Frank Upwards' Concept van het *Records Continuum* zo belangrijk is.³ Hierbij wordt met de idee van *the life cycle model* gebroken, dat stelt dat alle archiefbescheiden (records) een vast stramien doorlopen van een dynamische naar statische levensfase. "The life cycle model sees records passing through stages until they eventually 'die', except for the 'chosen ones' that are reincarnated as archives." Terwijl een "continuum is something continuous of which no separate parts are discernible, a continuous series of elements passing into each other."⁴ Het concrete gevolg hiervan dat van bij de creatie van een document al moet ingegrepen worden.

In het kader van deze problematiek heb ik een stage gevolgd aan het Vlaams Parlement. "Het Vlaams Parlement belichaamt de wetgevende pijler en de democratie in Vlaanderen. Het staat in voor de totstandkoming van de Vlaamse decreten en voor de goedkeuring van de Vlaamse begroting. Het Vlaams Parlement benoemt de Vlaamse Regering en controleert de Vlaamse Regering op de uitvoering van die decreten."⁵

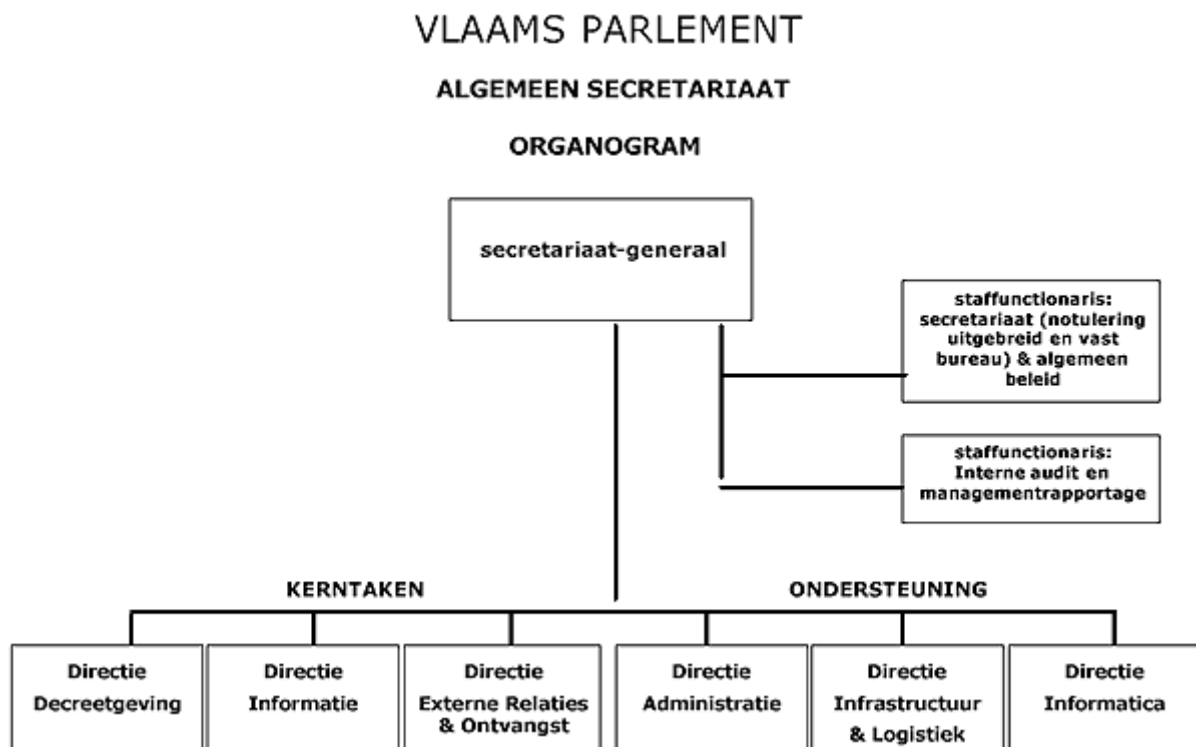
Het Vlaams Parlement heeft een sterk afgeplatte structuur. De dagelijkse leiding wordt waargenomen door de voorzitter van het Parlement, 4 ondervoorzitters en 3 secretarissen. Dit wordt het (Vast) Bureau genoemd. Bij het begin van elk zittingsjaar worden tijdens de plenaire vergadering de leden gekozen uit de Vlaamse volksvertegenwoordigers, volgens het principe van de evenredige verdeling. Zij zijn verantwoordelijk voor het administratieve beheer. Voor de regeling van zuiver politieke aangelegenheden wordt het Bureau aangevuld met de fractievoorzitters van de erkende fracties, d.i. het Uitgebreid Bureau. De belangrijkste politieke taken worden door de voorzitter waargenomen. In de uitoefening van zijn ambt wordt hij bijgestaan door het Bureau en eventueel het Uitgebreid Bureau. De diensten van het Vlaams Parlement zorgen voor de administratieve ondersteuning. Aan het hoofd staat een griffier met de rang

³ Upward, F. In Search of the continuum: Ian McLean's 'Australian experience' essays on recordkeeping. In: McKemmish, S. & Piggot, M. (eds.), *The Records Continuum: Ian McLean and the Australian Archives first fifty years*, Canberra: Ancora Press, 1994, pp. 110-130

⁴ McKemmish, S. Yesterday, today and tomorrow: a continuum of responsibility. In: Horsman, P.J.; Ketelaar, F.C.J. & Thomassen, T.H.P.M., *Naar een nieuw paradigma in de archivistiek*, 's-Gravenhage: Stichting Archiefpublicaties, 1999, p. 195

⁵ Missieverklaring, <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/HoeWerktHetVlaamsParlement/Missie/MissieVerklaring.html>

van secretaris-generaal, benoemd door de plenaire vergadering. Momenteel worden ruim 200 ambtenaren te werk gesteld door het Secretariaat-Generaal en de verschillende Directies.⁶ Aan het hoofd van het Secretariaat-Generaal staat dus de secretaris-generaal, bijgestaan door de Directieraad. Daaronder staan enerzijds twee staffunctionarissen en anderzijds de Directies Decreetgeving, Informatie, Externe Relaties & Ontvangst, Administratie, Infrastructuur & Logistiek en Informatica. Deze worden weer verder onderverdeeld in een aantal Cellen. Zo valt de Cel Archief & Biografische Dienst onder de Directie Informatie.



Bron: Website van het Vlaams Parlement,
<http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/HoeWerktHetVlaamsParlement/AlgemeenSecretariaat/InKortBestek.html>

Voor het papieren archief werd reeds een duidelijke regeling getroffen⁷. Voor het digitale archief moeten nog richtlijnen opgesteld worden. Het was mijn taak om Mevrouw Van Ongevalle, de huidige archivaris, bij te staan in haar zoektocht naar een adequate digitale archiveringsstrategie. Tijdens mijn stage, werd getracht de digitale werkomgeving in kaart te brengen. Via een enquête werd nagegaan welke digitale middelen gebruikt worden tijdens het uitvoeren van de processen door de verschillende archiefvormende eenheden. Vervolgens werd gekeken in hoeverre de opgeslagen bestandsformaten waarin de gegevens die hierbij gegenereerd worden, geschikt zijn voor archiveringsdoeleinden. Het betreft hier niet alleen data die rechtstreeks

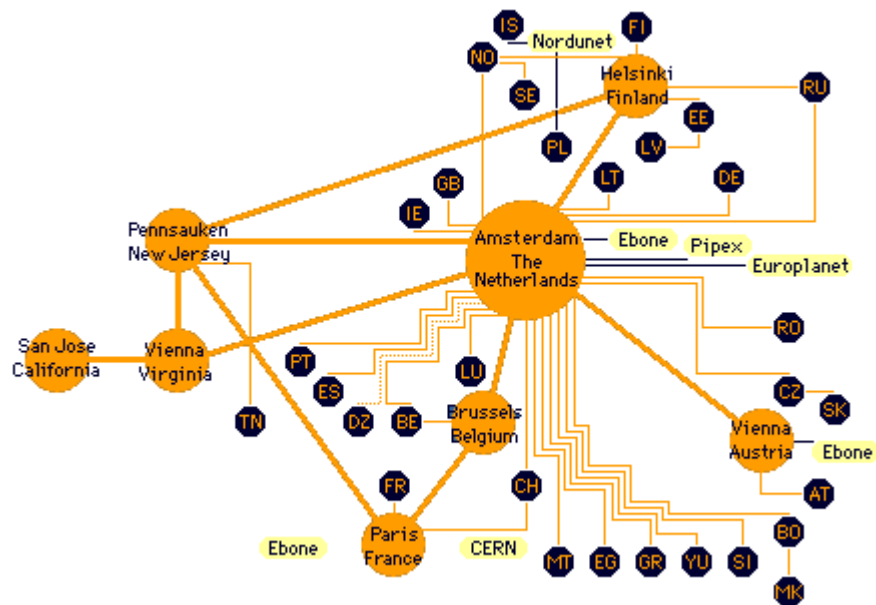
⁶ Van Esbroeck, M. & Goossens, M. Het Vlaams Parlement. Tielt: Uitgeverij Lannoo, 2003, pp. 171-173

⁷ <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/Pdf/20042005/Archief/Archiefreglement.pdf>, 08/10/05

bewerkt worden door de administratie, zoals tekstverwerkingsbestanden, spreadsheets, databases en e-mails, maar ook audiovisuele opnames. “Het ... is een globaal onderzoek en spitst zich niet toe op één specifiek soort archief.” We spreken dus over een “inventarisatie en analyse van de op het Vlaams Parlement gebruikte servers, applicaties, systemen, databanken en het volume en de aard van documenten en gegevens die zij genereren. Het is de bedoeling om in het licht van de digitale archivering de knelpunten inzake de digitale bewaarstrategieën, gebruikte formaten, metadata aan het licht te brengen.”⁸ Hieruit voortvloeiend zal dan een advies opgesteld worden naar de toekomst, waarbij aandacht wordt besteed aan de digitale noden van de instelling.

⁸ Vanongevale, H. Hearing rond digitaal archiveren. Gent: IPEA, 12 december 2005 (onuitgegeven)

Deel I: Het theoretische verhaal in kaart gebracht



Kaart van internetverbindingen in Europa
Bron: International Policy Fellowships, <http://www.policy.hu/inetclass/index.html>

1. Digitale problematiek

It is to information professionals that future generations will look if they find no surviving history of the early electronic age. It will not be the fault of those who brought the electronic age into being; it will be purely the fault of those who did not adept their thinking and their solutions to the changing requirements of a new information environment.¹

1.1 A fine romance...²: computerbestanden en archiefstukken

Sinds de opkomst van de *personal computer* tijdens de jaren tachtig van de vorige eeuw wordt het werk in overheden (en privé-bedrijven) meer en meer geautomatiseerd. Stilletjes, men zou bijna kunnen zeggen op kousenvoeten, zijn we aan het evolueren van een fysieke naar een digitale realiteit op de werkvloer. Een grote hoeveelheid interne en externe documenten wordt nu digitaal aangemaakt en zal dat ook blijven. Een papieren versie van een tekstbestand kan dan nog afgedrukt worden omwille van het leesgemak, een databank echter niet.

Elke overheid produceert een aanzienlijke hoeveelheid documenten (zowel op papier als digitaal), die het bewijs vormen van een handeling, van de uitvoering van een taak of om bepaalde rechten te handhaven. Om in de toekomst dit doel te kunnen dienen, moet gedacht worden aan een duurzame manier van bewaring. Bij digitale documenten stellen zich een aantal specifieke problemen.

In tegenstelling tot “klassieke” archiefstukken, kunnen digitale bescheiden niet met het blote oog worden waargenomen. Voor het bekijken van digitale archiefdocumenten hebben we een welbepaalde hard- en softwareconfiguratie nodig.³ Sinds de jaren zestig hebben computers en de gebruikte software een enorme evolutie gekend. “One of the marvels of the information technology revolution is the continuous improvement in computer memory and storage performance and their simultaneous drop in cost. Thanks to what has been called “silicon scaling” the processing power of a 1980s vintage mainframe computer now fits on miniscule silicon chips ... One might suspect that archiving and preserving digital information would become easier and cheaper as a consequence of these improvements. But ... The technologies, strategies, methodologies, and resources needed to manage digital information for the long term have not kept pace with

¹ Stuckey, S. Forward. In: Judith A. Ellis (ed.), *Selected essays in electronic recordkeeping in Australia*, Victoria: Australian Society of Archivists Inc., 2000, p. (v)

² Billie Holiday, A fine romance. In: *Lady Day Swings*, Sony, 2002 (eerste opname: 29 september 1936)

³ Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief DAVID Project, 2004, p. 52

innovations in the creation and capture of digital information.”⁴ Een gevolg hiervan is de snelle veroudering van de gebruikte technologie. Een document dat nu wordt geproduceerd, zal in datzelfde formaat naar alle waarschijnlijkheid binnen vijf jaar niet meer kunnen gelezen worden. Een wordperfect-document, bijvoorbeeld, kan vandaag in een XP omgeving niet meer worden geopend. Niet alle formaten zijn bovendien geschikt voor langdurige bewaring vanuit een archivistisch standpunt (bvb. omwille van copyright).

Deze evolutie heeft eveneens gezorgd voor een grote verscheidenheid in formaten en vormen. De traditionele soorten, zoals geluidsopnames, foto's en tekstverwerking, bestaan nog steeds, maar hiernaast werden nieuwe elementen geïntroduceerd, zoals hypertext, dynamische pagina's, geografische informatiesystemen, multimedia en interactieve video. Hierdoor ondergaat de inhoud van een document een aantal wijzigingen. Websites bijvoorbeeld bevatten eveneens links naar andere webpagina's en extern bewaarde databanken. Een bladzijde opgemaakt met een tekstverwerkingsprogramma kan, buiten de eigenlijke tekst, grafieken, afbeeldingen, voetnoten en interactieve boodschappen bevatten.⁵

Een document is pas begrijpbaar wanneer het doel ervan duidelijk is. In andere woorden, een stuk krijgt betekenis indien het binnen zijn context⁶ kan geplaatst worden. Voor het papieren archief kan dit nog achterhaald worden achteraf door middel van historisch onderzoek. Binnen een digitale omgeving is dit bijzonder moeilijk, al dan niet onmogelijk. Eigenlijk is het vereist om, indien mogelijk, van bij het ontstaan of, ten laatste, voor de archivering bijkomende, contextuele gegevens toe te voegen door middel van metadata. Hierdoor kan het (blijvende) administratieve en/of historische nut behouden blijven. Het zoeken en terugvinden wordt ook vergemakkelijkt door het toekennen van metadata.⁷ Bovendien moet de integriteit en de authenticiteit van het bewaarde document gewaarborgd blijven. In termen van gebruiksgemak is het grootste voordeel van een digitale omgeving dat documenten steeds kunnen gewijzigd worden. Dit staat uiteraard haaks op

⁴ Hedstrom, M. It's about time: Research challenges in digital archiving and long-term preservation. Final report. Workshop april 12-13 2002. The Library of Congress: Digital Preservation, National digital information infrastructure and preservation program, 2003, p. vii

http://www.digitalpreservation.gov/repot/NSF_LC_Final_Report.pdf , 16/02/06

⁵ Chen, S. The paradox of digital preservation. University of Missouri-Columbia, 2001, p. 3

<http://www.gseis.ucla.edu/us-interpares/pdf/ParadoxOfDigitalPreservation.pdf> , 17/02/06

⁶ “Context ... het geheel van administratief-organisatorische, bestuurlijk-juridische en technische gegevens, waarbinnen de functie van het archiefstuk, archiefbestanddeel of archiefafdeling in relatie tot de activiteiten en taken van de archiefvormer moet worden geïnterpreteerd.” den Teuling, A.J.M. (ed.) Archiefterminologie voor Nederland en Vlaanderen. 's Gravenhage: Stichting Archiefpublicaties, 2003, artikel 20

⁷ De toekenning van metadata moet echter zo veel mogelijk automatisch gebeuren, omdat voor de gemiddelde gebruiker zelfs het invullen van de eigen naam en een onderwerp soms al te veel moeite is. Boudrez, F. Digitaal archiveren van archiefdocumenten. Casus: e-mail. Antwerpen : Stadsarchief Antwerpen, Onuitgegeven lezing, 6 maart 2006; Cory Doctorow stelt zelfs dat een aantal intrinsieke eigenschappen van mensen (algemeen) aan de basis liggen van slechte metadata indien ze die zelf moeten ingeven. Relevant voor archivistische metadata zijn de opmerkingen dat mensen lui zijn en niet nauwgezet omgaan met gegevens, waardoor (volgens de auteur) de enige oplossing “sending out squads of vengeful info-ninjas to add metadata to the average user's files” kan zijn. Doctorow, C. Metacrap: Putting the torch to seven straw-men of the meta-utopia. 26 august 2001, <http://www.well.com/~doctorow/metacrap.htm#0> , 17/02/06

de premisse dat archiefdocumenten vast en ongewijzigd moeten worden overgeleverd. De archivaris moet bijgevolg erover waken dat de bescheiden die hem⁸ werden toevertrouwd niet onrechtmatig veranderd kunnen worden en dat eventuele manipulaties opgespoord en ongedaan kunnen gemaakt worden.⁹

1.2 Het digitale archiefdocument

Indien we de term *digitaal archiefdocument* ontleden, vinden we een aantal elementen terug. De belangrijkste vooronderstelling is het besef dat *electronic records*, als complexe structuren, een aantal eigenschappen erven van verschillende klassen van objecten.¹⁰

Een eerste belangrijk kenmerk is uiteraard het digitaal zijn. Een digitaal archiefdocument is per definitie een digitaal object. Een digitaal object kan summier omschreven worden als “an object composed of a set of bit sequences.”¹¹ Bij uitbreiding kan gesteld worden dat het een informatie object betreft, die eender welk type van informatie kan bevatten of uit eender welk formaat kan bestaan, in een digitale vorm.¹² Het betreft hier zowel informatie die *born-digital* is, als analoge informatie die gedigitaliseerd werd.¹³ Een direct begrijpbare definitie wordt ons aangeboden door The National Archives of Australia: “an object that can be represented by a computer, such as a file type generated by a particular system or software application.”¹⁴

Een tweede element vinden we in het woord *archiefdocument*. Een document kan omschreven worden als “recorded information or object which can be treated as a unit.”¹⁵ Den Teuling voegt hier nog aan toe dat het een “geheel van samenhangende gegevens, vastgelegd op een of meer gegevensdragers” betreft, waarbij “samenhangende digitale gegevens, vastgelegd op meer dan

⁸ Of haar uiteraard. Voor het leesgemak echter zal ik bij volgende verwijzingen enkel “hem” gebruiken.

⁹ Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief DAVID Project, 2004, p. 53

¹⁰ Thibodeau, K.; Moore, R. & Baru, C. Persistent object preservation : Advanced computing infrastructure for digital preservation. In: *Proceedings of the DLM-Forum 1999, European citizens and electronic information: the memory of the information society*, Luxemburg: European Communities, Brussel 18-19 oktober 1999, http://europa.eu.int/ISPO/dlm/fulltext/full_thib_en.htm, 24/03/06

¹¹ Consultative Committee for Space Data Systems, Reference model for an Open Archival Information System (OAIS). Blue Book. Washington: CCSDS, National Aeronautics and Space Administration, 2002, p. 1-10 <http://public.ccsds.org/publications/archive/650x0b1.pdf>, 08/03/06

¹² Digital object. In: *Wikipedia, the free encyclopedia*, http://en.wikipedia.org/wiki/Digital_object, 23/03/06

¹³ Horsman, P.J.; van Bussel, G.J. & Waalwijk, H. Softwarespecificaties voor Records Management Applicaties voor de Nederlandse Overheid. Remano 2004. Amsterdam: Archiefschool, maart 2004, p. 19
Zie ook: http://www.archiefschool.nl/docs/ReMANO_2004.pdf

¹⁴ National Archives of Australia. Functional specifications for electronic records management systems software. Exposure draft. Commonwealth of Australia, 2006, p. 63, <http://www.naa.gov.au/recordkeeping/er/erms/ERMSspecifications.pdf>, 20/02/06

¹⁵ IDA, Model Requirements for the management of electronic records. MoReq Specification. Prepared by Cornwell Management Consultants plc, October 2002, p. 7
<http://www.cornwell.co.uk/moreqdocs/MoReq%20Specification%20v5-2.4.doc>, 06/11/05

één gegevensdrager, één document [vormen] als zij bestemd zijn om in één actie te worden opgeroepen.” Maar ook “een document waarvan gedeelten op verschillende typen gegevensdragers zijn vastgelegd.”¹⁶

Ten derde moet benadrukt worden dat het niet zo maar een document betreft, maar een archiefstuk (*record*). Dit is een “document, ongeacht zijn vorm, naar zijn aard bestemd om te berusten onder de persoon, groep personen of organisatie die het heeft ontvangen of opgemaakt uit hoofde van zijn of haar activiteiten, zijn of haar taken of ter handhaving van zijn of haar rechten.”¹⁷

ISO 15489 geeft als karakteristieken van een digitaal archiefstuk¹⁸ de volgende elementen aan:

- *Authenticiteit*: het moet kunnen bewezen worden dat een archiefstuk is wat het beweert te zijn.¹⁹
- *Betrouwbaarheid*: het moet mogelijk zijn erop te vertrouwen dat de inhoud van een stuk een accurate voorstelling is van de transactie waarvan het getuige is.
- *Integriteit*: het moet mogelijk zijn om een stuk te beschermen tegen ongeautoriseerde veranderingen, én geautoriseerde veranderingen na te gaan.
- *Herbruikbaarheid*: garanties moeten ingebouwd worden om een stuk terug te vinden, op te halen, ter beschikking te stellen en begrijpelijk te maken binnen de originele context.

Hier kunnen we nog twee bijkomende eisen aan toevoegen²⁰:

- *Uitvoerbaarheid*²¹: de dragers moeten intacte bitstreams bevatten.
- *Visualiseerbaar*: de bitstreams moeten door de computer correct weergegeven worden op het scherm.

Omwille van het feit dat het digitale objecten zijn, zijn digitale bescheiden anders in een aantal opzichten dan papieren documenten. Boudrez somt ze als volgt op²²:

- De opslagwijze en de verschijningsvorm zijn niet dezelfde: op de gegevensdrager wordt de informatie in bits en bytes opgeslagen, die vervolgens op het scherm in zijn intellectuele vorm wordt gepresenteerd.

¹⁶ den Teuling, A.J.M. (ed.) *Archiefterminologie voor Nederland en Vlaanderen*. 's Gravenhage: Stichting Archiefpublicaties, 2003, artikel 3

¹⁷ Den Teuling, A.J.M. (ed.), *Op. Cit.*, artikel 8

¹⁸ ISO 15489-1: 2001 (E), Information and documentation – Records Management –Part 1: General, p. 7

¹⁹ “it must be possible to prove that a record is what it purports to be and that it has been created or sent by the alleged person and at the time purported. Records need to be protected against unauthorised addition, deletion, alteration, use or concealment and the creation, receipt, transmission of records needs to be controlled to ensure that records creators are authorised and identified.”

²⁰ Deze zitten eigenlijk al vervat in het vorige, maar ik voeg ze bij voor de duidelijkheid. Boudrez, F. & Dekeyser, H. *Digitaal archiefbeheer in de praktijk*. Antwerpen: Stadsarchief DAVID Project, 2004, p. 56

²¹ Uitvoeren wordt hier gebruikt als vertaling voor ‘to render’, en niet voor exporteren.

²² Boudrez, F. & Dekeyser, H., *Idem*, p.54

- Drager en archiefdocument vormen geen eenheid meer: wijzigingen zijn niet meer visueel waarneembaar.
- Digitale objecten zijn enkel machineleesbaar: hard- en software is vereist om de bits om te zetten naar een documentaire vorm, die enkel raadpleegbaar is aan de hand van de overeenkomstige apparatuur en programmatuur.
- In een digitale omgeving bestaat het verschil tussen origineel en kopie niet: de gekopieerde bitstream is technisch identiek aan de originele.
- Digitale objecten hebben verschillende verschijningsvormen: weergave is mede afhankelijk van de computer- en gebruikersinstellingen.
- Één document kan verschillende bitrepresentaties hebben: een Word-document bijvoorbeeld kan zowel opgeslagen worden in Rich Text Format als in XML of PDF.
- Er is geen vaste relatie tussen digitale bescheiden en computerbestanden. De relatie varieert van:
 - Één op één: het document wordt in één bestand opgeslagen.
 - Één op veel: het document bestaat uit meerdere bestanden.
 - Veel op één: meerdere documenten worden in één bestand bewaard.

2. Wetenschappelijk onderzoek

2.1 Preserveringsstrategieën

A. Hard copy

Een hard copy is een fysieke weergave van een digitaal document. Bij deze strategie wordt het stuk overgezet naar microfilm of afgedrukt op papier.

Dit heeft het voordeel dat bepaalde documenten makkelijker leesbaar kunnen worden. Op de korte termijn is deze manier van werken soms aan te raden om onherroepelijk verlies van gegevens tegen te gaan bij het ontbreken van een degelijke procedure voor het beheer van digitale stukken.¹ Microfilm (gebruikt voor het bewaren van foto's bvb.) kent, bij een degelijke bewaring, een bewaartermijn tot 500 jaar.² Binnen de vertrouwde papieren werkomgeving kunnen deze documenten makkelijk ingepast worden.³

De nadelen zijn echter groter dan de voordelen. Nu wordt nog impliciet gewerkt vanuit een "papiermentaliteit", maar dit zal niet lang meer blijven duren. Meer en meer wordt enkel de digitale kopie als het origineel beschouwd (en gebruikt). Vanuit archivistisch perspectief is de digitale versie bovendien het origineel, en is een hard copy, zoals het woord al aangeeft, een kopie. Bijgevolg mag het digitale origineel enkel vernietigd worden na goedkeuring van de Algemeen Rijksarchivaris of diens gemachtigde (Archiefwet 24 juni 1955, artikel 5).⁴

Bij overzetting naar papier gaan de digitale eigenschappen van het document verloren. Voordelen zoals geautomatiseerde zoekopdrachten, herbruikbaarheid, ter beschikking stellen via een netwerk, ... zijn niet meer mogelijk.⁵ Niet alle documenten kunnen afgedrukt worden, niet in zijn geheel of niet

¹ Heazlewood, J. Management of electronic records over time. In: Judith A. Ellis (ed.), *Selected essays in electronic recordkeeping in Australia*, Victoria: Australian Society of Archivists Inc., 2000, pp. 106-107

² Position statement on microfilm preservation. ICON: International Coalition on Newspapers, June 15, 2006, <http://icon.crl.edu/icon/statement.htm>, 21/07/06

³ Coy, W. Perspektiven der Langzeitarchivierung multimedialer Objekte. Frankfurt am Main: Nestor c/o Die Deutsche Bibliothek – Berlin: Humboldt-Universität zu Berlin, 2006, p. 24-26, <http://nbn-resolving.de/urn:nbn:de:0008-20051214015>, 16/03/06

⁴ Janssens, G. & Vermote, M. Het archiefwezen in België. Berchem: VVDAB, 2004, p. 1 <http://www.fondspascaldecroos.org/organisaties/opleidingen/researchjournalistiek/hennekam/archiefwezenBelgie.PDF>, 05/04/06; Coppein, B. Een onneembare burcht? Een toelichting bij de Belgische archiefwet en de pogingen om haar te vervangen. Lezing gehouden op de FAAD-studiedag "50 jaar Belgische archiefwet" van 23 april 2005, <http://www.vub.ac.be/archivistiek/FAAD/files/bc230405.doc>, p. 2

⁵ Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief DAVID Project, 2004, p. 58

zonder verlies aan informatie. Hierbij denken we aan databanken, multimediaobjecten, GIS, ...⁶

B. Computermuseum

De onderliggende idee hier is het bewaren van de originele omgeving waarin de documenten werden gecreëerd, i.e. het onderhouden van de originele hard- en software. Op de korte termijn kan op deze manier worden bewaard (5 tot 10 jaar). Voor de middellange en lange termijn is het niet haalbaar. Dit betekent dat alle mogelijk configuraties, applicaties en machines die ooit werden gebruikt, bewaard moeten blijven. De kennis om met verouderde systemen en computers te werken verdwijnt na verloop van tijd, evenals de technische en de materiële ondersteuning.⁷

Oude digitale documenten worden in de meeste gevallen niet blijvend bewaard op de originele digitale media. Een tekstbestand dat aanvankelijk op een floppy werd opgeslagen, zal nu misschien zijn overgezet naar een DVD. Het gevolg hiervan is dat de oorspronkelijke computer die werd onderhouden om het tekstbestand in zijn verouderd formaat te kunnen lezen in zijn “natuurlijke” omgeving, geen DVD-drive bevat, noch een aangepaste interface of *driver software*.⁸

Een computermuseum kan weliswaar dienen als bron voor historisch onderzoek naar het vroegere gebruik van IT, maar daar bestaan al gespecialiseerde instituten voor.⁹

C. Conversie

Wanneer digitale documenten worden overgezet naar een *up-to-date* systeem configuratie, spreken we van conversie. Het originele bestandsformaat wordt dan omgezet naar de laatste, nieuwe versie. Zo kan een *Post Script* bestand worden geconverteerd naar PDF, of een MS Word 3.1 bestand naar MS Word 2003. Op deze manier blijven documenten functioneel.

Ook hier echter zijn een aantal opmerkingen te maken over authenticiteit. Een conversie naar een nieuw formaat creëert in feite een nieuw document en kan dus als kopie worden aanzien. Bovendien veranderen bij elke nieuwe versie van een applicatie bepaalde eigenschappen, waardoor

⁶ Coy, W. *Perspektiven der Langzeitarchivierung multimedialer Objekte*. Frankfurt am Main: Nestor c/o Die Deutsche Bibliothek – Berlin: Humboldt-Universität zu Berlin, 2006, p. 24 <http://nbn-resolving.de/urn:nbn:de:0008-20051214015>, 16/03/06

⁷ Boudrez, F. & Dekeyser, H. *Digitaal archiefbeheer in de praktijk*. Antwerpen: Stadsarchief DAVID Project, 2004, p. 59

⁸ Rothenberg, J. *Avoiding technological quicksand: Finding a viable technical foundation for digital preservation*. Washington: Council on Library and Information Resources, 1998, <http://www.clir.org/pubs/reports/rothenberg/inadequacy.html>, 05/04/06

⁹ Enkele voorbeelden: <http://www.computermuseum.nl/>, <http://www.computermuseum-muenchen.de/>, <http://www.compustory.com/>, <http://www.computermuseum.li/>

conversieproblemen en verlies aan functionaliteiten en informatie zich kunnen voordoen. Bij moet elke versiewijziging moet geconverteerd worden, wat het kostenplaatje, zowel in termen van tijd als van geld, fors de hoogte injaagt.¹⁰

Een bijkomend probleem ligt in het feit dat de oorspronkelijke formaten vaak software-, producent- en versiegebonden zijn. Technische ondersteuning op de lange termijn kan dus niet gegarandeerd worden.¹¹

D. Migratie

In de literatuur wordt conversie en migratie nog al eens door elkaar gebruikt wanneer het gaat over het omzetten van één formaat naar een ander. Een mogelijke definitie werd voorgesteld in een rapport van de *Task Force on Archiving of Digital Information*: “Migration is the periodic transfer of digital materials from one hardware/software configuration to another, or from one generation of computer technology to a subsequent generation.”¹² In die zin slaat conversie dan op het laatste deel van de definiëring. Om verwarring te vermijden, geef ik dus even aan dat de term conversie hier gebruikt wordt in de enge zin, aangegeven door het DAVID-project van het Stadsarchief Antwerpen, en slaat dus terug op de *upgrade* van een document binnen een zelfde bestandsformaat.

De term migratie verenigt een aantal elementen. *The National Preservation Office* stelt de volgende onderverdeling voor¹³:

- *Media wijzigen*: opslag op optische of magnetische media, maar ook op papier of microfilm.
- *Backward compatibility*: dit hangt af van de vraag of de maker van de software ervoor zorgt dat nieuwe versies van de applicatie compatibel is met oudere versies.
- *Interoperabiliteit*: de mogelijkheid om objecten van het ene platform en/of applicatie naar het andere te verplaatsen zonder verlies van informatie, functionaliteit en authenticiteit.
- *Conversie naar standaardformaten*: verplaatsing van bestandsobjecten vanuit het originele applicatiespecifieke formaat naar een open, gedocumenteerd standaardformaat.¹⁴

¹⁰ Aschenbrenner, A. Long-term preservation of digital material – Building an archive to preserve digital cultural heritage. [2001?], http://www.ifs.tuwien.ac.at/~aola/publications/thesis-ando/Conversion_standard.html#foot469, 05/04/06

¹¹ Boudrez, F. & Dekeyser, H. *Idem*, p. 61

¹² Task Force on Archiving of Digital Information, Preserving digital information. Commission on Preservation and Access/ Research Libraries Group, 1996, p. 6, <http://www.rlg.org/legacy/ftpd/pub/archtf/final-report.pdf>, 05/04/06

¹³ Testbed Digitale Bewaring, Migratie: Context en huidige stand van zaken. Den Haag: ICTU, 2001, pp. 9-11 ; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white_paper_migratie.pdf, 05/04/06

De meest gebruikte toepassing van de term, die aansluit zich aan bij de omschrijving dat migratie “changes the structure of the original data file”¹⁵, slaat terug op de omzetting van het originele formaat naar een officiële of *de facto* standaard. Dit is eveneens de meest gebruikte bewaarstrategie op dit moment.¹⁶ Het voordeel van open standaarden is de relatieve stabiliteit van deze vastgelegde formaten. Zelfs wanneer een bepaalde standaard na verloop van tijd wordt vervangen door een andere, zal technische ondersteuning mogelijk blijven.¹⁷ Standaarden nemen echter tijd in beslag om te ontwikkelen, waardoor ze kunnen voorbijgestreefd zijn alvorens geïmplementeerd te kunnen worden. Voorts spelen een aantal zelfde problemen mee als bij andere formaten.: niet alle standaarden zijn even wijdverspreid, ze hebben geen onbeperkte levensduur, ze worden soms *ad hoc* uitgebreid om bijkomende mogelijkheden in te bouwen. Bovendien zijn voor bepaalde bestanden geen adequate archiveringsformaten voor handen.¹⁸

Een andere kritiek slaat terug op de technische problemen die zich kunnen voordoen. Migratie vereist in de meeste gevallen een transformatie van de originele bitstream, waardoor de integriteit en authenticiteit bij elke omzetting in gevaar is. Een zeker verlies in informatie, functionaliteit en voorstelling behoort tot de mogelijkheden, alsook de introductie van fouten of software *bugs*. De nieuwe formaten blijven uiteraard ook niet eeuwig bestaan. Nieuwe omzettingen moeten bijgevolg op geregelde tijdstippen worden uitgevoerd, waardoor deze strategie vrij kostelijk en arbeidsintensief is.¹⁹

¹⁴ Een voorbeeld hiervan is PDF/A (*Portable Document Format/Archive*), ontwikkeld door Adobe in samenwerking met AIIM (<http://www.aiim.org/standards.asp?ID=25013>, 05/04/06), die werd vastgelegd als ISO standaard. ISO 19005-1. Document management – Electronic document file format for long-term preservation – Part 1 : Use of PDF (PDF/A). 2005, [http://www.aiim.org/documents/standards/ISO_19005-1_\(E\).doc](http://www.aiim.org/documents/standards/ISO_19005-1_(E).doc), 05/04/06

¹⁵ Lawrence, G.; Kehoe, W.; Rieger, O.; Walters, W.; Kenney, A. Risk management of digital information. A file format investigation. Washington: Washington: Council on Library and Information Resources, 2000, <http://www.clir.org/pubs/reports/pub93/contents.html>, 05/04/06

¹⁶ Rothenberg, J. Avoiding technological quicksand: Finding a viable technical foundation for digital preservation. Washington: Council on Library and Information Resources, 1998, <http://www.clir.org/pubs/reports/rothenberg/inadequacy.html>, 05/04/06

¹⁷ Omdat de groep die hem heeft ontwikkeld nog actief is, of aan de hand van de broncode die bekend is.

¹⁸ Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief DAVID Project, 2004, p. 61

¹⁹ Hedstrom, M. Digital preservation: problems and prospects. In: Koichi Tabata, Shigeo Sugimoto, Tetsuo Sakaguchi (eds.), *Digital Libraries Journal*, N° 20, March 2001, http://www.dl.slis.tsukuba.ac.jp/DLjournal/No_20/1-hedstrom/1-hedstrom.html, 06/04/06

E. Emulatie

“The central idea of the approach ... is to enable the emulation of obsolete systems on future, unknown systems, so that a digital document's original software can be run in the future despite being obsolete.”²⁰ Niet de originele hard- en software wordt bewaard, maar het vereiste platform zal op een toekomstige computerconfiguratie worden gereconstrueerd. Op deze manier blijven bestanden in hun oorspronkelijk formaat raadpleegbaar. Dit kan op verschillende niveaus worden toegepast, i.e. nabootsing van hardware en besturingssystemen, specifieke software of een combinatie van het vorige.²¹

Testbed Digitale Bewaring definiëert emulatie als volgt: “Een emulator is een programma dat op één computer draait (het ‘hostsysteem’ van de emulator) en [a] ervoor zorgt dat die computer zich als een andere computer gedraagt”, of “[b] daarbij zorg draagt voor de virtuele herschepping van een andere computer (het ‘doelsysteem’ van de emulator).”²² Een verouderd programma zal dan draaien op een verouderde computer, die virtueel wordt herschapen op een nieuwe computer. Het is noodzakelijk dat de structuur van het te emuleren programma goed gedocumenteerd wordt.

Er zijn verschillende manieren om emulatie te implementeren²³:

- *Emulation Virtual Machine* (Jeff Rothenberg):

Drie soorten data worden ingekapseld. De originele bitstream wordt samen met de originele applicatie bewaard. Vervolgens wordt een specificatie voor een emulator geschreven die in de toekomst moet ontwikkeld worden. Hier wordt documentatie over het systeem in

²⁰ Rothenberg, J. Avoiding technological quicksand: Finding a viable technical foundation for digital preservation. Washington: Council on Library and Information Resources, 1998, <http://www.clir.org/pubs/reports/rothenberg/criteria.html#emulation>, 06/04/06

²¹ Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief DAVID Project, 2004, p. 59

²² Testbed Digitale Bewaring, Emulatie: Context en huidige stand van zaken. Den Haag: ICTU, 2003, pp. 17-18

²³ Deze zijn theoretisch gebaseerd op de idee van de Turing Machine. In informatica is een Turing Machine een model van berekening en berekenbaarheid, ontwikkeld door de wiskundige Alan Turing. Het betreft hier een eenvoudig mechanisme dat symbolen manipuleert. Ondanks het eenvoudige karakter kan een Turing Machine aangepast worden om de logica van eender welke computer, die ooit gebouwd zou kunnen worden, te simuleren. Hoewel het technisch mogelijk is ze te bouwen, was dit nooit de bedoeling. De idee van de Turing Machine diende enkel als gedachte experiment rond de limieten van mechanische berekeningen. Turing Machine. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Turing_machine, 06/04/06; Barker-Plummer, D. Turing machines. In: Edward, N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2005 Edition, <http://plato.stanford.edu/archives/spr2005/entries/turing-machine/>, 06/04/06; voor een algemeen overzicht zie ook: Granger, S. Digital preservation and emulation : from theory to practice. Leeds: Cedars, 2001, <http://www.leeds.ac.uk/cedars/pubconf/papers/ichim01SG.html>, 06/04/06

menselijk leesbare vorm aan toegevoegd, waar instructies voor de toegang en het gebruik van het document in vermeld staan.²⁴

- *Universal Virtual Machine* (Raymond Lorie):

Lorie stelt dat de techniek die Rothenberg voorstelt, zich schuldig maakt aan *overkill*, indien we niet het programma trachten te bewaren en enkel denken aan het archiveren van data. Daarom wordt een werkwijze op basis van een Universal Virtual Computer (UVC) voorgesteld (wat een combinatie van migratie en emulatie inhoudt).²⁵ Een UVC functioneert zoals een *virtual machine*²⁶, in die zin dat ze een laag creëert tussen het onderliggende computerplatform en de bovenliggende software. Dit geeft als voordeel dat ze kan uitgewisseld worden tussen verschillende platformen. Een UVC echter, in tegenstelling tot een normale *virtual machine*, wordt ontwikkeld met de bedoeling universeel te zijn. Een platform onafhankelijke laag wordt aangeboden die in principe onveranderlijk is.²⁷ “The main idea consists of archiving a program P [an executable program for a UVC] along with the data file that decodes the data and returns the information to a future client based on a logical view.”²⁸ De UVC een emulator die als een software applicatie gedraaid zal worden op een nog onbekend platform van nog niet bestaande hardware van de toekomst. Aangezien we nu nog niet weten hoe die eruit zullen zien, moet de eigenlijke emulator op dat tijdstip worden geschreven. De UVC vormt een platform waarop programma’s kunnen gedraaid worden die specifiek voor de UVC zijn geschreven. Zo een programma decodeert het bestandsformaat van digitaal object en haalt *element tags* op die specifieke informatie bevatten over de inhoud van de data. Op basis van die gegevens wordt een *Logical Data View* gemaakt dat eruitziet als XML. Het bestand kan vervolgens worden gelezen in een *Logical Data Viewer*, die eveneens in de toekomst moet worden geschreven. Het verschuiven van een gedeelte van de oplossing naar de toekomst toe houdt echter wel een zeker gevaar in, zolang deze methode nog niet volledig is getest.²⁹

- *Migration on request* (Cedars’ Camileon Project):

Emulatie volgens Cedars bestaat uit een drietal elementen. De originele bitstream wordt bewaard. Deze kan uitgevoerd worden op het originele

²⁴ Rothenberg, J. *Idem*, [.../criteria.html#emulation](#), 06/04/06; Rothenberg, J. *Ensuring the longevity of digital information*. Santa Monica: RAND, 1999, pp. 15-17, <http://www.clir.org/pubs/archives/ensuring.pdf>, 06/04/06

²⁵ Lorie, R. A. *A project on preservation of digital data*. In: RLG DigiNews, Vol.5, Nr. 3, June 2001, <http://www.rlg.org/legacy/preserv/diginews/diginews5-3.html#feature2>, 04/04/06

²⁶ Voor meer info, zie: *Virtual Machine*. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Virtual_machine, 06/04/06

²⁷ *Universal Virtual Computer*. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Universal_Virtual_Computer, 06/04/06

²⁸ Lorie, R. *The UVC: a method for preserving digital documents. Proof of concept*. Den Haag: Koninklijke Bibliotheek/Amsterdam: IBM External Communications, 2002, p. V, http://www.kb.nl/hrd/dd/dd_onderzoek/reports/4-uvc.pdf, 06/04/06

²⁹ *Universal Virtual Computer. Idem*

platform. Naar de toekomst toe wordt een emulator gebouwd die zal draaien op een nieuw platform. Door het verstrijken van de tijd zal dit platform verouderen en vervangen worden. De emulator kan dan aangepast worden door middel van een upgrade.³⁰ Het onderzoek dat in dit kader werd uitgevoerd, betrof het emuleren van programma's. Voor data objecten is zo een uitgebreide procedure niet altijd werkbaar. Om de bekomen ervaringen hiervoor te kunnen herbruiken, werd het concept van *migration on request* uitgewerkt. "Klassieke" migratie is, zoals al eerder aangehaald, zeer kostelijk en tijdsrovend. Camileon maakt gebruik van specifieke migratietools. De oorspronkelijke bitstream wordt bewaard. Om deze uitvoerbaar te maken wordt een welbepaalde tool ontwikkeld, die het digitale object migreert naar een "hedendaags", leesbaar formaat op het moment dat het wordt opgevraagd.³¹

De voornaamste voordelen van emulatie als bewaarmethode houden verband met het feit dat digitale bescheiden, theoretisch, in het oorspronkelijke formaat kunnen gearchiveerd worden (eigenschappen en functionaliteiten blijven behouden, geen wijzigingen ten gevolge van omzettingen, ...). Hiernaast wordt de kostprijs (uitgedrukt in geld en arbeid) niet bepaald door de documenten zelf, maar op basis van het systeem dat onderhouden wordt. Boudrez vat de nadelen als volgt samen³²:

- *Technisch complex*: de meeste archieven zijn technisch niet voldoende onderbouwd om dergelijke systemen te bouwen en onderhouden, en zijn bijgevolg afhankelijk van externe diensten en partners.
- *Kosten*: niet alle archieven kunnen de hoge ontwikkelings- en onderhoudskosten (blijven) dragen.
- *Overkill*: sommige benaderingen gaan uit van een volledige reconstructie van applicaties, terwijl het vaak al voldoende is een viewer te programmeren om het archiefstuk weer te geven.
- *Auteursrecht*: de bescherming op auteursrecht houdt beperkingen in. Bovendien is het emuleren op basis van *reverse engineering* voor gesloten, niet-gedocumenteerde formaten in het beste geval zeer risicovol, waardoor migratie naar een open standaard veelal aan te raden valt.
- *Verscheidenheid*: archieven moeten dan tal van verschillende systemen onderhouden, waarvan een aantal *ad hoc* aanpassingen zijn binnen het kader van de eigen organisatie. Hierdoor kunnen kosten niet gedeeld worden met andere archieven.

Hoewel emulatie nieuwe mogelijkheden schept voor de toekomst, moeten de verschillende technieken hun nut nog bewijzen. Enige voorzichtigheid is dus aangewezen.

³⁰ Holdsworth, D. & Wheatley, P. *Emulation, preservation and abstraction*. Leeds: University of Leeds CAMiLEON Project, 2001, <http://cedarsgw.leeds.ac.uk/CAMiLEON//dh/ep5.html>, 06/12/05

³¹ Mellor, P.; Wheatley, P.; Sergeant, D. *Migration on request, a practical technique for preservation*. Leeds: University of Leeds CAMiLEON Project, 2002, pp. 1-4 <http://www.si.umich.edu/CAMiLEON/reports/migreq.pdf>, 06/04/06

³² Boudrez, F. & Dekeyser, H. *Digitaal archiefbeheer in de praktijk*. Antwerpen: Stadsarchief DAVID Project, 2004, p. 59-60

2.2 Het digitaal depot

A. Institutioneel depot

Door de sterke toename van (digitale) informatie tijdens de jaren negentig ontstond de nood deze gegevens adequaat te beheren. Nieuwe ideeën werden ontwikkeld met de bedoeling kennis te bevorderen en efficiënter te gebruiken. Aanvankelijk werden ze gegroepeerd onder de noemer *information management*, wat later *content management* werd, en nu *knowledge management* wordt genoemd. Bedrijven en organisaties introduceerden *Electronic Document en Record Management Systems* (EDRMS) in het werkveld. Binnen deze context ontstond het concept van een institutioneel depot (*institutional repository*) in de academische wereld (ontwikkeld voor het beheer van de intellectuele *output*). In de beginfase (late jaren negentig) werd software ontwikkeld voor en door de eigen organisatie (cfr. CERN¹ en Humboldt-Universität zu Berlin²). Systemen moesten geprogrammeerd worden door professionele computerspecialisten omdat op dat moment nog niets voor handen was. Omwille van de hoge kosten die hiermee gepaard gaan, konden enkel grote instellingen zich dit veroorloven. Pas tijdens een later stadium werden algemene systemen met aanpasbare pakketten op de markt gebracht.³ Een groot aantal instellingen werkt ondertussen met deze pakketten.⁴

Aanvankelijk werd zeer veel aandacht besteed aan de eigenlijke technologische problematiek binnen het kader van een aantal projecten (die zich aansloten bij de *Open Access* beweging⁵). Na het verschijnen van de

¹ Conseil Européen pour la Recherche Nucléaire (CERN), voor algemene informatie zie www.cern.ch, voor publicaties : <http://cdsweb.cern.ch/>

² <http://www.hu-berlin.de/>

³ Onder meer: OPUS (<http://elib.uni-stuttgart.de/opus/>), GNU Eprints (<http://www.eprints.org/>), FEDORA (<http://www.fedora.info>) en DSpace (<http://www.dspace.org>). Aschenbrenner, A. & Kaiser, M. White paper on digital repositories. reUSE! Digital Master Files of Printed Materials, March 2005, pp. 9-10, http://www.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf, 12/04/06

⁴ Voor een uitgebreid overzicht zie: the Directory of Open Access Repositories (DOAR), <http://www.opendoar.org/> en the Institutional Archives Registry, <http://archives.eprints.org/>.

⁵ The Open Access Initiative stuurt aan op vrije toegang van wetenschappelijk onderzoek. “De open access-beweging wil actuele informatie gratis online beschikbaar maken. De algemene principes van open access vragen dat de auteur de onherroepelijke toestemming geeft aan iedereen om een werkstuk te lezen, te downloaden, te kopiëren en te herverdelen. De auteur moet altijd vermeld worden. Aanpassingen en massale commerciële verdeling vallen hier meestal buiten. Door het onveranderlijk karakter en de auteursvermelding van de publicatie, wordt de verantwoordelijkheid van de auteur voor zijn artikel gegarandeerd. Dit is belangrijk voor wetenschappelijke publicaties. Het doel van open access is de verspreiding en feedback van wetenschappelijke literatuur verbeteren.”

eerste *best practices*, verplaatste de focus zich naar problemen op het niveau van beheer. Een studie in opdracht van het *Joint Information Systems Committee* (JISC) in Engeland, toonde aan dat de meeste problemen voor de preservatie van, doorgaans statische, tekstuele, inhoud van institutionele depots, eerder organisatorisch dan technologisch van aard zijn.⁶ Meerdere initiatieven werden ondernomen om meer inzicht in deze problematiek te bekomen.⁷ De algemene trend die hierin zichtbaar is, is een grotere samenwerking met de nationale bibliotheken en tussen de verschillende instellingen op nationaal en internationaal vlak.⁸

B. Betrouwbaar digitaal depot

De idee van een *trusted digital repository* (betrouwbaar digitaal depot) werd omschreven in een rapport van OCLC⁹ en RLG¹⁰ en gedefinieerd als “one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future.” Om deze taak uit voeren moet bij de gebruiker een zeker vertrouwen kunnen opgewekt worden. De voornaamste gebruikers van een archief zijn uiteraard de leden van de eigen organisatie in de eerste instantie, en de gemeenschap (omwille van transparant bestuur en/of historisch onderzoek) in de tweede. De eigen organisatie wil degelijk bewaarde bronnen om ze te kunnen gebruiken bij de uitvoering van de eigenlijke job of als bewijskracht van die uitvoering. De gemeenschap kan enkel zichzelf kennen aan de hand van zijn geschiedenis (bekomen dus op basis van de bronnen die voor handen zijn)

<http://wbib.kuleuven.be/?node=227>, 26/04/06, Zie onder andere: Open Society Institute: Budapest Open Access Initiative, <http://www.soros.org/openaccess/>, Open Archives Initiative (OAI [niet te verwarren met OAIS]), <http://www.openarchives.org/>, Bailey, C.B., Jr. *Scholarly electronic publishing bibliography*. <http://epress.lib.uh.edu/sep/sepb.html>

⁶ James, H.; Ruusalepp, R.; Anderson, S.; Pinfield, S. *Feasibility and requirements study on preservation of E-Prints*. Report commissioned by the Joint Information Systems Committee (JISC). October 2003, http://www.jisc.ac.uk/uploaded_documents/e-prints_report_final.pdf, 26/04/06

⁷ Zie onder meer: Project DAVID (Digitale Archivering in Vlaamse Instellingen en Diensten), <http://www.expertisecentrumdavid.be/>, Digitale Duurzaamheid, <http://www.digitaleduurzaamheid.nl/home.cfm>, Curl Exemplars in Digital Archives (CEDARS), <http://www.leeds.ac.uk/cedars/>, Records Continuum Research Group, <http://www.sims.monash.edu.au/research/rcrg/>, International Research on Permanent Authentic Records in Electronic Systems (InterPARES), <http://www.interpares.org/>, an Effective Strategic model for the Preservation and disposal of Institutional Digital Assets (eSPIDA), <http://www.gla.ac.uk/espida/>, SHERPA DP: Creating a persistent preservation environment for Institutional repositories, <http://ahds.ac.uk/about/projects/sherpa-dp/>, Digital Academic Repositories (DARE), <http://www.darenet.nl>, Digitala Vetenskapliga Arkivet (DiVA), <http://www.diva-portal.net/>, Deutsche Initiative für Netzwerkinformation (DINI), <http://www.dini.de/>, the National Digital Information Infrastructure and Preservation Program (NDIIP), <http://www.digitalpreservation.gov/>, Victorian Electronic Records Strategy (VERS), www.prov.vic.gov.au/vers/

⁸ Aschenbrenner, A. & Kaiser, M. *White paper on digital repositories*. reUSE! Digital Master Files of Printed Materials, March 2005, pp. 12-13, http://www.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf, 12/04/06

⁹ Online Computer Library Center, <http://www.oclc.org/>

¹⁰ Research Libraries Group, <http://www.rlg.org>

zouden we enigszins naïef kunnen stellen. In beide gevallen wordt van de opslagplaats (en dus van de archivaris) verwacht dat de opgevraagde informatie betrouwbaar is. De traditionele archieven hebben dit vertrouwen opgebouwd door jaren (zelfs eeuwen) van ervaring. In de vluchtige, digitale wereld hebben we, ironisch genoeg, die luxe van tijd niet. We moeten nu, onmiddellijk betrouwbare documenten kunnen voorleggen indien we het verleden in de toekomst willen veilig stellen (en nee dat is geen *contradictio in termines*). Na verloop van tijd kunnen de archieven enkel betrouwbaarder worden.

De karakteristieken van een digitaal archiefstuk die werden vastgelegd in ISO 15489 kunnen ons hierbij helpen: authenticiteit, betrouwbaarheid, integriteit, herbruikbaarheid (en hierbij samenhangend visualiseerbaarheid en uitvoerbaarheid). Bewaring in een digitaal archief is dus meer dan enkel het bijhouden van de bitstream. De doelgroep (*designated community*, zie verder) moet erop kunnen vertrouwen dat het depot in staat is om de voorgaande elementen te waarborgen. Om deze taak te kunnen uitoefenen naar behoren, moeten aan een aantal voorwaarden worden voldaan (aldus het RLG/OCLC rapport).¹¹ Elk digitaal depot moet:

- Verantwoordelijkheid opnemen voor het onderhoud van de opgenomen digitale bronnen op de lange termijn
- Een organisatorische structuur hebben die de uitvoerbaarheid van zowel het depot als van de opgenomen informatie ondersteunt
- Fiscale verantwoordelijkheid en duurzaamheid kunnen aantonen
- Een systeem ontwerpen in overeenstemming met algemeen aanvaarde conventies en standaarden
- Methodologieën opstellen voor systeemevaluatie die in overeenstemming zijn met de verwachtingen van de gemeenschap ten opzichte van betrouwbaarheid
- Een transparant beleid kunnen voorleggen

Kort samengevat, het concept van een betrouwbaar e-depot wordt opgebouwd rond twee basisvereisten¹²:

1. De bewaarplaats, met de samenhangende beleidsmaatregelen, standaarden en technologische infrastructuur, verzorgt het kader waarbinnen aan digitale preserving wordt gedaan
2. Het e-depot is een betrouwbaar systeem, in die zin dat het een systeem betreft van soft- en hardware waarop vertrouwd kan worden om aan bepaalde regels te voldoen

¹¹ RLG/OCLR, Trusted digital repositories. Attributes and responsibilities. An RLG-OCLR Report. Mountain View: RLG, 2002, p. 5; <http://www.rlg.org/en/pdfs/repositories.pdf>, 15/03/06

¹² Jantz, R. & Giarlo, M. Digital preservation. Architecture and technology for trusted digital repositories. In: *D-Lib Magazine*, Volume 11, Number 6, June 2005, <http://www.dlib.org/dlib/june05/jantz/06jantz.html>, 12/04/06

RLG/OCLC stelt een aantal attributen voor die een betrouwbaar digitaal depot opmaken¹³:

- De regels naleven zoals ze beschreven zijn in *The reference model for an Open Archival Information System (OAIS)*: Het OAIS zorgt voor een functioneel en een informatie model. Door deze normen na te leven, kan een waardering gemaakt worden doordat met een vaste set van concepten en terminologie wordt gewerkt. Hier kom ik dadelijk nog uitgebreider op terug.
- *Administratieve verantwoordelijkheid*: de organisatie expliciteert haar toewijding voor de ontwikkeling van een betrouwbaar e-depot, in overeenstemming met de gangbare standaarden, procedures en gebruiken. Het opstellen van een rampenplan mag hierbij niet vergeten worden. Een betrouwbaar depot zal steeds in alle transparantie werken.
- *Organisatorische levensvatbaarheid (organizational viability)*: de organisatie kan aantonen dat haar missie, wettelijk statuut en werkingsmiddelen voldoende zijn om een betrouwbaar e-depot blijvend te onderhouden.
- *Financiële duurzaamheid*: de organisatie kan aantonen dat ze in een sterke financiële basis kan voorzien.
- *Technologische en procedurele geschiktheid*: op een open, weldoordachte manier zullen preserveringsstrategieën worden opgesteld en geïmplementeerd, in samenspraak met het management, de ICT-afdeling en de archiefdienst.
- *Systeembeveiliging*: alle gebruikte systemen zullen zo ontworpen worden dat de veiligheid van de digitale bescheiden verzekerd kan worden. Hierbij worden bepaalde procedures en beleidsmaatregelen vastgelegd die het gebruik regelen. Verder wordt speciale aandacht besteed aan processen die verlies van gegevens detecteren, vermijden en herstellen.
- *Procedurele verantwoordingsplicht*: de organisatie verzorgt op een systematische manier een uitgebreide documentatie van alle digitale preserveringsacties die werden (en zullen worden) ondernomen.

C. Open Archival Information System (OAIS)

Het *Open Archival Information System (OAIS)* model werd ontwikkeld door NASA's Consultative Committee for Space Data Systems¹⁴, oorspronkelijk voor data bekomen uit observaties van de Aarde en de ruimte. Het OAIS is een conceptueel referentiemodel, dat een kader biedt waarbinnen procedures voor de lange termijn archivering van digitale informatie wordt uitgewerkt.¹⁵

¹³ RLG/OCLC, Trusted digital repositories. Attributes and responsibilities. An RLG-OCLC Report. Mountain View: RLG, 2002, pp. 13-15; zie ook: [Digital preservation management: Implementing short-term strategies for long-term problems](http://www.library.cornell.edu/iris/tutorial/dpm/foundation/tdr/index.html), 20/04/06

¹⁴ En vastgelegd als ISO standaard in 2002: ISO-14721:2002

¹⁵ Boudrez, F. & Dekeyser, H. *Digitaal archiefbeheer in de praktijk*. Antwerpen: Stadsarchief DAVID Project, 2004, pp. 71-72

Het model wordt gedefinieerd als “an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community.”¹⁶ Een OAIS archief heeft als taak informatie van archiefvormers te behandelen en bewaren. Het is noodzakelijk om voldoende rechten op controle van de beheerde inhoud te verkrijgen, opdat lange termijn preservatie kan verzekerd worden.¹⁷ Na de opname moet de kwaliteit van het document verzekerd worden, in samenspraak met de *designated community*, die deze kunnen begrijpen zonder externe hulp. Vervolgens moeten de opgenomen documenten op zo een manier beheerd worden dat ze beschermd zullen zijn tegen alle mogelijke voorvallen, waardoor hun authenticiteit verzekerd blijft. Ten laatste, moet de toegang tot de bewaarde informatie¹⁸ mogelijk worden gemaakt voor de gebruiker.¹⁹

Een OAIS archief bestaat uit drie gescheiden maar gerelateerde delen. Het eerste deel beschrijft de externe omgeving waarbinnen een OAIS werkt. Het tweede deel omschrijft de functionele componenten (of interne mechanismen) die de verantwoordelijkheden inzake preservatie vervullen. In het derde deel worden de informatie objecten beschreven die ingevoegd, beheerd en ter beschikking worden gesteld door het OAIS.²⁰

Binnen de OAIS-omgeving zijn eveneens een aantal externe **stakeholders** van tel. Het referentiemodel identificeert en beschrijft deze belanghebbenden én bepaaldt de *interface* tussen hen onderling en met het OAIS. Drie soorten entiteiten worden onderscheiden:

- *Management*: verantwoordelijk voor, onder meer, het formuleren, herzien en soms opleggen van het beleidskader op het hoogste niveau die de OAIS activiteiten bepalen.
- *Producent(en) (producers)*: d.i. de archiefvormers die de te bewaren informatie en de bijbehorende metadata overhandigen.
- *Gebruikers (consumers)*: d.i. degenen die gebruik maken van het archief.

¹⁶ Consultative Committee for Space Data Systems, Reference model for an Open Archival Information System (OAIS). Blue Book. Washington: CCSDS, National Aeronautics and Space Administration, 2002, p. 1-1

<http://public.ccsds.org/publications/archive/650x0b1.pdf>, 08/03/06

¹⁷ “The problems of assuming sufficient control ... are addressed in three related categories ...: copyright implications, intellectual property and other legal restrictions in use; authority to modify Representation information; agreements with external organizations [to assist in its preservation objectives].” Consultative Committee for Space Data Systems, *Ibidem*, pp. 3-2 – 3-3

¹⁸ Informatie wordt gedefinieerd als “any type of knowledge that can be exchanged, and this information is always expressed (i.e., represented) by some type of data.” Consultative Committee for Space Data Systems, *Ibidem*, p. 2-3

¹⁹ Aschenbrenner, A. Long-term preservation of digital material – Building an archive to preserve digital cultural heritage from the internet. [2002?], <http://www.ifs.tuwien.ac.at/~aola/publications/thesis-ando/OAIS.html>, 08/03/06

²⁰ Lavoie, B. The Open Archival Information System Reference Model: Introductory guide. DPC Technology Watch Series Report 04-01, Dublin: OCLC, Inc., 2004, pp. 4-5, http://www.dpconline.org/docs/lavoie_OAIS.pdf 08/03/06

Een speciale groep gebruikers wordt omschreven met de term **Designated Community**. Van hen wordt verwacht dat ze de gearchiveerde informatie kunnen begrijpen in de vorm waarin ze werd bewaard, zonder bijkomende hulp, omwille van hun gespecialiseerde kennis. Het werkingsgebied van deze *designated community* bepaaldt de inhoud en de vorm in dewelke de informatie zal worden bewaard.

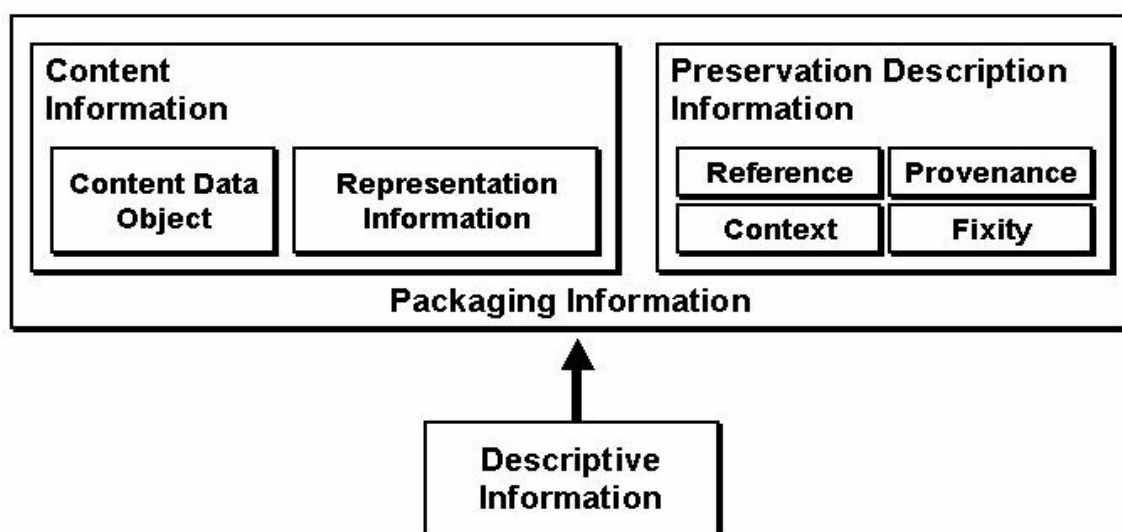
De bovenstaande concepten verwoorden eerder functionele, dan organisationele rollen. Essentieel is de logische opdeling van het beleidsproces en de belangen van de *stakeholders*.²¹

Het OAIS informatie model is opgebouwd rond een **information package (IP)**. Het betreft hier een conceptualisatie van de structuur van de informatie en haar bewegingen in en uit het archiefsysteem. Ze bestaat het gearchiveerde digitale object en haar metadata in één logisch geheel (of pakket). Drie varianten zijn mogelijk²²:

- *Submission Information Package (SIP)*: via een opname proces wordt de informatie overgedragen van de producent naar het OAIS. De exacte vorm wordt bepaald door onderlinge afspraken tussen de beiden. Tegelijkertijd wordt de nadruk gelegd op het feit dat de vorm waarin zal bewaard worden kan veranderen.
- *Archival Information Package (AIP)*: dit is de versie van het IP die opgeslagen zal worden. In één logische eenheid worden informatie en metadata bewaard. De effectieve vorm kan veranderd worden op basis van archivistische vereisten.
- *Dissemination Information Package (DIP)*: deze versie krijgt de consument te zien. Het is mogelijk dat de vorm en inhoud niet dezelfde zijn als die van het gearchiveerde object (bvb. formaat, hoeveelheid van de inhoud of metadata, ...).

²¹ Lavoie, B. The Open Archival Information System Reference Model: Introductory guide. DPC Technology Watch Series Report 04-01, Dublin: OCLC, Inc., 2004, pp. 5-7

²² Lavoie, B. *Idem*, pp. 10-11



Archival Information Package

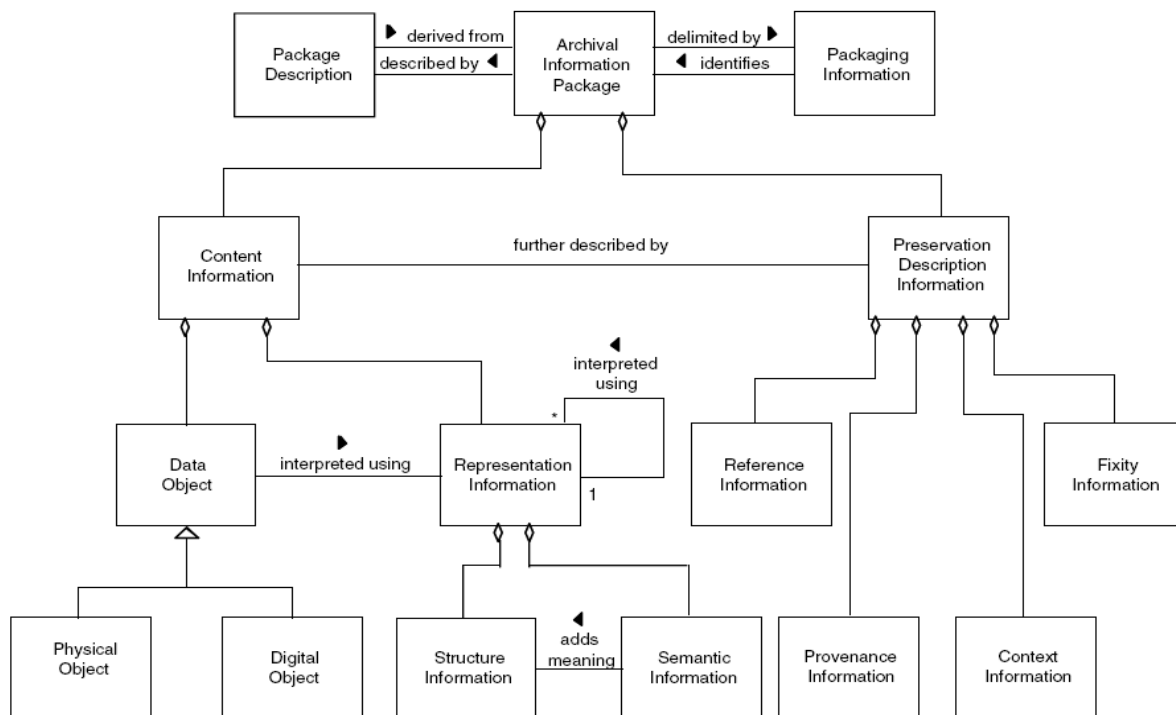
Bron: Lavoie, B. The Open Archival Information System Reference Model: Introductory guide. DPC Technology Watch Series Report 04-01, Dublin: OCLC, Inc., 2004, pp. 12, http://www.dpconline.org/docs/lavoie_OAIS.pdf 08/03/06

De nadruk ligt op het **AIP**, d.i. het archiefstuk. De essentiële elementen zijn²³:

- *Packaging Information*:
 - *Content Information*:
 - *Content Data Object*: de informatie die bewaard moet blijven. Het kan gaan om een enkel, onafhankelijk (*self contained*) bestand of om meerdere samenhangende bestanden. Het OAIS is verantwoordelijk voor de preservering van het content data object op de lange termijn en de ter beschikking stelling in een vorm die begrijpbaar is voor de designated community.
 - *Representation Information*: d.i. informatie noodzakelijk om de bitstream verstaanbaar en uitvoerbaar te maken. Dit kan bijvoorbeeld een beschrijving van de hard- en software zijn of een samenvatting van de juiste interpretatie van het Content Data Object. Het werkingsgebied van de Designated Community bepaaldt hoe uitgebreid de Representation Information zal zijn.
 - *Preservation Descriptive Information (PDI)*: d.i. bijkomende metadata ter ondersteuning en documentatie van het preservingsproces van een OAIS. Deze bestaat uit vier elementen:

²³ Consultative Committee for Space Data Systems, Reference model for an Open Archival Information System (OAIS). Blue Book. Washington: CCSDS, National Aeronautics and Space Administration, 2002, p. 2-3 – 2-7
<http://public.ccsds.org/publications/archive/650x0b1.pdf> , 08/03/06

- *Reference Information*: identificeert de Content Information zowel binnen de interne systemen, als ten opzichte van de externe systemen en entiteiten.
- *Context Information*: beschrijft de verhouding van Content Information ten opzichte van andere Content Information Objects (bvb. dezelfde informatie opgeslagen in verschillende formaten).
- *Provenance Information*: documenteert de geschiedenis van de Content Information, inclusief het ontstaan, veranderingen qua inhoud of formaat, acties ter preservering en overdracht van beheer.
- *Fixity Information*: valideert de authenticiteit en/of integriteit van de Content Information (bvb. check sums, digitale handtekening, digitaal watermerk, ...).
- *Descriptive Information*: ondersteunt het terugvinden en ophalingen van Content information door de Designated Community via de OAIS zoekfuncties (bvb. op basis van de Dublin Core Metadata set²⁴).



Archival Information Package (gedetailleerd)

Bron: Consultative Committee for Space Data Systems, Reference model for an Open Archival Information System (OAIS). Blue Book. Washington: CCSDS, National Aeronautics and Space Administration, 2002, p. 4-37

Verder worden de basis functies van een archief omschreven. Deze omvatten onder meer de manier en tijdstip van opname en ter beschikking stellen, opstellen van een rampenplan, een afbakening van de *designated community*, een afbakening van verantwoordelijkheden, toegangsbeperkingen en –controles, ...

²⁴ Dublin Core Metadata Initiative, <http://dublincore.org/>

2.3 Standaarden

Zoals hierboven aangehaald, moeten archieven een betrouwbaar karakter uitstralen. Hierom werden bepaalde concepten uitgewerkt, zoals beschreven in het *Trusted digital repositories* en het OAIS rapport. Maar, omdat het moeilijk is appels met peren te vergelijken, werd (hieraan voorafgaand én volgend) onderzoek gedaan met het oog op de ontwikkeling van een aantal standaarden.¹ Hoe meer organisaties deze implementeren, hoe gemakkelijker het is na te gaan in hoeverre de bewaring van de bescheiden verzekerd is. Het doel is uiteraard te zorgen voor een kwaliteitslabel en de uniformiteit te bevorderen.

Verschillende niveaus bestaan binnen de groep van standaarden:

- Archiveringsstandaarden
 - Documentbeheerssystemen en Records Management Systemen (vb. ISO 15489)
 - Beschrijvingsstandaarden (ISAD)
 - Archiefschema's (EAD)
 - Metadatastandaarden (vb. Dublin Core)
- Technische standaarden
 - Architectuur van systemen (vb. WMFC, W3C)
 - Code tabellen (vb. ASCII)
 - Bestandsformaten (vb. PDF/A)
 - Dragere (vb. ISO 9660)

Het gebruik van standaarden heeft ook zijn nadelen. De ontwikkeling ervan neemt tijd in beslag, met het risico voorbijgestreefd te zijn alvorens op de markt te worden gebracht. Bovendien is de term *lange termijn* in de wereld van computers een relatief begrip. Technologie kan (en zal naar alle waarschijnlijkheid) voorbijgestreefd worden, waardoor aanpassingen en/of omschakelingen naar nieuwe systemen en/of formaten vereist zijn. Omwille van het jeugdige karakter van IT in het algemeen en de standaarden in het bijzonder, moeten de merites van deze laatste nog bewezen worden. Zoals bij alles zijn er voor- en nadelen, maar we moeten ergens beginnen.

¹ Deze rapporten kunnen eveneens als standaarden worden beschouwd, maar omwille van het lezersgemak behandel ik ze apart.

A. Electronic Document and Records Management Systems

ISO 15489

De *International Organization for Standardization* (ISO)² is, zoals de naam al aangeeft, een internationaal orgaan, samengesteld uit nationale standaardontwikkende organisaties, dat technische standaarden ontwikkelt in samenspraak met de sectoren voor dewelke ze bedoeld zijn.³ “ISO standards specify the requirements for state-of-the-art products, services, processes, materials and systems, and for good conformity assessment, managerial and organizational practice.”⁴

*ISO 15489: 2001 (E), Information and documentation – Records Management*⁵ werd specifiek ontwikkeld voor de archiefwereld met als doel het dynamisch archiefbeheer te verbeteren. De oorsprong van deze norm ligt in Australië. In 1990 werd door *Standards Australia* het comité IT/21 opgericht, dat de opdracht kreeg te werken aan de standaardisatie van het *records management*. Dit leverde in 1996 de Australische standaard AS 4390 op. Vervolgens werd het ISO gecontacteerd om er een internationale standaard van te maken. Na een aantal aanpassingen werd dit een feit in 2001 als ISO 15489.⁶ Binnen de archiefwereld werd het nut van deze norm snel erkend. Niet lang erna werd een Nederlandse vertaling⁷ op de markt gebracht door het Nederlandse Normalisatie-instituut (NEN).⁸

De norm bestaat uit twee delen. “Deel 1: Algemeen” verzorgt een conceptueel kader voor archiefbeheer. De nadruk wordt gelegd op de voordelen van records management en het belang van een duidelijke verdeling van verantwoordelijkheden inzake archiefbeheer. Een duidelijke omschrijving wordt gegeven van de vereiste elementen bij het opstellen van een Records Management Systeem. “Deel 2: Richtlijnen” zorgt voor een meer gedetailleerde begeleiding bij de implementatie van het kader zoals het

² <http://www.iso.org/>

³ ISO is trouwens geen acroniem, maar afgeleid van het Griekse woord *ισος* (*isos*), wat “gelijk” betekent. [International Organization for Standardization](http://en.wikipedia.org/wiki/ISO). In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/ISO>, 04/05/06

⁴ International Organization for Standardization, *ISO in brief*. Genève: ISO, 2005, <http://www.iso.org/iso/en/aboutiso/isoinbrief/isoinbrief.html>, 04/05/06

⁵ ISO 15489-1: 2001 (E), Information and documentation – Records Management –Part 1: General en ISO/TR 15489 –2: 2001 (E), Information and documentation – Records Management –Part 2: Guidelines

⁶ Cumming, K. *Evolution and requirements of ISO 15489*. 24 June 2003, pp. 1-2 <http://www.archives.govt.nz/continuum/dls/pdfs/katec-03-10.pdf>, 04/05/06

⁷ NEN-ISO 15489-1 (nl), Informatie en documentatie – Informatie- en archiefmanagement – Deel 1: Algemeen (ISO 15489-1: 2001, IDT), Delft: Nederlands Normalisatie-instituut, 2001; en NPR-ISO/TR 15489-2 (nl), Informatie en documentatie – Archiefbeheer – Deel 2: Richtlijnen (ISO/TR 15489-2: 2001), Delft: Nederlands Normalisatie-instituut, 2002

⁸ www.nen.nl

geschetst wordt in deel 1. Zo wordt bijvoorbeeld de DIRS methodologie uit de doeken gedaan.⁹

DIRS of DIRKS staat voor *Designing and Implementing Recordkeeping Systems*. Deze methodologie vindt haar oorsprong in AS 4390-1996 en werd bijgevolg geïncorporeerd in de ISO-norm die haar heeft opgevolgd. In *The DIRKS manual: a strategic approach to managing business information*¹⁰ wordt de ontwikkeling van een records management systeem beschreven in acht stappen. Het betreft hier geen controlelijst waarbij na het uitvoeren van elk onderdeel een vakje moet worden aangevinkt, maar eerder een theoretisch kader dat bij implementatie aangepast moet worden aan de specifieke noden van de eigen organisatie.

Model Requirements for the Management of Electronic Records (MoReq)

Tijdens het Europese DLM-Forum¹¹ van 1996 werd een eerste maal gesproken over de nood aan een uitgebreide specificatie van vereisten voor records management applicaties. De ontwikkeling van een model werd aan Cornwall Management Consultants uitbesteed onder auspiciën van het Interchange of Data between Administrations (IDA) programma van de Europese Commissie. Dit resulteerde in *Model Requirements for the Management of Electronic Records*, een model waarin de basis vereisten worden beschreven voor een EDRMS (zowel voor een privé onderneming als een overheidsinstelling). Het doel is een praktisch hulpmiddel aan te reiken dat een organisatie helpt haar (papieren en elektronisch) documentbeheer efficiënt te beheren. Het kan gebruikt worden door (potentiële) EDRMS gebruikers en organisaties die training geven als referentie, door academische instellingen als een onderwijsmiddel, door EDRMS leveranciers als richtlijn bij productontwikkeling, door record management service providers als richtlijn bij dienstenontwikkeling. Het is echter niet de bedoeling om voor een concrete checklist te zorgen die blindelings moet gevolgd worden. Het model richt zich op archief- en documentbeheer in de brede zin van het woord. Elk land (en elke organisatie) heeft zijn eigen gewoontes op dit vlak, waaraan de implementatie zal moeten worden aangepast.¹²

⁹ Bradley, M. *ISO 15489: A practical guide*. In: *eGov Monitor*, 16 June 2005, <http://www.egovmonitor.com/node/1401>

¹⁰ The National Archives of Australia, *The DIRKS Manual: A strategic approach to managing business information*. Commonwealth of Australia, 2001 (revised 2003), <http://www.naa.gov.au/recordkeeping/dirks/dirksman/dirks.html>, 06/11/05

¹¹ DLM staat voor *Données Lisibles par Machine*, later verandert naar *Document Lifecycle Management*. Het DLM Forum is een vereniging van organisaties uit de publieke sector, het onderzoeksveld en de private sector, waaronder de 25 leden van de Europese Unie en de Europese Commissie. Deze vereniging verricht onderzoek naar en verzorgt ondersteuning voor document lifecycle management in de brede zin. <http://www.dlm-network.org/>, 30/06/06

¹² IDA, *Model Requirements for the Management of Electronic Records. MoReq Specification*. Prepared by Cornwell Management Consultants plc, October 2002, pp. 1-3, <http://www.cornwell.co.uk/moreq.html>, 06/11/05

De focus ligt op de ontwikkeling van een specificatie van een EDRMS (van software dus) ter ondersteuning van het records management schema (met behulp van onder meer ISO 15489) opgesteld door de organisatie in kwestie.¹³ MoReq is onderverdeeld in hoofdstukken aangaande classificatieschema's, controlefuncties en beveiliging, selectie en vernietiging, zoekfuncties, workflow, ... Een sjabloon wordt voorzien waarin eerst elke vereiste wordt gedefinieerd en vervolgens elke functie gedetailleerd wordt omschreven. Omwille van de brede opvatting van het model wordt een onderscheid gemaakt tussen verplichte (*must*) en wenselijke (*should*) vereisten.¹⁴

Ondertussen zijn vertalingen beschikbaar in het Spaans, Tsjechisch, Frans, Hongaars, Italiaans, Portugees, Russisch en Sloveens.¹⁵ De Archiefschool Nederland heeft een eigen vertaling op de markt gebracht, met specifieke aanpassingen gericht op de eigen organisaties en instellingen, onder de titel *Softwarespecificaties voor Records Management Applicaties voor de Nederlandse Overheid (ReMANO)*.¹⁶

DoD 5015.2-STD

In 1997 werd door het Amerikaanse Department of Defense een set van basisvereisten voor records management application software uitgegeven voor intern gebruik, onder de titel *Design Criteria Standard for Electronic Records Management Software Applications*, beter gekend als *DoD 5015.2* (met een herziening in 2002).¹⁷ Vrij snel werd dit model overgenomen door andere overheidsdiensten, waardoor het een *de facto* standaard werd in de Verenigde Staten. Bovendien hebben de beschreven vereisten in meerdere of mindere mate een basis verzorgd voor andere standaarden zoals PRO¹⁸, MoReq en ReMANO. In de privé-sector bezorgt het DoD-certificaat een RMA een kwaliteitslabel.¹⁹ Wanneer een programma in overeenstemming is met de

¹³ Waldron, M. MoReq: Model Requirements for the Management of Electronic Records. In: B. Duhon (ed.), *AIIM E-Doc Magazine*, September/ October 2002, http://www.edocmagazine.com/article_new.asp?ID=25504, 30/06/06

¹⁴ Cain, P. MoReq: The standard of the future? In: *Information Management Journal*, March/April 2003, p. 2, http://www.findarticles.com/p/articles/mi_qa3937/is_200303/ai_n9184315/pg_2, 30/06/06

¹⁵ <http://www.cornwell.co.uk/moreq.html>

¹⁶ Horsman, P.J.; van Bussel, G.J. & Waalwijk, H. Softwarespecificaties voor Records Management Applicaties voor de Nederlandse Overheid. Remano 2004. Amsterdam: Archiefschool, 2004, http://www.archiefschool.nl/docs/ReMANO_2004.pdf

¹⁷ Department of Defense, DoD 5015.2-STD. Design criteria standards for electronic records management. Washington: Department of Defense, Washington Headquarters Services, 2002, http://www.dtic.mil/whs/directives/corres/pdf/50152std_061902/p50152s.pdf, 09/03/06

¹⁸ D.i. de vroegere Public Record Office van het VK (www.pro.gov.uk), nu simpelweg the National Archives (www.nationalarchives.gov.uk).

¹⁹ Gable, J. Everything you wanted to know about DoD 5015.2. In: *Information Management Journal*, November/ December 2002, http://www.findarticles.com/p/articles/mi_qa3937/is_200211/ai_n9164040, 30/06/06

aangegeven basisvereisten, kan het Defense Information Systems Agency (DISA) Joint Interoperability Test Command (JITC) het voornoemde certificaat uitreiken. Los hiervan kunnen nog extra functionaliteiten worden toegevoegd, maar hierover spreekt het JITC zich niet uit.²⁰

De DoD 5015.2 standaard “sets forth mandatory baseline functional requirements, and identifies non-mandatory features deemed desirable for Records Management Application software.”²¹ Er wordt een beschrijving gegeven van de verschillende vereisten. Vervolgens worden deze verder opgesplitst, waarbij wordt aangegeven welke onderdelen verplicht en welke onderdelen wenselijk zijn.

Het was nooit de bedoeling dat de standaard een maatstaf werd voor commerciële producten. Een DoD certificatie vertelt ons dat een product voldoet aan de eisen opgelegd door de *design criteria* na controle van het JITC. Het vertelt ons echter niets over de gebruikersvriendelijkheid of de elegantie van de onderliggende code. Bovendien is interoperabiliteit geen criterium dat in het model werd opgenomen. Dit heeft te maken met het feit dat de tekst werd ontwikkeld toen digitaal archiefbeheer nog in zijn kinderschoenen stond. Naar alle waarschijnlijkheid zal het JITC nog aanpassingen doorvoeren in de toekomst.²²

The National Archives

In 1999 bracht de Engelse regering de wil tot uiting de werking van de overheid te moderniseren.²³ The National Archives van Engeland (voorheen Public Record Office) ontwikkelde in het kader hiervan een aantal richtlijnen, waaronder *Requirements for Electronic Records Management Systems*.²⁴ We vinden hierin een soortgelijke onderverdeling (en bepalingen) zoals in de voorgaande standaarden, aangepast weliswaar aan de Engelse realiteit. Het geheel is wel opgesplitst in vier documenten. In *Part 1: Functional Requirements*²⁵ worden de basisvereisten voor een EDRMS uit de doeken gedaan. Het doel is in een kader te voorzien, waaraan een (aan te kopen of zelf te ontwikkelen) systeem kan getoetst worden. In *Part 2: Metadata standard*²⁶ ligt de nadruk op de implicaties voor metadata in systemen die

²⁰ Department of Defense, *op.cit.*, p. 23

²¹ Department of Defense, *Idem*

²² Gable, J. Everything you wanted to know about DoD 50515.2. In: *Information Management Journal*, November/ December 2002, http://www.findarticles.com/p/articles/mi_qa3937/is_200211/ai_n9164040/pg_8, 30/06/06

²³ Blair, T. Modernising government. White paper. London: The Stationery Office, March 1999, <http://www.archive.official-documents.co.uk/document/cm43/4310/4310.htm>, 07/07/06

²⁴ <http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/>

²⁵ The National Archives, Requirements for Electronic Records Management Systems. Part 1: Functional requirements. Surrey: Public Record Office, 2002, <http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/requirementsfinal.pdf>, 06/11/05

²⁶ The National Archives, Requirements for Electronic Records Management Systems. Part 2: Metadata standard. Surrey: Public Record Office, 2002,

in overeenstemming zijn met deel één. De basisconcepten die aan de basis liggen van het model, inclusief een woordenlijst, worden verklaard in *Part 3: Reference document* ²⁷. In het laatste deel, *Implementation guidance* ²⁸, worden (zoals de titel aangeeft), richtlijnen aangegeven bij de effectieve implementatie.

The National Archives of Australia

Australië kent een lange traditie van dynamisch documentbeheer.²⁹ Ze lag onder andere aan de basis van het *records continuum* model en, via de Australische Standaard 4390, van ISO 15489 (zie boven). Deze concepten worden nog steeds verder uitgewerkt via een relevante wetgeving (sinds 1996 op regelmatige basis aangepast) en door verschillende organisaties binnen en buiten de nationale archieven.³⁰

Een volledig overzicht van het werk verricht in Australië zou ons te ver leiden. De reden om ze aan te halen, is omwille van de volledigheid. De aangehaalde standaarden inzake archief- en documentbeheer zijn, op dit moment, wereldwijd de meest aanvaarde modellen. In de Verenigde Staten speelt de Library of Congress en het Department of Defense een toonaangevende rol. Op het Europese vasteland is MoReq van belang. In Engeland en Australië wordt gewerkt met de raadgevingen van hun respectievelijke National Archives. Overkoepelend wordt rekening gehouden met de ISO-norm(en).

<http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/metadafinal.pdf>,
06/11/05

²⁷ The National Archives, Requirements for Electronic Records Management Systems. Part 3: Reference document. Surrey: Public Record Office, 2002,
<http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/referencefinal.pdf>,
06/11/05

²⁸ The National Archives, Requirements for Electronic Records Management Systems. Part 4: Implementation guidance. Surrey: Public Record Office, 2004,
<http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/implementation.pdf>,
06/11/05

²⁹ Curtin, J. Australian contributions to recordkeeping. In: John Curtin, *Understanding society through its records*, [2003], <http://john.curtin.edu.au/society/australia/index.html>,
11/07/06

³⁰ bvb. de *Records Continuum Research Group* van de Monash universiteit (<http://www.sims.monash.edu.au/research/rcrg/>), Australian Government Locator Service (AGLS) Metadata Standaard (http://www.naa.gov.au/recordkeeping/gov_online/agls/summary.html); voor een volledig overzicht van de aangereikte adviezen door de nationale archieven, zie: <http://www.naa.gov.au/recordkeeping/default.html>

B. Internationale Archiefraad

De Internationale Archiefraad (International Council on Archives – Conseil International des Archives – ICA/CIA) is, zoals de naam al aangeeft, een internationale vereniging, met als doel internationale contacten, opleiding en vorming van archivariissen te stimuleren en ondersteunen. Het is een beroepsvereniging van archivariissen en een vereniging van archiefdiensten (met 1600 leden uit 180 landen) die de preserving, de ontwikkeling en het gebruik van archivistisch erfgoed wereldwijd wil bevorderen.³¹

Met de bedoeling archieftoegangen te standaardiseren (en bijgevolg internationaal uitwisselbaar te maken) werd door een *ad hoc* commissie de *International Standard for Archival Description (General)*, ISAD(G)³², ontwikkeld. Deze werd gepresenteerd in 1992 op het ICA archiefcongres en aanvaard. In 2000 werd een herziene versie te Sevilla gepresenteerd. De ISAD-norm is geen methode en geeft geen regels voor het ordenen van archieven. Het is een handleiding om de beschrijvingsgegevens systematisch te verdelen in een *multilevel description*. Als uitgangspunt worden een aantal beginselen geformuleerd: de basiseenheid van archiefbeheer is het archief(bestand), een toegang moet informatie over de achtergrond, het ontstaan en de context bevatten, binnen een archief worden verschillende ordeningsniveaus onderscheiden (archief, deelarchief, archiefreeks, eventueel deelreeks, bestanddeel, stuk) die van het algemene naar het bijzondere worden beschreven. 26 beschrijvings-elementen die in een beschrijving kunnen voorkomen zijn aangegeven, samen met aanwijzingen hoe en in welke volgorde ze in een toegang moeten toegepast worden. De nadruk ligt vooral op statische archiefstukken.³³ De officiële versie werd in het Engels en Frans gepubliceerd, maar voor andere taalgebieden zijn ook vertalingen beschikbaar.³⁴ De Nederlandse vertaling werd in 2004 uitgebracht als *Algemene Internationale Norm voor Archivistisch Beschrijven*.³⁵

Parallel met ISAD werd de *International Standard Archival Authority Record for Corporate Bodies, Persons and Families*, ISAAR(CPF)³⁶, ontwikkeld voor het beschrijven van contextuele informatie over archiefvormers. De tweede herziene versie werd in 2004 op het congres in Wenen voorgesteld. Voor elke archiefvormer wordt een autoriteitsfiche opgesteld, die bestaat uit een

³¹ Janssens, G. & Put, E. *Geschiedenis, principes en terminologie van de archivistiek. Syllabus*. Brussel: VUB-KUL-UG-UA, 2005-2006, p. 35

³² Committee on descriptive standards, *ISAD (G): General Standard International Archival Description*. Stockholm: ICA, 2000, zie ook: [http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf)

³³ Janssens, G. & Put, E. *Ibidem*, pp. 11-13

³⁴ En kunnen hier gedownload worden: <http://www.ica.org/biblio.php?pdocid=1>

³⁵ Hokke, E.; Horsman, P., Nelissen, M. & Vanneste, W. *ISAD(G)Algemene Internationale Norm voor Archivistisch Beschrijven. Vertaling van de tweede uitgave*. Antwerpen-Leuven-Amsterdam, 2004, zie ook: <http://www.ica.org/biblio/ISAD2NL.pdf>

³⁶ Committee on descriptive standards, *ISAAR (CPF). International Standard Archival Authority Record For Corporate Bodies, Persons and Families. Second edition*. Cranberra: ICA, 2004, zie ook: http://www.icacds.org.uk/eng/isaar2ndedn-e_3_1.pdf

autoriteitshoofding en bijkomende velden die naar andere fiches verwijzen of de archiefvormer beschrijven.³⁷

C. Encoded Archival Description (EAD)³⁸

De wijde verspreiding van het internet heeft ons de mogelijkheid gegeven de archiefinstelling bijna rechtstreeks in contact te brengen met het publiek. Inventarissen kunnen op het web gepubliceerd worden, alsmede digitale en gedigitaliseerde bescheiden. De snelle ontwikkeling van de bestaande technologieën hebben er echter voor gezorgd dat deze publicaties snel verouderen. Hiernaast bestaat het probleem van archiefterminologische onduidelijkheid op het internationaal vlak. Dit werd ondertussen al grotendeels verholpen door de ontwikkeling van ISAD en ISAAR.

Binnen het kader van deze problematiek begon in 1993 een project van the University of California (Berkeley) Library (in samenwerking met andere universiteiten) aan de ontwikkeling van de EAD DTD.³⁹ Het doel van dit project was na te gaan in hoeverre het nuttig zou zijn een niet merkgebonden beschrijvingsstandaard te ontwikkelen voor machine leesbare nadere toegangen tot archieven, bibliotheken, musea en bewaarplaatsen van handschriften.⁴⁰ Een aantal belangrijke voorwaarden werden vastgesteld waaraan de gebruikte techniek zou moeten voldoen. “Het moest in ieder geval mogelijk zijn:

1. de grote hoeveelheden beschrijvende informatie in inventarissen en toegangen van archieven te presenteren, alsmede de onderlinge relaties tussen deze beschrijvingen;
2. de hiërarchische relaties tussen verschillende niveaus van beschrijvingen ook in digitale vorm te behouden;
3. beschrijvingen die op meerdere niveaus van toepassing zijn weer te geven;
4. binnen een hiërarchische informatiestructuur goed en gemakkelijk te navigeren;
5. specifieke elementen in beschrijvingen te indexeren zodat tevens op deze elementen kan worden gezocht (retrieval).⁴¹

Van de mogelijke kandidaten, bleek de *Standard Generalized Markup Language* (SGML) het meeste te voldoen aan deze eisen (in een later stadium

³⁷ Janssens, G. & Put, E. *Geschiedenis, principes en terminologie van de archivalieken*. Syllabus. Brussel: VUB-KUL-UG-UA, 2005-2006, p. 13

³⁸ Encoded Archival Description, Version 2002, Official Site: <http://www.loc.gov/ead/>

³⁹ DTD: Document Type Description. Een DTD legt vast welke *tags* en attributen gebruikt worden om de inhoud van een SGML, XML of HTML document te beschrijven. *DTD*. In: *Webopedia Computer Dictionary*, <http://www.webopedia.com/TERM/D/DTD.html>, 29/06/06; hier kom ik later nog op terug.

⁴⁰ *Development of the Encoded Archival Description DTD*. Encoded Archival Description, Version 2002, <http://www.loc.gov/ead/eaddev.html>, 29/06/06

⁴¹ Siepel, A. & de Vries, H. *Encoded Archival Description (EAD). Het digitaliseren van toegangen op archieven*. Amsterdam: Archiefschool, 2001, p. 6, <http://www.archiefschool.nl/docs/siepenco.pdf>, 29/06/06

werd de *eXtensible Markup Language*, XML, eveneens geïncorporeerd) en werd vervolgens gebruikt als basis voor de technische uitwisseling. Als standaard voor de beschrijvings-elementen werd geopteerd voor ISAD(G).⁴²

D. Dublin Core

The Dublin Metadata Core Element Set, of kortweg *Dublin Core*, is een kern lijst van metadata elementen zoals overeengekomen op de OCLC/NCSA Metadata workshop in maart 1995 in Dublin, Ohio.⁴³ Dit rapport werd verder ontwikkeld en gepubliceerd als een *National American Standard* ⁴⁴ “for cross-domain information resource description.”⁴⁵ Sinds 2003 werd ze eveneens aanvaard als ISO-norm.⁴⁶ Deze workshop verenigde groeperingen van verschillende professionele achtergronden die zich bezig hielden met de problematiek van metadata. Het doel was de ontwikkeling van een gemakkelijk te gebruiken vaste set van metadata elementen die door verschillende doelgroepen (archieven, bibliotheken, musea, ...) kan worden gebruikt. Ze zijn bijgevolg uitwisselbaar en verhogen de mogelijke zoekacties op een gestandaardiseerde manier. In de eerste instantie ging de aandacht naar netwerk omgevingen en het bevragen van HTML(*HyperText Markup Language*)-pagina's.⁴⁷

Uit dit initiatief werd het *Dublin Core Metadata Initiative (DCMI)* geboren.⁴⁸ Deze organisatie heeft als missie de promotie van “the widespread adoption of interoperable metadata standards and developing specialized metadata

⁴² Pitti, D. Encoded Archival Description. An introduction and overview. In: *D-Lib Magazine*, Volume 5, Number 11, November 1999, <http://www.dlib.org/dlib/november99/11pitti.html>, 29/06/06 ; voor meer info zie o.m. : Metadata Advisory Group, EAD – Encoded Archival Description. MIT Libraries, 2004, <http://libraries.mit.edu/guides/subjects/metadata/standards/ead.html>, 29/06/06; Dempsey, L. & Heery, R. Specification for resource description methods Part 1: A review of metadata: a survey of current resource description formats. Bath: The UK Office for Library and Information Networking (UKOLN), The University of Bath, 1999, http://www.ukoln.ac.uk/metadata/desire/overview/rev_06.htm, 29/06/06; RLG, Encoded Archival Description Activities. RLG, 2006, http://www.rlg.org/en/page.php?Page_ID=411, 29/06/06; Cover, R. (ed.) Technology reports. Encoded Archival Description (EAD). In: *Cover Pages, Online resource for markup language technologies*, 2002, <http://xml.coverpages.org/ead.html>, 29/06/06

⁴³ Weibel, S.; Godby, J.; Miller, E.; Daniel, R. OCLC/NCSA Metadata Workshop Report. <http://dublincore.org/workshops/dc1/report.shtml>, 29/06/06

⁴⁴ ANSI/NISO, The Dublin Core Metadata Element Set. Bethesda: NISO Press, 2001, <http://www.niso.org/standards/resources/Z39-85.pdf>, 29/06/06

⁴⁵ Dublin Core. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Dublin_Core, 29/06/06

⁴⁶ ISO 15836:2003, Information and documentation – The Dublin Core metadata element set

⁴⁷ Dempsey, L. & Heery, R. Specification for resource description methods Part 1: A review of metadata: a survey of current resource description formats. Bath: The UK Office for Library and Information Networking (UKOLN), The University of Bath, 1999, http://www.ukoln.ac.uk/metadata/desire/overview/rev_05.htm, 29/06/06

⁴⁸ Dublin Core Metadata Initiative, <http://dublincore.org/>

vocabularies for describing resources that enable more intelligent information discovery systems.”⁴⁹

Op dit moment wordt Dublin Core toegepast in 75 projecten.⁵⁰

E. Technische standaarden

Naast de archiefwetenschappelijke kennen we ook nog technische (industriële) standaarden. Deze slaan terug op de kwaliteit van het gebruikte materiaal, zowel softwarematig als van de fysieke drager.

Voor de dragers kunnen we ons beroepen op ISO-normen, bvb. ISO 9660⁵¹ voor CD's en ISO 9529⁵² voor 3,5” diskettes⁵³, en op ANSI-normen⁵⁴, bvb. ANSI INCITS 27⁵⁵ voor het aanbrengen van labels op magnetische banden. Voor software maakt Boudrez een onderscheid tussen officiële en *de facto* standaarden. Officiële standaarden worden vastgelegd door erkende standaardiseringorganisaties, waarin (inter)gouvernementele organisaties zetelen (bvb. ISO, ANSI, IEC, ...). *De facto* standaarden worden verder verdeeld in specificaties, open en gesloten formaten. De specificaties en open formaten zijn beide publiek gedocumenteerd. De eerste groep worden opgesteld door niet-officiële standaardiseringorganisaties (en zijn dus niet afhankelijk van één enkele producent), terwijl de tweede groep worden beheerd door één producent. De gesloten formaten zijn eveneens afhankelijk van één commerciële actor, maar de technische specificatie werd niet vrijgegeven. Ze kunnen als standaard beschouwd worden omwille van de brede marktpenetratie.⁵⁶

⁴⁹ About the Initiative. DCMI, <http://dublincore.org/about/>, 29/06/06

⁵⁰ Voor een volledige lijst zie: <http://dublincore.org/projects/>

⁵¹ ISO/IEC DIS 9660: 1999 (E), Information processing – Volume and file structure of CD-ROM for Information Interchange, http://www.y-adagio.com/public/standards/iso_cdromr/tocont.htm, 11/07/06

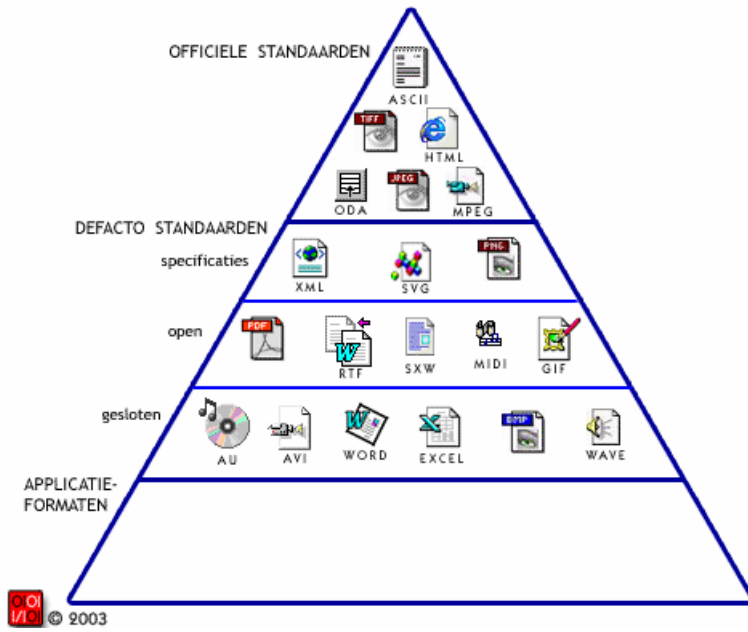
⁵² ISO/IEC 9529-1 : 1989, Information Processing Systems - Data Interchange on 90 mm (3.5 in) Flexible Disk Cartridges Using Modified Frequency Modulation Recording at 15 916 ftprad on 80 tracks on Each Side – Part 1: Dimensional, physical and magnetic characteristics; ISO/IEC 9529-2 : 1989, *Idem*, Part 2: Track Format

⁵³ Voor een overzicht, zie: <http://www.msscience.com/iso.html>

⁵⁴ American National Standards Institute, <http://www.ansi.org/>

⁵⁵ ANSI INCITS 27-1987 (R1989), File structure and labeling of magnetic tape for information interchange

⁵⁶ Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief, 2004, pp. 67-70; voor een technisch overzicht, zie: Boudrez, F. Standaarden voor digitale archiefdocumenten. Technisch rapport. Antwerpen: Stadsarchief, 2001–2003, <http://www.expertisecentrumdavid.be/davidproject/teksten/DAVIDbijdragen/Standaarden.pdf>, 28/11/06



Bron: Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief, 2004, p. 68

3. Bevindingen

Tot nu toe werd in vogelvlucht een overzicht gegeven van de theoretische onderbouw voor een goed begrip van archivistische vereisten voor de uitbouw van een digitaal depot. De vraag is nu: wat doen we daar mee? Wat is het nut voor de digitale archivering zoals ze uiteindelijk zal geïmplementeerd worden in het Vlaams Parlement? Anders gesteld: tot nu toe hebben we de bomen gezien, de volgende taak is een afbakening van het bos.

Alvorens we van een archiefbeheer kunnen spreken moeten we zeker kunnen zijn dat de archiefbescheiden die aan het depot worden toevertrouwd authentiek, betrouwbaar, integer en herbruikbaar zijn. Essentiële eigenschappen moeten bewaard blijven. Beslissingen over de uiteindelijke bewaarvorm van het document, de hoeveelheid en kwaliteit van de toe te voegen metadata, de manier waarop het archiefstuk zal ter beschikking zal worden gesteld aan de gebruiker, de wijze van opzoeken, ... hebben een invloed op de keuze van de te volgen conserverings- en archiveringsstrategie(ën).

De strategieën die op dit moment voor handen zijn op het vlak van preservatie, zijn het afdrukken van een hard copy, de uitbouw van een computermuseum, converteren van oude bestanden naar een nieuwe versie van hetzelfde bestandsformaat, migreren naar een ander formaat of het emuleren van de oorspronkelijke softwareomgeving waarbinnen het bestand werd gecreëerd. De eerste twee mogelijkheden kunnen al onmiddellijk worden afgewezen. De hard copy tactiek kan handig zijn om kleine dossiers te beheren door de archiefvormer zelf, maar daar stoppen de voordelen dan ook. Binnen het geheel van digitale documenten gebruikt in het Vlaams Parlement lenen enkel sommige kantoordocumenten zich tot deze strategie. Afgedrukte tekstbestanden, presentaties en spreadsheets kunnen bij dossiers gevoegd worden zodat op een fysieke manier de volledigheid kan worden behouden. Voor de eigenlijke tekst van een e-mail kan hetzelfde worden gedaan. Deze manier van werken doet echter een beetje Amish aan.¹ Op die manier wordt, uit gewoonte, vastgehouden aan een manier van werken die stilaan aan het verdwijnen is. Afprinten van documenten wordt tegenwoordig in feite enkel nog gedaan omwille van het leesgemak. In het Vlaams Parlement werd die evolutie zelfs formeel aangegeven door uittredend voorzitter Norbert De Batselier in zijn beleidsplan voor de periode 2004-2009. Een goede werking wordt enkel mogelijk geacht wanneer een optimale toegang tot informatie wordt gegarandeerd. Daarom moet er “verder gewerkt worden aan een modern maar betaalbaar informatiebeleid, gericht op toegankelijkheid van de informatie en op het meer efficiënt maken van de informatiestromen.” Een passage die van toepassing is op de archieven vermeldt dat een goed informatiebeleid rekening houdt met het verbeteren

¹ Orthodoxe leden van de Amish geloofsgemeenschap zweren het gebruik van moderne technologieën af. Hun manier van leven is nog altijd hetzelfde als tijdens de negentiende eeuw, inclusief het gebruik van paard en kar als vervoermiddel.

van het zoeken naar en het aanbieden van relevante informatie. Dit beleid moet er “op gericht zijn om de werking van het Vlaams Parlement ... te optimaliseren en te ondersteunen. Het Algemeen Secretariaat zal blijvend de snelle technologische evoluties op het vlak van ICT opvolgen en evalueren. Daar waar de nieuwe technieken een toegevoegde waarde kunnen hebben voor de organisatie zullen er gecoördineerde projecten worden opgestart. Zo zal men op korte termijn het Vlaams Parlement omdopen tot een virtuele Hot-Spot.”²

Het andere uiterste van het spectrum wordt ingenomen door de computermuseumstrategie. Vanuit historisch oogpunt is het interessant om oude hard- en software aan de praat te houden, maar dit is niet de taak van een parlementaire instelling. De bestaande computermusea zijn daar ongetwijfeld beter voor uitgerust. Deze strategie maakt het nodeloos moeilijk (voor ICT en de gebruiker) en kostelijk voor de organisatie zonder daarom een concrete meerwaarde te bieden.

Het gebruik van conversie is, opnieuw, een strategie die op de korte termijn een aantal voordelen biedt aan, voornamelijk, de archiefvormer van de stukken. Door te *upgraden* blijft de volledige functionaliteit van het document behouden. Dit is vooral praktisch indien nog veranderingen moeten aangebracht worden en dus wanneer de dynamische fase nog niet is afgelopen. Vanaf het moment dat de statische (en meestal al vanaf de semi-statische) fase een aanvang heeft genomen, hebben de stukken meer een informatieve waarde. Bij hergebruik van dezelfde gegevens is het aan te raden een nieuw document te creëren. Door een bestand te converteren maken we in feite sowieso een nieuw document aan, waardoor authenticiteitseisen in het gedrang komen.

Emulatie heeft een aantal grote voordelen voor het behoud van volledige systemen en dus bijgevolg de werking van de documenten die hiermee worden aangemaakt. Bij het lezen van de theorie kunnen we niet anders dan denken dat we de ideale oplossing hebben gevonden. Deze strategie heeft echter ook een aantal nadelen. Een zeer volledige oplossing wordt voorgesteld, waarbij het gebruikte systeem in zijn geheel wordt nagebootst. Bij het archiveren door middel van emulatie van computerspelletjes zoals Chuckie Egg³ is dit nodig omdat anders geen correcte representatie mogelijk is, maar in hoeverre dit een vereiste is voor tekstbestanden, bijvoorbeeld, blijft de vraag. Vooral het concept van een *Universal Virtual Computer* is

² De Batselier, N. Een slagvaardig Vlaams Parlement, 2004 – 2009. 3^e Beleidsplan. Februari 2005, p. 26-27, <https://intranet.vlaamsparlement.be/docs/portlet/p615dp26lijn487.doc>, 12/07/06

³ Chuckie Egg is een platformspelletje dat werd uitgebracht in 1983 om te worden gespeeld op een van de eerste home computers, BBC Micro. In kader van het CAMiLEON project, werd onderzocht in hoeverre gebruikers een verschil opmerken tussen gemigreerde en gemuleerde digitale objecten. Merkwaardig genoeg bleken beide strategieën gelijkaardige afwijkingen met zich mee te brengen in vergelijking met een speelsessie op de originele soft- én hardware. Hedstrom, M. & Lampe, C. Emulation vs. Migration: Do users care? In: *RLG DigiNews*, Volume 5, Number 6, December 15, 2001, <http://www.rlg.org/preserv/diginews/diginews5-6.html#feature1>, 06/04/06

bijzonder veelbelovend. Simplistisch gesteld: we maken nu een UVC-platform waarbinnen onze stukken gearchiveerd worden en ons werk is gedaan. Dit is een emulator die als een software applicatie gedraaid zal worden op een nog onbekend platform van nog niet bestaande hardware van de toekomst. Aangezien we nu nog niet weten hoe die eruit zullen zien, moet de eigenlijke emulator op dat tijdstip worden geschreven. De UVC vormt een platform waarop programma's kunnen gedraaid worden die specifiek voor de UVC zijn geschreven. Zo een programma decodeert het bestandsformaat van digitaal object en haalt *element tags* op die specifieke informatie bevatten over de inhoud van de data. Op basis van die gegevens wordt een *Logical Data View* gemaakt dat eruitziet als XML. Het bestand kan vervolgens worden gelezen in een *Logical Data Viewer*, die eveneens in de toekomst moet worden geschreven. Het verschuiven van een gedeelte van de oplossing naar de toekomst toe houdt echter wel een zeker gevaar in, zolang deze methode nog niet volledig is getest. De opgesomde nadelen tot nu toe, slaan terug op verschillen in visie op de archiveringsfunctie. Het grootste nadeel ligt in het feit dat de verschillende vormen van deze strategie hun nut op de lange termijn nog niet hebben bewezen.

Aangezien in het Vlaams Parlement voornamelijk gewerkt wordt met software die eigendom is van commerciële bedrijven, is het af te raden documenten te bewaren in het oorspronkelijke bestandsformaat. Migreren naar *open source* of internationale standaarden is aangewezen. Ook hier zijn nadelen aan verbonden. Het nieuwe formaat zal niet eeuwig behouden kunnen worden. Op termijn zullen nieuwe migraties noodzakelijk zijn. De rode draad in het verhaal is dat een methode die voor eens en altijd alle problemen oplost niet voor handen is. En eender welke oplossing wordt geïmplementeerd, nieuwe investeringen zullen periodiek terugkomen. Maar, keuzes moeten gemaakt worden, een beetje op basis van *choosing the lesser evil* (hoewel dat een beetje overdreven klinkt). Op dit moment lijkt migratie de beste oplossing te zijn. Zoals gezegd, zal na verloop van tijd echter opnieuw gemigreerd moeten worden. Dit zal in bulk moeten uitgevoerd worden. Het is dan zeer belangrijk een robuust foutenopsporingsmechanisme te hanteren om de integriteit van alle stukken te blijven behouden. In dit opzicht lijkt Cedars' *migration on request* een interessante piste te bieden. Indien eerst gemigreerd wordt naar, bijvoorbeeld, een *open source* formaat, kan daar aanvullend op een later tijdstip een migratietool voor worden geschreven. Op het moment dat de aanvraag van archiefstukken wordt verwerkt, genereert deze tool een geschikte DIP voor de gebruiker in een formaat dat dan leesbaar is. Aangezien het bewaarde formaat op basis van *open source* is, is de broncode ook gekend en kan een tool door de Directie Informatica worden ontwikkeld. Op die manier is het niet nodig om gigantische hoeveelheden op geregelde tijdstippen om te zetten naar nieuwe formaten, maar moeten enkel de verschillende tools op geregelde tijdstippen worden aangepast. Een goede organisatie is wel vereist om het geheel van de IT-infrastructuur te beheren. In eerste instantie lijkt het mij aangewezen migratie *pur sang* toe te passen. Er moet dan een gelimiteerde hoeveelheid van formaten worden toegekend aan de verschillende klassen van AIP. In een later stadium kan de *migration on request* strategie onderzocht worden en getest op haalbaarheid.

Het gebruik van standaarden is aan te raden omwille van verschillende redenen. Een eerste reden is het verhogen van de betrouwbaarheid van en het vertrouwen dat gesteld wordt in het archief. Het concept van een betrouwbaar digitaal depot zorgt voor een aantal duidelijke richtlijnen waaraan een archief moet voldoen om haar opdracht te vervullen. Aan de hand hiervan kan een concreet plan opgesteld of aangevuld worden voor de uitbouw van digitaal archief. Bovendien zijn de ideeën van het OAIS-model erin verwerkt. Dit model zorgt voor een bruikbaar hulpmiddel waarop een concrete strategie, wat betreft de technische en procedurele aspecten van digitale archivering, gemodelleerd kan worden. De concepten die door deze werken worden aangereikt, hebben een hoog *common sense* – gehalte. Veel elementen behoren ergens wel tot de basisidee dat men heeft van hoe een digitaal archief eruit moet zien. Door het allemaal uitgeschreven voor ogen te hebben worden ze echter niet vergeten, én aangevuld met een aantal dingen waaraan niet zou gedacht worden en *best practices*.

Op het vlak van *electronic document and record management systems* ligt het voor de hand dat ISO 15489 wordt geïmplementeerd. Deze standaard is ondertussen voldoende gekend binnen de archiefwetenschappelijke gemeenschap om als kwaliteitslabel gehanteerd te worden. Hierbuiten wordt een certificatie *tout court* en een ISO-certificatie in het bijzonder, eveneens aanzien als een element waarmee kwaliteit kan afgelezen worden. Los hiervan, wordt dit document, net zoals het *trusted digital repositories*-rapport en OAIS, gekenmerkt door een hoge *common sense*. En ook hier wordt een handig instrument geboden, waarop het onderzoek naar de vereiste specificaties voor een EDRMS aangepast (aan de eigen organisatie) kan gemodelleerd worden. Het enige nadeel van het rapport is dat het nogal summier is (26 pagina's voor deel één, 53 voor deel twee). Een uitbreiding op de uitwerking van een stappenplan vinden we in de DIRKS Manual. "The DIRKS methodology is an 8-step process ... based on and expands the best-practice approach outlined in Australian Standard AS 4390-1996 ... and International Standard ISO 15489."⁴

ISO 15489 kan een filosofisch model genoemd worden, met de nadruk op een geheelconcept rond de opbouw van een EDRMS. Het grootste belang ervan ligt in het aanreiken van een aangepast vertaling van het archivistische gedachtengoed naar een elektronische werkomgeving. Voor de opbouw van technische, functionele vereisten biedt MoReq een praktisch werkmiddel. Deze specificatie voor elektronische records kan aangepast worden aan de specifieke noden van een organisatie, als aanvulling van het onderzoek verricht in het kader van de bovenstaande ISO-norm. Bijkomend kan de Nederlandse aanpassing, ReMANO, er langs gelegd worden. Dit werk is een aanpassing van de MoReq-norm aan de Nederlandse realiteit⁵, die

⁴ The National Archives of Australia, [The DIRKS Manual: A strategic approach to managing business information](http://www.naa.gov.au/recordkeeping/dirks/dirksman/dirks.html). Commonwealth of Australia, 2001 (revised 2003), p. 6, <http://www.naa.gov.au/recordkeeping/dirks/dirksman/dirks.html>, 06/11/05

⁵ "Met inachtneming van vooral Nederlandse wet- en regelgeving op het gebied van overheidsinformatievoorziening en archiefbeheer, de in Nederland gebruikelijke methoden,

werd aangevuld met concordanties met het Amerikaanse DoD 5015.2 en het Duitse DOMEA-koncept. Door de (relatieve) taaleenheid tussen Vlaanderen en Nederland worden bepaalde concepten sneller uitgewisseld, en kunnen bepaalde elementen overgenomen of aangepast worden aan de Belgische realiteit. Ondertussen is het bovendien mogelijk een ReMANO-certificatie te bekomen.⁶

Op het vlak van archiefbeschrijvingen lijkt het mij aangewezen de internationale standaard ISAD(G) te gebruiken. Ze wordt gedragen door de Internationale Archiefraad en is specifiek aangepast aan archiefbeschrijvingen. Bovendien kan ze gekoppeld worden aan EAD. In het kader van het te ontwikkelen online archiefloket van het Vlaams Parlement (zie Deel II), kan al een groot deel van dit werk op automatische manier verlopen, door de gemaakte beschrijvingen te importeren. Op die manier kan bovendien een inventaris opgebouwd worden die internationaal begrijpelijk is. Door de opbouw van het algemene naar het bijzondere wordt eveneens een soort van algemeen aanvaarde standaard gebruikt (ik gebruik de term zeer los nu). De meeste mensen kunnen ondertussen met een GSM of een IPOD werken. Daarin wordt op dezelfde manier een typische boomtakstructuur gehanteerd voor het terugvinden van, respectievelijk, telefoonnummers en liedjes. Koppelen we daaraan het principe van het doorklikken op een website, dan kan op een algemeen begrijpbare manier een inventaris op het web gepubliceerd worden. Door het aangeven van het gebruik van de ISAD-structuur, kan ook verduidelijkt worden op welke manier automatische zoekacties resultaten zullen genereren.

technieken en terminologie voor informatie- en archiefbeheer, en verschillende al bestaande standaards en modellen, zijn software specificaties voor Records Management Applicaties voor de Nederlandse Overheid (ReMANO) ontwikkeld.” Horsman, P.J.; van Bussel, G.J. & Waalwijk, H. Softwarespecificaties voor Records Management Applicaties voor de Nederlandse Overheid. ReMANO 2004. Amsterdam: Archiefschool, maart 2004, p. 8

⁶ Stichting Certificatie ReMANO, <http://www.remano.nl/>

Deel II: De eerste stappen in de praktijk



1. Het Vlaams Parlement als Betrouwbaar Digitaal Depot uitbouwen

In dit hoofdstuk wordt eerst een samenvatting gegeven van de vereisten om een betrouwbaar e-depot te worden. Vervolgens wordt een samenvattende conclusie geformuleerd, toegepast op het Vlaams Parlement. Indien we opnieuw kijken naar de definitie van een betrouwbaar digitaal depot zoals ze gesteld wordt door the Online Computer Library Center (OCLC) en the Research Libraries Group (RLG), zien we dat ze gelijk is aan die van het Vlaams Parlement: “one whose mission is to provide reliable, long-term access to managed digital resources to its designated community.”¹ In de strategische nota van het Archief en de Biografische dienst wordt de missie als volgt verwoord: 1° “het beheer en de ontsluiting op de lange termijn van het administratief en historisch-wetenschappelijk archief van het Vlaams Parlement”, 2° “het beheer en de ontsluiting van de biografische dossiers van de leden.”² Het archief ondersteunt de dagelijkse werking van het parlement en zorgt tot de mogelijkheid van controle op eerder genomen beslissingen en uitgevoerd werk. Om dit te kunnen doen moet ze betrouwbare documenten (blijvend) kunnen voorleggen. In de volgende paragrafen zal ik de verschillende attributen en verantwoordelijkheden zoals ze worden voorgelegd in *Trusted Digital Repositories* (TDR) nagaan en toetsen aan het Parlement. Een bruikbaar instrument hierbij is de RLG-uitgave *An audit checklist for the certification of trusted digital repositories*.³ In dit document wordt een overzicht en een hulpmiddel aangereikt, waarmee een betrouwbaar digitaal depot (in wording) kan geëvalueerd worden. In hoofdstuk II (Verificatie en Certificatie Criteria, Audit & Certification Criteria) worden de verschillende elementen, waarmee een digitaal depot beoordeeld kan worden op haar betrouwbaarheid, verklaard. In hoofdstuk III wordt een concreet *audit instrument for the certification of Trusted Digital Repositories* aangereikt.⁴ De criteria worden opgesplitst in vier delen: organisatie, functies, processen en procedures van een depot, de *designated community* en de bruikbaarheid van informatie, technologieën en technische infrastructuur.

¹ RLG/OCLC, *Trusted digital repositories. Attributes and responsibilities. An RLG-OCLC Report*. Mountain View: RLG, 2002, p. 5; <http://www.rlg.org/en/pdfs/repositories.pdf> , 15/03/06

² Directie Informatie, Archief en Biografische Dienst, *Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement*. 17 november 2005, p. 1

³ RLG, *An audit checklist for the certification of trusted digital repositories. Draft for public comment*. Mountain View: Research Libraries Group, 2005

⁴ RLG, *op.cit.*, pp. 8-58

1.1 Criteria

A. Organisatie⁵

Organisatorische attributen hebben een invloed op performantie, aansprakelijkheid en duurzaamheid van een depot. TDR onderscheidt vier groepen: administratieve verantwoordelijkheid, organisatorische levensvatbaarheid, financiële duurzaamheid en procedurele aansprakelijkheid. Dit geheel kan omschreven worden als de organisatorische infrastructuur.

Organisatorische levensvatbaarheid

Een organisatie kan aantonen dat haar missie, legale status en werkzaamheden afdoende zijn om een TDR te onderhouden. Ze moet levensvatbaarheid en betrouwbaarheid kunnen aantonen.

Het depot:

- Heeft een missieverklaring dat de toewijding weerspiegelt voor het beheer van de haar toevertrouwde digitale informatie. De missieverklaring moet toegankelijk zijn voor de archiefvormers en andere stakeholders.
- Heeft een formeel rampenplan en een plan voor de overname van het archief indien het depot zou opgeheven worden of drastisch van doelstelling zou veranderen. Doordat een archief de belofte maakt de opgenomen (historische) stukken blijvend te beheren, is het noodzakelijk een regeling te treffen voor de overname van het archief wanneer het noodzakelijk zou zijn. Zulke regelingen kunnen beter worden getroffen wanneer de organisatie gezond is, en niet tijdens crisismomenten, om mogelijk verlies van bescheiden tegen te gaan.

Organisatorische structuur en personeel

Een depot moet eigen personeel hebben dat theoretisch en technisch onderlegd is om ontwikkeling op de lange termijn te waarborgen. Er moet voldoende personeel zijn om het aangenomen werk naar behoren te kunnen uitvoeren. De archiefmedewerkers hebben voldoende *skills* en expertise om hun opdracht(en) te kunnen uitvoeren. Op geregelde tijdstippen worden professionele trainingen ingericht om deze up-to-date te houden.

⁵ RLG, An audit checklist for the certification of trusted digital repositories. Draft for public comment. Mountain View: Research Libraries Group, 2005, pp. 9-13

Procedurele aansprakelijkheid en beleidskader

Een depot zorgt voor een duidelijke documentatie van de vereisten, beslissingen, ontwikkeling en ondernomen handelingen gericht op toegankelijkheid op de lange termijn van de opgenomen bescheiden. Door in alle transparantie te werken, kan alleen maar duidelijker aangetoond worden dat het depot betrouwbaar is en/of wil zijn.

Het beleid en de ingestelde procedures moeten meegaan met de tijd. Daarom is het aan te raden op regelmatige tijdstippen deze aan een herbeoordeling te onderwerpen en indien nodig aan te passen. Aanvullend moet voorzien worden in officiële kanalen langs dewelke gebruikers en archiefvormers kunnen terugkoppelen naar de archiefdienst. Door middel van constructieve kritiek worden problemen sneller opgelost, vereisten sneller aangepast en wordt de ontwikkeling bevorderd. Aangebrachte veranderingen moeten wel minutieus worden bijgehouden, opdat latere gebruikers en beheerders op de hoogte zijn van de gevolgde conserveringsstrategieën.

De integriteit van de beheerde stukken moet bewaard blijven. Aangepaste voorzorgsmaatregelen om die te verzekeren moeten ontwikkeld, waar nodig aangepast, worden. Verlies van gegevens bij migraties of andere ondernomen handelingen moet worden vermeld.

Financiële duurzaamheid

Een betrouwbaar depot moet kunnen aantonen dat ze over voldoende financiële middelen beschikt om het beheer van de opgenomen stukken over de tijd heen voort te blijven zetten. Hierom wordt een duurzaam business plan opgesteld (zowel voor de korte als de lange termijn), dat jaarlijks wordt herzien. Door pro-actief te werken, wordt vermeden dat financiële moeilijkheden plots uit de lucht komen vallen. Ook hier wordt gewerkt in alle transparantie om betrouwbaarheid te verhogen. Een voortdurende inzet wordt aangetoond op het vlak van risico-, voordelen-, investerings- en uitgave analyse en rapportering. Het doel is uiteraard deze verschillende elementen in evenwicht te houden.

Contracten, licenties en aansprakelijkheden

Indien een depot digitale bescheiden beheert voor derden, zijn aangepaste overeenkomsten voor contracten en overdrachten vereist. Hierin worden de rechten, verantwoordelijkheden en verwachtingen van de verschillende partijen op een wettelijk bindende manier omschreven. Op die manier worden mogelijke problemen vermeden. Beperkingen op het vlak van copyright en aangekochte licenties worden eveneens expliciet gesteld.

B. Functies, processen en procedures⁶

In dit deel worden de functies, processen en procedures van een depot besproken, die nodig zijn voor de opname, het beheer en de toegankelijkheid van digitale objecten op de lange termijn. Deze functies zijn verdeeld in vijf onderdelen, die toelaten ze te plaatsen onder de functionele entiteiten van OAIS.

Opname/ acquisitie

Opname (*ingest*) is een algemene term die het proces beschrijft dat plaats vindt voor de uiteindelijke vorm van bewaring van een stuk aanwezig is in het depot. Deze fase is beëindigd wanneer een AIP, samen met de metadata, veilig aanwezig is in het depot. De manier waarop de stukken wordt overgedragen is afhankelijk van de relatie die het depot heeft met de archiefvormers. In het geval van overheidsinstellingen wordt dit doorgaans geregeld aan de hand van een archiefreglement. De digitale objecten die aanvaard worden voor preservatie zouden een weerspiegeling moeten zijn van de missieverklaring en de interessegebieden van de *designated community*. Gebruikers moeten gemakkelijk de relatie kunnen begrijpen tussen het depot, haar missieverklaring en haar collectie. Duidelijke overeenkomsten betreffende de overdracht bepalen de toe te voegen metadata die ervoor zorgen dat gebruikers de objecten kunnen gebruiken zonder nood aan toegang tot de archiefvormers, andere experts, en eventueel de archivaris. Op basis van schriftelijke afspraken worden alle aspecten van acquisitie, onderhoud en toegang vastgelegd.

De essentiële eigenschappen worden vastgesteld voor elke klasse van digitaal object. Dit gebeurt in overeenstemming met de archiefvormers. Op basis hiervan zal de manier van bewaren en de toe te voegen technische metadata bepaald worden. Er wordt eveneens vastgelegd, in een identificeerbare, uitgeschreven definitie, hoe elke SIP eruit moet zien. Voor de opname worden de authenticiteit, integriteit en de volledigheid van de SIP's nagekeken volgens geijkte procedures. De metadata die al werd toegevoegd wordt eveneens gecontroleerd en waar nodig aangevuld. Op vastgestelde momenten moet SIP's AIP's worden (en dus opgenomen in het depot) of vernietigd worden. De archiefvormer wordt hiervan tijdig op de hoogte gesteld.

⁶ RLG, An audit checklist for the certification of trusted digital repositories. Draft for public comment. Mountain View: Research Libraries Group, 2005, pp. 14-31

Bewaring in het archief: management van de gearchiveerde data

Acties moeten ondernomen worden om de opgenomen objecten blijvend te bewaren en een versie ter beschikking te stellen aan de eindgebruikers die op een sterke wijze gelinkt blijft aan het originele object.

Elke (klasse van) AIP wordt op een identificeerbare, uitgeschreven manier gedefinieerd, op zulke wijze dat bewaring op de lange termijn wordt verzekerd. Door een expliciete, duidelijke beschrijving worden problemen vermeden bij latere acties ondernomen ter preservering. Verder moet ook worden beschreven hoe SIP's worden getransformeerd tot AIP's. Zo kan later aangetoond worden dat het gepreserveerde object werd afgeleid van het oorspronkelijke, overgedragen stuk.

Elke AIP moeten unieke *identifiers* hebben gebaseerd op een vastgelegde conventie voor benaming. Op die manier worden duplicaten vermeden en kunnen alle handelingen die van toepassing zijn op een welbepaalde AIP getraceerd worden door de tijd heen en los van veranderingen in systemen of bewaring.⁷

Elke AIP wordt voor finale opname in het depot nagekeken op volledigheid, authenticiteit en integriteit.

Planning van preservatie, migratie en andere strategieën

Een depot moet hedendaagse, gedocumenteerde preserveringsstrategieën, die hun nut al hebben bewezen, instellen en implementeren. Data moet bewaard worden in overeenstemming met voorop vastgestelde, gedocumenteerde procedures en beleidsplannen. Met deze documenten kan aangetoond worden, in alle transparantie, welke praktijken en procedures worden toegepast.

Een depot moet een aantal dingen kunnen aantonen om toekomstige toegankelijkheid te waarborgen:

- Beslissingen over aanvaardbare formaten (bvb. beslissingen omtrent toegestane formaten)
- Een uitgebreid *workflow* systeem voor de overdracht van digitale objecten (bvb. transferprotocollen)
- Verwachte en/of toegepaste preserveringsacties die terugslaan op bepaalde (klassen van) AIP's (bvb. preservatieplannen, logboeken)
- Beleidsbeslissingen, procedures en praktijken betreffende bewaring die effectieve overdracht en betrouwbare bewaring verzekeren. Alsook de bereidheid tot handelen garanderen wanneer ingrijpende

⁷ Indien een SIP voor opname over een unieke *identifier* beschikt, wordt deze best mee gearchiveerd op zulke wijze dat een duidelijke relatie met de latere AIP behouden blijft.

- technologische veranderingen zich voordoen (bvb. veiligheidsplannen)
- Onafhankelijke manieren om de verwachte inhoud van het depot na te gaan op basis van een beveiligd spoor van de ontvangen objecten (bvb. een inventaris die niet kan gewijzigd worden)

Een preservatiestrategie zorgt voor een blauwdruk, waarmee de bovenstaande elementen kunnen worden uitgevoerd in een evoluerende omgeving. Ze zorgt voor het opvolgen van veranderingen die een invloed kunnen hebben op de preservatie, een manier om de implicaties van deze veranderingen te interpreteren, een gepland antwoord hierop (inclusief een manier om dit antwoord te implementeren). Zo een strategie moet ook de condities vastleggen voor de mogelijke vernietiging van AIP's.

Data management

Een systeem moet op een efficiënte manier beschrijvende informatie (metadata) bewaren en kunnen gebruiken. Dit is essentieel voor het terugvinden van bestanden. Wanneer de gevraagde informatie is gevonden, moet aan gebruikers een bruikbare kopie ter beschikking worden gesteld. Ze moeten ondermeer weten welke software vereist is voor het openen van het bestand, hoe een bepaalde kopie kan opgevraagd worden (aan de hand van uniek identificatiemiddel of een archiefreferentie).

Op een formele manier moet vastgelegd worden hoe, wanneer en door wie (of wat) metadata wordt toegevoegd. Deze kan ingevuld worden door de archiefvormers, de archiefdienst, automatisch gegenereerd worden door een systeem of een combinatie van het voorgaande. Referentiele integriteit tussen de AIP en de geassocieerde metadata moet gecreëerd en nagekeken worden.

Toegankelijkheid (access management)

Het toegankelijkheidsbeleid moet op een correcte en duidelijke manier worden toegepast. Toegang kan geregeld worden door computers, door mensen of binnen een hybride systeem. Het is uiteraard de bedoeling dat enkel personen met de juiste autorisatie toegang krijgen tot bepaalde bestanden. Indien ze geen toegang krijgen omwille van fouten in het systemen, moeten deze foutmeldingen worden genoteerd in een logboek ten einde snel tot een oplossing te komen.

De gegenereerde DIP moet binnen een redelijke termijn op een correcte manier worden aangeboden aan de gebruiker. De gebruiker moet het bestand of bestanden die werden aangevraagd (en die hij dus verwacht) ook krijgen. Bij ontvangst krijgt hij ook een bevestiging hiervan. Indien dit niet het geval is, omwille van bepaalde beperkingen of omdat een fout is opgetreden bij de generatie van het document

bijvoorbeeld, moet de gebruiker hiervan op de hoogte worden gesteld. De manier waarop een DIP wordt opgebouwd in relatie tot de oorspronkelijke AIP moet aangetoond kunnen worden. Deze informatie moet toegankelijk zijn.

C. De *designated community* en de bruikbaarheid van informatie⁸

Een *designated community* (DC) wordt in het OAIS-model omschreven als een identificeerbare groep van mogelijke gebruikers die een bepaalde set van informatie zouden moeten begrijpen. De DC kan uit verschillende gebruikersgroepen bestaan. Naar alle waarschijnlijkheid bestaat ze zowel uit de klanten van het depot als uit archiefvormers. De noden en interessegebieden verschillen al naar gelang over welke groep we het hebben en met beide moet rekening gehouden worden. Om de limieten vast te leggen van de dienstverlening, worden de verwachtingen van het gebruik van de opgenomen informatie objecten, d.i. *Content Information* en PDI, duidelijk omschreven. Op die manier wordt een basis geboden waarop de reikwijdte van de aangeboden diensten, die nodig zijn om te voldoen aan de informatieve verwachtingen van de gebruikers, kan worden gebaseerd. Uiteindelijk moet de *content information* en PDI begrepen kunnen worden door een welbepaalde DC zonder externe hulp. Nood aan experts die de inhoudelijke informatie zeer goed kennen verhoogt de kostprijs en vermindert de tijd dat data kan bewaard blijven zonder ze te verliezen. Door de (verschillende groepen binnen de) DC te omschreven, en deze omschrijving toegankelijk te maken, kan aan gebruikers duidelijk gemaakt worden wat ze kunnen verwachten.

Documentatie

Een duidelijke definitie van de DC omvat: wie er in zit, wat de basisvereisten op het vlak van kennis zijn, welk niveau van dienstverlening wordt aangeboden, ... Bijvoorbeeld, om het digitaal archief van het Vlaams Parlement te valoriseren, zal een Nederlandstalig archiefloket online kunnen geraadpleegd worden, waarin PDF-bestanden te vinden zijn. Van de gebruiker wordt verwacht dat hij de Nederlandse taal machtig is, met een webbrowser kan werken en een PDF-reader heeft geïnstalleerd op zijn computer. Dit kan meegedeeld worden aan de DC door een *disclaimer* op de website te plaatsen (met misschien zelfs een link naar Adobe waar de reader kan gedownload worden).

⁸ RLG, An audit checklist for the certification of trusted digital repositories. Draft for public comment. Mountain View: Research Libraries Group, 2005, pp. 32-38

Beschrijvende metadata geschikt voor de designated community

Een set van minimum vereisten voor metadata wordt omschreven. Deze is gebaseerd op de types van vragen die een archief kan verwachten. Hierbij zal uitgegaan worden van een soort van algemene deler, omdat het onmogelijk is voor elke soort bevraging aangepaste metadata voor te leggen. De vereisten moeten dus overeenkomen met de verwachtingen van de gemiddelde archiefgebruiker.

Gebruik en bruikbaarheid

Het depot omschrijft op welke manier de openbaarheid van de archieven wordt geregeld en deelt dit mee aan de DC. Op die manier wordt verduidelijkt wat kan en wat niet kan, of kopies kunnen bekomen worden van stukken en aan welke prijs, welke stukken worden ingezien en waar, ... Dit kan verschillen naar gelang de subgroep van DC. Vervolgens worden alle aanvragen tot toegang bijgehouden en opgevolgd. Afspraken hier rond moeten worden afgedwongen, zowel in de negatieve (zonder autorisatie geen toegang) als in de positieve (indien geautoriseerd moet toegang effectief verleend worden) zin.

Nagaan van begrijpelijkheid

Het hoofddoel van een (e-)depot is de preservatie van informatie, zodat deze later kan hergebruikt worden. Het depot heeft de verantwoordelijkheid er voor te zorgen dat ze de *content information* en PDI verkrijgt, beheert en beschikbaar maakt. Dit in een vorm die toelaat dat het digitale object begrijpelijk en bruikbaar is. Het depot legt daarom een gedocumenteerd proces vast om de begrijpelijkheid na te gaan van de *content information* en PDI.

D. Technologieën en technische infrastructuur⁹

Dit deel behandelt systemen, technologieën en technische infrastructuur die nodig is om AIP's op de lange termijn te bewaren. Met systeem wordt niet alleen IT systemen bedoeld, maar ook rampenplannen en systemen die de terugslaan op de handelingen van mensen.

Systeem infrastructuur

Een digitaal depot moet gebouwd zijn op stevige funderingen om betrouwbaar te kunnen zijn. Het gebruikte *operating system* en andere essentiële infrastructurele software moet goed onderbouwd zijn en onderhouden worden. Als eerste veiligheidsmaatregel in het geval van

⁹ RLG, An audit checklist for the certification of trusted digital repositories. Draft for public comment. Mountain View: Research Libraries Group, 2005, pp. 39-44

falen van het systeem moeten voor alle platformen backup functies ingebouwd worden. Deze functies moeten voldoende garanties geven dat het geheel van de diensten van het depot en opgeslagen data niet verloren gaan. Een tweede maatregel is het bijhouden van veiligheidskopies (niet andere versies) van de AIP's. Hoeveel kopies worden gemaakt en waar ze worden opgeslagen moet vastgelegd worden. De verschillende kopies moeten gesynchroniseerd worden. Door de synchronisatie worden alle kopies aangepast, wanneer bepaalde veranderingen worden uitgevoerd.

Data corruptie of verlies moet via efficiënte mechanismen gedetecteerd worden en stappen moeten ondernomen worden om gecorrumpeerde data te restaureren en verloren data te vervangen. Zowel de detectie als de ondernomen stappen moeten gedocumenteerd worden en doorgegeven aan de hogere instanties binnen de organisatie.

In het geval dat bepaalde veranderingen worden doorgevoerd bij de processen inzake data management, opname, opslag, toegankelijkheid en veiligheid, moet dit kunnen vastgesteld worden via gedocumenteerde processen. Zo kunnen de effecten van bepaalde veranderingen worden nagegaan. Essentiële veranderingen in het systeem moeten eveneens nagegaan kunnen worden en getest. Voor de veiligheid van het systeem te verhogen moet het geheel *up-to-date* blijven en bijgevolg de laatste *security fixes* ontvangen.

Aangepaste technologie

Een digitaal depot beschikt over de benodigde hard- en software om aan de noden van de DC te kunnen voldoen. Procedures worden opgesteld om op te volgen en aan te geven wanneer de beschikbare technologie moet aangepast worden. Aanpassingen kunnen nodig zijn omwille van de groeiende hoeveelheid data die moet verwerkt worden, evolutie van de technologische omgeving of omdat de behoeftes van DC veranderd zijn.

Veiligheid

Een depot moet bewust blijven van haar omgeving. Een analyse van de beschikbare data, systemen, personeel, fysieke site, ... zijn niet overbodig. De rollen, verantwoordelijkheden en autorisaties van elke medewerker worden duidelijk afgelijnd. Op die manier blijft het duidelijk wie, bijvoorbeeld, gebruikers mag toevoegen in een systeem, metadata moet controleren, ...

In navolging van deze systematische analyse worden processen geïmplementeerd die, in geval van nood, voor een oplossing kunnen zorgen bij alle gekende (en liefst ook ongekende) veiligheidsproblemen. Hiernaast wordt ook een rampenplan opgesteld, dat op regelmatige tijdstippen wordt getest.

1.2 Implicaties voor het archief van het Vlaams Parlement

A. Organisatie

De missie van de archiefdienst is het bewaren, ontsluiten, ter beschikking stellen en valoriseren van de archieven van het Vlaams Parlement. In het licht van de digitale bescheiden verandert deze opdracht niet. De missieverklaring van het Vlaams Parlement stelt dat ze het centrum wil zijn “van een democratisch Vlaanderen en de burgers nauw bij de politieke besluitvorming betrekken, en streeft daartoe naar maximale transparantie en optimale communicatie. Het wil een modern parlement zijn dat slagvaardig, vooruitziend en dynamisch werkt.”¹⁰ In dit opzicht werden in het beleidsplan voor de periode van 2004 tot 2009 verschillende elementen opgenomen om het informeren van de burger te verbeteren.¹¹ Daar sinds de tweede helft van de jaren negentig meer en meer digitaal wordt gewerkt, is de uitbouw van een toegankelijk betrouwbaar digitaal depot een vereiste.

Tussen 1971 en 2001 kende België een evolutie van een unitair naar een federaal model door de uitvoering van vijf staats hervormingen. Ons land werd onderverdeeld in verschillende deelstaten met elk eigen bevoegdheden. Deze werden bepaald door de verschillende visies op de indeling van ons land. Vlaanderen wou in de eerste plaats garanties in verband met de culturele ontvoogdingsstrijd. De Vlamingen vroegen daarom meer bevoegdheid over zaken die te maken hebben met de afdwinging van het gebruik van de eigen taal. Daarom lag de nadruk op de uitbouw van een Vlaamse, Franse en Duitse gemeenschap. De Walen wilden voornamelijk een eigen sociaal-economisch beleid kunnen voeren en vroegen daarom een opsplitsing in gewesten. Het resultaat van de onderhandelingen zorgde voor het typische Belgische compromis. Het land werd zowel in gewesten als in gemeenschappen onderverdeeld, met elk een eigen regering en parlement. In Vlaanderen werden deze onmiddellijk samengevoegd tot één deelstaat.¹²

Het bestaansrecht van het Vlaams Parlement wordt dus gegarandeerd door de staats hervorming. Om haar werking te financieren krijgt de Vlaamse overheid middelen van de federale overheid gebaseerd op een

¹⁰

Missieverklaring, <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/HoeWerktHetVlaamsParlement/Missie/MissieVerklaring.html>, 20/11/05

¹¹ De Batselier, N. *Een slagvaardig Vlaams Parlement, 2004 – 2009. 3^e Beleidsplan*. Februari 2005, pp. 17-21, <https://intranet.vlaamsparlement.be/docs/portlet/p615dp26lijn487.doc>, 12/07/06

¹² Voor een uitgebreid overzicht, zie oa.: Witte, E.; Craeybeckx, J.; Meynen, A. *Politieke geschiedenis van België van 1830 tot heden*. Brussel: VUB Press, 1997; Platel, M. *Communautaire geschiedenis van België, van 1830 tot vandaag*. Leuven: Davidsfonds, 2004; Vande Lanotte, J.; Bracke, S. & Goedertier, G. *België voor beginners. Wegwijs in het Belgische labyrint*. Brugge: Die Keure, 2000

jaarlijks te herziene verdeelsleutel.¹³ Op basis van een dotatie van de Vlaamse Gemeenschap worden de gewone werkingskosten van het Parlement gefinancierd.¹⁴ De begroting van de Vlaamse overheid werd dit jaar voor de zevende maal op rij afgesloten met een positief begrotingssaldo. “Het is dankzij deze begrotingsoverschotten dat de schuld verder afgebouwd kan worden en dat Vlaanderen tegen 2009 schuldenvrij is.”¹⁵

Continuïteit van de instelling wordt door het federale model verzekerd. Zelfs bij een terugkeer naar een unitaire staat, zullen de uitgevoerde functies worden overgenomen door de federale overheid. In dat geval wordt het archief ook overgeheveld.

Het Archief en de Biografische Dienst (Mevr. Van Ongevalle en Mr. Van Paemel) valt onder de Directie Informatie, waardoor ze niet beschikt over eigen middelen. Jaarlijks worden de voorziene uitgaven, op de korte en lange termijn, verdedigd. In de *nota van de Directieraad aan het Uitgebreid Bureau* (17/11/05) nemen Mevr. Van Ongevalle en Mr. De Graeve (Directeur Informatie) de verantwoordelijkheid op om “het lange termijn geheugen van de democratische werking in Vlaanderen ... te vrijwaren en te valoriseren”, omdat dit kan “bijdragen tot de visibiliteit en verdere promotie van het Vlaams Parlement en haar leden.” De archiefdienst vraagt hiertoe “het nodige mandaat, evenals de medewerking van het Secretariaat-generaal, de betrokken Directies en de ondersteuning van de Directie Informatie.”¹⁶ Tijdens de vergadering van het Vast Bureau van 19 december 2005 verklaarde het Bureau zich eenparig akkoord met de strategische nota (2005-2009) en met het actieplan voor het zittingsjaar 2005-2006. Bovendien werd besloten “aan het secretariaat-generaal en de directies opdracht te geven de archivaris te betrekken in alle projecten van geautomatiseerd en digitaal documentbeheer en dit met de bedoeling de archivering van de door deze systemen gegenereerde documenten en gegevens veilig te stellen” en “de archivaris de nodige medewerking te verlenen in het vastleggen van selectiecriteria en bewaartermijnen.”¹⁷ Zoals hier boven wordt aangetoond engageert het Vlaams Parlement als instelling in het algemeen, en de archiefdienst in het bijzonder, zich op formele wijze tot de uitbouw van een modern archief. Intussen heeft het concept van het TDR ingang gevonden en zal de uit te werken digitale archiveringsstrategie voor het multimediaal archief van het Parlement

¹³ De Belgische Grondwet, Titel V, Art. 171, Art. 175, Art. 177, http://www.senate.be/doc/const_nl.html#t5, 20/07/06

¹⁴ De werkingsbegroting van het Vlaams Parlement, <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/HoeWerktHetVlaamsParlement/MeerInformatieOverHetVlaamsParlement/BudgetFinanciering.html>, 10/02/06

¹⁵ Van Mechelen, D. *De Vlaamse Begroting 2006*. Brussel: Vlaamse Regering, 2006, p. 3, http://www2.vlaanderen.be/ned/sites/financien/Info/Vlaamse_begroting_2006.pdf, 14/07/06

¹⁶ Directie Informatie, Archief en Biografische Dienst, *Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement*. 17 november 2005

¹⁷ Notulen van de vergadering Vast Bureau, 19 december 2005, tiende blad

geplaatst worden binnen het totaalconcept van het digitaal depot conform het OAIS-model.¹⁸

B. Functies, processen en procedures

Het beheer van archiefbescheiden in de brede zin van het woord, begint met het opstellen van concrete procedures en richtlijnen. Indien archiefvormers (en gebruikers) duidelijk weten wat van hun verwacht wordt, wordt de kans op fouten verminderd. Om die reden is het aangewezen criteria voor de creatie, benaming, selectie en overdracht van bestanden te bepalen.

“Voor een groot deel van het archief van het Vlaams Parlement zijn geen selectiecriteria vastgelegd en ontbreken goedgekeurde bewaartermijnen.” Hierdoor vinden ongeselecteerde overdrachten plaats, met als gevolg dat te vernietigen archief wordt beheerd. Dat is een kostelijke zaak omdat het geheugen opneemt, servers belast en de toegankelijkheid en raadpleegbaarheid bemoeilijkt. Om aan efficiëntie selectie te kunnen doen moet een archiefbeheersplan worden opgesteld. In functie daarvan zullen “volgende taken uitgevoerd worden:

- contact opnemen met de verschillende archiefvormers, de werkprocessen in kaart brengen en de archiefstukken die zij genereren
- ontwerp in samenspraak met archiefvormers – en desgevallend juridische dienst – van selectielijsten; deze selectielijsten worden ter goedkeuring voorgelegd aan de Directieraad en het Bureau en verwerkt in het archiefbeheersplan
- het archiefbeheersplan kenbaar maken
- het op geregelde tijdstippen actualiseren van het archiefbeheersplan
- de vernietigingsprocedure voor het archief uitwerken en kenbaar maken (onder meer uitwerken van een aanvraagformulier voor archiefvernietiging)¹⁹

De overdracht wordt geregeld aan de hand van het archiefreglement: “alle statische archiefbescheiden van het Vlaams Parlement (die geen direct administratief nut meer hebben) worden zo snel mogelijk aan de archiefdienst overgedragen.” De voorwaarden voor een overdracht, in goede en geordende staat, zijn echter enkel van toepassing op het papieren archief. Enkel de eerste twee elementen kunnen toegepast worden op het digitaal archief: “de archivaris wordt vooraf op de hoogte gebracht” en voor elke dienst stelt de overdragende dienst een beknopte overdrachtslijst

¹⁸ Van Ongevalle, H. Digitaal depot Vlaams Parlement. Conceptnota voor een plan van aanpak. Onuitgegeven, Mei 2006, p. 2

¹⁹ Directie Informatie, Archief en Biografische Dienst, Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement. 17 november 2005, p. 6

op.”²⁰ Om deze reden is voorzien in de uitwerking van een richtlijn voor overdracht van digitale objecten door de archiefdienst, in samenwerking met de Directie Informatica. “Deze richtlijn legt vast onder welke voorwaarden bestanden aan het Archief worden overgedragen. De richtlijn houdt rekening met:

- aan te leveren formaat
- specificatie voor de dragers
- geordende staat
- benaming van de bestanden
- aan te leveren metagegevens
- selectie (moet zijn toegepast conform het Archiefbeheersplan; in concreto moeten gegevens geselecteerd zijn of moet de selectie na afloop bewaartijd makkelijk toepasbaar zijn)
- authenticiteit
- openbaarheid”²¹

Een ontwerp van een tussentijdse richtlijn – in afwachting van de effectieve implementatie van het TDR-concept – werd aan de Directie Informatica ter discussie overgemaakt.²²

De concrete archiefwerving zal gebeuren aan de hand van een archiefbeheersplan, afgesproken overdrachtstermijnen en gestandaardiseerde elektronische overdrachtsformulieren. De overdrachtsprocedure zal geautomatiseerd worden en overdrachtstermijnen worden vastgelegd in onderling overleg met de archiefvormers.²³

Na overdracht zullen de overgedragen gegevens in quarantaine worden geplaatst. Het SIP wordt dan gecontroleerd op virussen, vervolgens op volledigheid, authenticiteit, geordende staat, integriteit, etc. Dit zal gebeuren conform de op te stellen richtlijn. In die richtlijn mag niet vergeten worden duidelijk af te bakenen hoe lang deze controles mogen duren. Vernietiging van de “originele” kopie, die zich bij de archiefvormer bevindt, kan uit veiligheidsoverwegingen pas na de transformatie van SIP naar AIP. Door het vastleggen van een minimale tijdsperiode van controle, wordt vermeden dat dit gedeelte vlug vlug wordt afgehaspeld, waardoor de veiligheid en de goede werking van het depot in het gedrang komen. Invoeren van een maximale tijdsbesteding vermijdt het ontstaan van een *twilight zone* waarin de in quarantaine geplaatste stukken een schaduwbestaan leiden. Op die manier weet de archiefvormer ook hoe lang

²⁰ Archiefreglement van het Vlaams Parlement, p. 1, <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/Pdf/20042005/Archief/Archiefreglement.pdf>, 20/11/05

²¹ Van Ongevalle, H. Digitaal depot Vlaams Parlement. Conceptnota voor een plan van aanpak. Onuitgegeven, Mei 2006, pp. 6-7

²² Naar analogie van de richtlijn van het archief van de Raad van Europa.

²³ Directie Informatie, Archief en Biografische Dienst, Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement. 17 november 2005, p. 7

het kan duren alvorens de stukken, die hij na overdracht nog heeft bijgehouden, mag vernietigen. Door op deze transparante manier te werken, wordt de efficiëntie verhoogd en worden servers niet nodeloos belast.

Bij de controle van het SIP wordt ook de aangeleverde metadata nagekeken. Een gedeelte kan al automatisch gegenereerd worden. Ideaal gebeurt deze toekenning volledig automatisch. In de praktijk is dit meestal niet mogelijk. Een deel zal nog steeds handmatig moeten ingegeven worden. Voor vele mensen kan dit soms te veel zijn, zelfs als het maar twee handelingen betreft. Controleren is dus aangewezen.²⁴ Waar nodig moeten de ontbrekende gegevens bijgevoegd worden. In de *conceptnota voor een plan van aanpak* wordt vermeld welke gegevens in- en aangevuld moeten worden. “Registratie ... van de metadata: beschrijvend (maken mogelijk: identificatie, interpretatie, authenticatie, vinden); administratieve of beheersmetadata (autorisatie, eigendom, formele herkomst, verantwoording van beheersactiviteiten); technische (software, hardware, opslagformaat).” Welk metadatamodel toegepast zal worden, moet nog worden bepaald. Om een wildgroei van gebruikte standaarden en modellen tegen te gaan, stelt Mevr. Van Ongevalle dat het aangewezen is met één model te werken, dat toegepast kan worden op alle objecten. Aangezien gewerkt wordt in een hybride omgeving (papieren en digitale objecten) is integratie met het papieren archief noodzakelijk.²⁵

Voor het terugvinden van bescheiden, zijn niet alleen de correcte en volledige invulling van metadata van belang. Het instellen van een duidelijke klasseringsstructuur, gekoppeld aan een unieke en duidelijke naamgeving is eveneens aangewezen. In een genetwerkte omgeving is het belangrijk om duidelijke richtlijnen op te stellen voor de benaming van documenten en mappen. Bij het ontbreken van concrete afspraken leidt dit tot een kleine chaos, waarin niemand buiten de archiefvormer in kwestie zijn weg in kan vinden. Het mooiste voorbeeld is de “archivering” van de website, waar “ronkende” benamingen als 001 en doc1 werden toegekend. Aan de hand van het archiefbeheersplan zal een virtuele structuur opgesteld worden. Hieraan gekoppeld krijgen de verschillende onderdelen unieke codes waaraan de benamingen van documenten gekoppeld zullen worden. Het is een beetje oneerlijk van mij hier de indruk te wekken dat de website door onkunde slecht gearchiveerd werd. De eigenlijke bedoeling was te voorzien in een soort van backup. Backups zijn van nature enkel gericht op de zeer korte termijn. Maar, door het ontbreken van een concrete procedure voor het archiveren van de website, zijn dit de enige bestanden die werden bewaard en overgedragen aan het archief. Intussen werden met ICT en de webmaster de nodige afspraken gemaakt om de huidige versie van de site te archiveren volgens de strategie van het bewaren van de bronbestanden en het filmen van een

²⁴ Boudrez, F. Digitaal archiveren van archiefdocumenten. Casus: e-mail. Antwerpen : Stadsarchief Antwerpen, Onuitgegeven lezing, 6 maart 2006

²⁵ Van Ongevalle, H. Digitaal depot Vlaams Parlement. Conceptnota voor een plan van aanpak. Onuitgegeven, Mei 2006, p. 7

surfsessie. In de nieuwe versie van de site worden conform de Dublin Core standaard de metadatavelden geïntegreerd.

Alvorens over te gaan tot de overzetting van een SIP naar een AIP, moet de digitale inhoud geanalyseerd worden. Samen met de archiefvormers en ICT moet nagekeken worden wat de essentiële eigenschappen (*significant properties*²⁶) zijn van de archiefstukken en bijgevolg aan welke authenticiteitseisen het gearchiveerde object moet voldoen. Bij het archiveren van e-mail, bijvoorbeeld, is het van belang om de gekoppelde bijlagen te bewaren en te kunnen openen. Bij Excel-documenten daarentegen zal het in vele gevallen niet noodzakelijk zijn om de onderliggende wiskundige structuur mee op te nemen. Duidelijke afspraken met zowel de archiefvormers als de Directie Informatica zijn dus een noodzaak.

Op het vlak van toegankelijkheid is het archiefreglement zeer duidelijk: alle aanvragen worden aan de archivaris gericht, die de vindplaats opzoekt, (na verificatie van de autorisatie) de stukken uit de archiefbewaarplaats haalt en achteraf terugplaatst. Aan ambtenaren mogen de stukken uitgeleend worden, derden kunnen ze enkel in de leeszaal van de informatheek van het Parlement raadplegen. De reden waarom toegang enkel door middel van de archivaris (of de archiefmedewerker) kan verzekerd worden, ligt in het feit dat “een groot deel van het archief van het Vlaams Parlement al dan niet tot op dossierniveau ontsloten [is] aan de hand van inventarissen opgemaakt in Excel, Word of Access. Een coherent, geautomatiseerd en met de andere archieftaken geïntegreerd systeem van ordenen en beschrijven ontbreekt. Deze diversiteit maakt het zoeken omslachtig en beperkt het zoekproces tot het archiefpersoneel.”²⁷

Het takenpakket van e archiefdienst wordt gedeeltelijk ondersteund door verschillende manieren van geautomatiseerde gegevenswerking: MS Office, Proteus, Mediargus²⁸, Arbor Audio, Efrat Future Technology. De bestaande situatie kent een aantal tekortkomingen:

- “verspreiding van de inventarisatiegegevens over verschillende pakketten met soms ingewikkelde zoektochten tot gevolg
- een niet gestandaardiseerde en consistente ontsluiting

²⁶ Het concept van *significant properties* werd door Cedars en the National Library of Australia uitgewerkt. The Cedars Project Team, *The Cedars Project Report*. April 1998-March 2001. June 2001, pp. 14-15, <http://www.leeds.ac.uk/cedars/OurPublications/cedarsrepmar01exec.html>, 20/07/06

²⁷ Directie Informatie, Archief en Biografische Dienst, *Nota van de Directieraad aan het Uitgebreid Bureau*. Opdracht van het archief in een modern Parlement. 17 november 2005, p. 4

²⁸ Mediargus is een databank van Vlaamse en Nederlandse berichtgeving, die toelaat “om het brede scala van artikels op een snelle en gebruiksvriendelijke manier te raadplegen en op een correcte manier te verspreiden.” <http://www.mediargus.be/NL/index.asp>

- niet meer onderhouden pakketten waarvan de informatie voor gebruik niet langer meer is gekend (Efrat Future Technology²⁹) met gegevensverlies tot gevolg
- ontbreken van een globaal overzicht en deelloverzichten van het archievenbestand
- ontbreken van beheersgegevens
- een niet geautomatiseerde verwerking van aanvragen, uitleningen, archiefoverdrachten en selecties met manueel bijhouden van allerhande lijsten tot gevolg
- beheer van het digitale archief aan de hand van een op termijn niet meer beheersbare mappenstructuur
- uiterst beperkte mogelijkheden in het toekennen van autorisaties; onmogelijk om autorisatierechten toe te kennen op basis van het type-gebruik en aard van het archiefstuk (alles-of-niets-autorisaties)

Een globale visie op de automatisering van de dienst moet uitgewerkt worden rekening houdend met volgende zaken in de uitwerking van een oplossing:

- integratie binnen bestaande context (mogelijkheden binnen bestaande context moet onderzocht worden)
- mogelijke koppeling met bestaande gegevensbeheerssystemen en systemen voor digitaal documentbeheer
- mogelijke integratie van de diverse archiefbeheersfuncties
- een gebruiksvriendelijke gebruikersinterface/ publiekscatalogus
- een open en modulair opgebouwd systeem
- multimediale ontsluiting
- koppeling van gegevens aan ordeningsplannen of classificatiesystemen
- flexibiliteit in de gegevensverwerking
- mogelijkheid tot integratie van de standaarden voor beschrijving
- mogelijkheden in het verzamelen en verwerken van beheersgegevens
- opmaak en uitprint van overzichten en mogelijke koppeling met Word en Excel³⁰

Om de toegankelijkheid te verhogen zal een strategie uitgewerkt worden voor de uitbouw van “een geautomatiseerd systeem voor inventariseren en terugzoeken van archief, en de opmaak van archieftoegangen:

²⁹ Efrat Future Technology was een onderaannemer van het Amerikaanse Comverse Technology Inc. Comverse heeft sterk geïnvesteerd in dit Israëlische bedrijf en het uiteindelijk, na een meerderheidsaandeel te hebben verzekerd van 99% het bedrijf omgedoopt tot een afdeling van het moederbedrijf onder de naam Comverse Network Systems. Ze leveren software en systemen gericht op *network-based enhanced services*, voornamelijk op het vlak van de telecommunicatie. http://www.comverse.com/company_profile, <http://www.cmvt.com/news2.asp?id=3&newsid=343>, <http://company.monster.com/comverse/>, 23/07/06

³⁰ Directie Informatie, Archief en Biografische Dienst, Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement. 17 november 2005, p. 4

- zoeken naar een geautomatiseerde oplossing van de archiefbeheerstaken ‘ordenen, ontsluiting en archiefraadpleging’
- opmaak van een archievenoverzicht ...
- de ontsluiting van het archief optimaliseren aan de hand van geschikte en gewenste toegangen (waaronder een globaal archievenoverzicht)
- onderzoeken in hoeverre op het vlak van beschrijving en ontsluiting kan samengewerkt worden met andere parlementaire instellingen of Europese projecten
- implementeren van de internationale standaarden voor beschrijving en uitwisseling van gegevens (ISAD, ISAAR, ...)”³¹

Na ontsluiting moeten archieven uiteraard ter beschikking worden gesteld opdat het Vlaams Parlement al een deel van haar streven naar maximale transparantie en optimale communicatie kan verwezenlijken. Het overgrote deel van de parlementaire documenten worden al digitaal ter beschikking gesteld (i.e. via de website). Het archief omvat meer dan dat. Doordat niet bekend wordt gemaakt welke archieven aanwezig zijn, wordt het historisch bezit van het Parlement niet ten volle benut. Een tweede probleem is het ontbreken van een eenvoudige en gebruiksvriendelijke interface voor de raadpleging van de digitale bescheiden.

Een strategie moet (en zal) uitgewerkt worden om de werking van de dienst te optimaliseren. De krachtlijnen zijn de volgende:

- “de werking van de dienst zowel als de aard en de inhoud van de door de dienst beheerde stukken kenbaar maken, daarbij rekening houdend met het onderscheid in gebruik door de eigen leden en medewerkers van het Vlaams Parlement enerzijds en belangstellende buitenstaanders anderzijds (intranet, internet, een brochure, ... en evolueren in de richting van een digitaal archiefloket)
- opportuniteit van een klantvriendelijke geautomatiseerde ‘gebruikersinterface of publiekscatalogus’ welke online ter beschikking wordt gesteld onderzoeken en desgevallend implementeren
- het archiefreglement up to date houden en kenbaar maken
- een minimale infrastructuur voor raadpleging door externen voorzien”³²

³¹ Directie Informatie, Archief en Biografische Dienst, Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement. 17 november 2005, p. 8

³² Directie Informatie, Archief en Biografische Dienst, *Idem*, p. 9

C. De *designated community* en de bruikbaarheid van informatie

“Content creators, system developers, custodians, and future users are all potential stakeholders in the preservation of digital materials, and this complicates the determination of responsibilities - who, when, and for how long.”³³ Concreet vertaald naar het Vlaams Parlement, zijn de belanghebbenden³⁴ de archiefdienst, de Directie Informatica en de klanten. Deze laatste omvatten “de leden van het Parlement, het Algemeen Secretariaat, de fracties, de politieke medewerkers, de paraparlementaire medewerkers, de andere Belgische overheden, de onderzoeksinstellingen (universiteiten, hogescholen, onderzoeksinstellingen en –centra, archieven, ...), de pers, het brede publiek.”³⁵ Duidelijk zal omschreven moeten worden wie wanneer toegang heeft tot welke stukken. Er moet rekening gehouden worden met beperkingen op de openbaarheid, de vertrouwelijkheid van bepaalde stukken, privacywetgeving, auteursrechten, etc. Zelfs intern bestaan verschillen op geautoriseerde toegang tot bepaalde stukken. Het is bovendien mogelijk dat met de openstelling van een archiefserver, de bescheiden sneller zullen overgedragen worden dan nu het geval is. Om deze redenen zal de beschrijvingen van autorisaties sterk verfijnd moeten worden. Een eis die gesteld zal worden aan het archiefbeheerssysteem is flexibiliteit in het parametriseren van gebruikersgroepen en het toekennen van vertrouwelijkheidscodes tot op een zeer laag niveau (*in casu* tot op het individuele stuk).

De modaliteiten betreffende raadpleging en openbaarheid van de opgenomen bestanden worden geregeld aan de hand van het archiefreglement. “Bij gebrek aan ... een archiefdecreet waarin de openbaarheid van de archieven van wetgevende instellingen geregeld wordt, is de openbaarheid van de archiefbescheiden van het Vlaams Parlement als volgt geregeld, in overeenstemming met het bureaubesluit van 5 mei 1997 over de administratieve aspecten van de nieuwe politieke cultuur:

- Archiefbescheiden betreffende alle plenaire vergaderingen en alle openbare commissievergaderingen zijn openbaar zodra ze aan de archiefdienst overgedragen en geïnventariseerd zijn

³³ RLG/OCLR, Trusted digital repositories. Attributes and responsibilities. An RLG-OCLR Report. Mountain View: RLG, 2002, p. 18

³⁴ “The word “stakeholder” was first recorded in 1708 as “a person who holds the stake or stakes in a bet”; the current definition is “a person with an interest or concern in something” (Bisset, personal communication, 1998). Freeman (1984, p. vi) defines a stakeholder as “any group or individual who can affect, or is affected by, the achievement of a corporation’s purpose.” Ramirez, R. Concept: Society. Chapter 5: Stakeholder analysis and conflict management. Canada: IDRC Books, 2005, http://www.idrc.ca/en/ev-27971-201-1-DO_TOPIC.html#to205ftn01, 18/07/06

³⁵ Directie Informatie, Archief en Biografische Dienst, Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement. 17 november 2005, p. 1

- Archiefbescheiden betreffende besloten commissievergaderingen die aan de archiefdienst overgedragen en geïnventariseerd zijn, worden openbaar 10 jaar na datum van de archiefbescheiden
- Archiefbescheiden betreffende besloten commissievergaderingen die aan de archiefdienst overgedragen en geïnventariseerd zijn, worden openbaar 10 jaar na datum van de archiefbescheiden
- Archiefbescheiden van besloten vergaderingen van onderzoekscommissies, van de commissie Vervolgen en van soortgelijke commissies worden openbaar 50 jaar na datum van de archiefbescheiden
- Archiefbescheiden betreffende de interne werking van het Vlaams Parlement zijn openbaar zodra ze aan de archiefdienst overgedragen en geïnventariseerd zijn

De openbaarheid van de archiefbescheiden die betrekking hebben op (kandidaat)personeelsleden en het personeelsbeleid van het Algemeen Secretariaat van het Vlaams Parlement of van de aan het Vlaams Parlement verbonden instellingen, of op overheidsopdrachten van het Vlaams Parlement of van de aan het Vlaams Parlement verbonden instellingen, wordt geregeld door de wetgeving inzake openbaarheid van bestuur

Andere archiefbescheiden waarvan de openbaarmaking afbreuk zou doen aan de bescherming van de persoonlijke levenssfeer van personen, worden slechts openbaar na het overlijden van de persoon, of eerder, indien de betrokkene met de openbaarmaking instemt³⁶

Volgens de beslissing van het Uitgebreid Bureau van 4 juli 2005 worden “de gearchiveerde publieke audiobestanden van de commissievergaderingen ter beschikking gesteld van de leden van de parlementaire gemeenschap, wanneer zij daarom vragen.”³⁷ Voor derden geldt dat alleen de bestanden van de openbare vergaderingen (of van besloten vergaderingen maar dan na afloop van een periode van tien tot vijftig jaar) kunnen beluisterd worden in de leeszaal van de Informatheek.

De openbaarheid van bestuur wordt geregeld door het *decreet van 26 maart 2004 betreffende de openbaarheid van bestuur (BS 01/07/2004)*.³⁸ Deze is van toepassing op het Vlaams Parlement en de eraan verbonden instellingen.³⁹ Artikel 7 bepaalt dat “de instantie verplicht [is] aan ieder natuurlijk persoon, rechtspersoon of groepering ervan die erom verzoekt, de gewenste bestuursdocumenten openbaar te maken door er inzage in te verlenen, er uitleg over te verschaffen of er een afschrift van te

³⁶ Archiefreglement van het Vlaams Parlement, pp. 2-3

<http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/Pdf/20042005/Archief/Archiefreglement.pdf>, 20/11/05

³⁷ Vergadering Uitgebreid Bureau van 4 juli 2005, Tiende blad

³⁸ <http://user.online.be/fschram/wet109.html>

³⁹ Art. 4, § 1

overhandigen.” Artikels 11 tot en met 15 bepalen de uitzonderingen op de wet.

D. Technologieën en technische infrastructuur

Het onderzoek naar een degelijke technische ondersteuning wordt op dit moment uitgevoerd. De effectieve uitbouw van een degelijke IT-infrastructuur heeft uiteraard pas zin, nadat het geheel van de noden en wensen in kaart zijn gebracht.

Het marktonderzoek is opgestart en een ontwerp van traject ligt ter onderhandeling op de Directie ICT.⁴⁰

| | Omschrijving | Uitvoering door | Timing (deadline) |
|------------|---|------------------------|------------------------------|
| I | Selectie | | |
| 1.1 | Opstellen van vragenlijsten | Archief | December 2005 |
| 1.2 | Bevragen van de diensten – institutioneel onderzoek | Archief | 2006 |
| 1.3 | Uitwerken van een archiefbeheersplan | Archief | December 2006 – Januari 2007 |
| II | ICT: Vooronderzoek digitale archivering | | |
| 2.1 | Opstellen van een vragenlijst | Archief | Maart 2005 |
| 2.2 | In kaart brengen van de technische omgeving waarbinnen de digitale objecten tot stand komen a.h.v. de bevraging van ICT en bijkomend onderzoek | Archief | Juli 2006 |
| III | Uitwerken van een richtlijn voor de overdracht van het digitaal archief naar de Archiefdienst. Specificaties m.b.t.: <ul style="list-style-type: none"> • dragers • formaten • ordening • benaming • metadata | Archief/ICT | Najaar 2006 – 2007 |
| IV | Ontwikkeling/Aankoop software voor een geautomatiseerd archiefbeheerssysteem (ABS) en de uitbouw van een digitaal depot | Archief/ ICT | 2007 |

⁴⁰ Van Ongevalle, H. Digitaal depot Vlaams Parlement. Conceptnota voor een plan van aanpak. Onuitgegeven, Mei 2006, pp. 9-11

| | | | |
|----------|---|---------------------|--------------------|
| 4.1 | Uitschrijven van de functionele vereisten waaraan het ABS moet voldoen | | |
| 4.2 | Vastleggen van het metadatamodel | | |
| 4.3 | Uitvoeren van een marktstudie – opmaken van een begroting | | |
| 4.4 | In geval aangekocht: opmaak van het lastenboek – aanbesteding | | |
| 4.5 | Implementatie | | |
| V | Uitschrijven van de uitvoeren processen binnen het digitaal depot – uitwerken van een systeem (incl. verantwoordelijken aanwijzen) – opmaken van een begroting | Archief/ ICT | 2008 - 2009 |
| 5.1 | Beheer: <ul style="list-style-type: none"> • toegangscontrole en beveiliging • audit trail • back up • authenticiteit • configuratie (installeren en actueel houden van de middelen die nodig zijn voor het in stand houden van het depot) | | |
| 5.2 | Opname: <ul style="list-style-type: none"> • aanleveren • technische controle aanwinst • administratieve controle aanwinst • inhoudelijke controle aanwinst • registreren aanwinst • authenticiteit | | |
| 5.3 | Bewerken (inhoudelijk): <ul style="list-style-type: none"> • extraheren en opslaan metadata • toevoegen metadata • uitvoeren selectie op basis van het archiefbeheersplan | | |
| 5.4 | Bewerken (technisch): | | |

| | | | |
|-------------|--|----------------------|--------------------|
| | <ul style="list-style-type: none"> • uitvoeren van de bewaarstrategieën • uitvoeren van de kwaliteitscontrole van de bestanden | | |
| 5.5 | Bewaren – plaatsen op de opslagserver | | |
| 5.6 | Zoeken en raadplegen – ter beschikking stellen: <ul style="list-style-type: none"> • conform criteria voor beveiliging • conform archiefreglement en toepassing privacywetgeving en auteursrecht | | |
| VI | Uitwerken, vastleggen en bijhouden van de bewaarstrategieën per type object (incl. authenticiteit) | Archief / ICT | 2008 – 2009 |
| 6.1 | Office documenten | | |
| 6.2 | E-mail | | |
| 6.3 | Website | | |
| 6.4 | Intranet | | |
| 6.5 | Audio | | |
| 6.6 | Beeld: <ul style="list-style-type: none"> • bewegend • stilstaand | | |
| VII | Digitale mappenstructuur voor het Algemeen Secretariaat | | |
| 7.1 | Project uitwerken | | |
| 7.2 | Project implementeren | | |
| VIII | Procedures – handleiding - opleiding | | |
| | | | |

2. De Technologische omgeving

“Significant properties ... of digital objects ... affect their quality, usability, rendering, and behaviour. In an ideal world ... archives would preserve their physical and digital collections in their original form with all significant properties intact. Resource constraints and technical limitations, however, often require ... to weigh decisions about which properties to preserve...”¹

2.1 Een korte introductie

Elke overheid produceert een aanzienlijke hoeveelheid documenten (zowel op papier als digitaal),² die het bewijs vormen van een handeling, van de uitvoering van een taak of om bepaalde rechten te handhaven.³ Om in de toekomst dit doel te kunnen dienen, moet gedacht worden aan een duurzame manier van bewaring. Bij digitale documenten stellen zich een aantal specifieke problemen.

Binnen het Vlaams Parlement worden uiteenlopende digitale documenten geproduceerd. Onderzoek moet verricht worden naar de aanwezige servers, operating systems, applicaties en de website (inclusief het intranet). Elk van deze onderdelen heeft eigen specifieke eigenschappen. Om te begrijpen welke problemen zich kunnen voordoen, moet eerst een inzicht verworven worden in de samenstelling van de gebruikte applicaties en formaten en de manier waarop ze werken. Dit zal behandeld worden in de volgende paragrafen.

¹ Hedstrom, M. & Lee, C. Significant properties of digital objects: definitions, applications, implications. In: *Proceedings of the DLM-Forum 2002, @ccess and preservation of electronic information: best practices and solutions*, Luxemburg: European Communities, 2002, p. 218

² “Document. Geheel van samenhangende gegevens, vastgelegd op een of meer gegevensdragers.” den Teuling, A.J.M. (ed.) Archiefterminologie voor Nederland en Vlaanderen. 's Gravenhage: Stichting Archiefpublicaties, 2003, artikel 3

³ “Archiefstuk. Document, ongeacht zijn vorm, naar zijn aard bestemd om te berusten onder de persoon, groep personen of organisatie die het heeft ontvangen of opgemaakt uit hoofde van zijn of haar activiteiten, zijn of haar taken of ter handhaving van zijn of haar rechten.” Den Teuling, A.J.M. (ed.), *Idem*, artikel 8

2.2 De gebruikte software en formaten

A. Tekstverwerking

Omdat moderne softwarepakketten uitgerust worden met WYSIWYG (What You See Is What You Get) en GUI (Graphical User Interface), blijft de band met de papieren omgeving sterk behouden. In die zin dat de eigenlijke tekst op een ons vertrouwde manier wordt weergegeven, net zoals we die zouden kunnen aantreffen in een boek. Hiernaast echter werden in de loop van de jaren bijkomende mogelijkheden toegevoegd. Afbeeldingen en grafieken kunnen worden ingevoegd, evenals bijkomende boodschappen die enkel op het scherm zichtbaar zijn. Bovendien moet rekening gehouden worden met *intelligent agents* (automatische spellingscontrole bvb.) die op de achtergrond lopen zonder dat we het zien.

Een tekstdocument kan (buiten tekst) de volgende elementen bevatten⁴:

- Tabellen
- Formulieren
- Koppen en paragrafen
- Secties
- Opsommingstekens en genummerde lijsten
- Voet- en eindnoten
- Tekens en symbolen
- Velden en kaders
- Waarden, die automatisch worden bijgevoegd
- Grafische/ingevogde items
- Gekoppelde of geïntegreerde objecten (Object Linking and Embedding)
- Gecompileerde, grote documenten kunnen navigatie en structurele wegwijzers inhouden (index, inhoudsopgave, bibliografie, ...)
- Documenten kunnen veiligheidskenmerken bevatten
- Het uiterlijk van een document kan
 - Verschillende lettertypes en –groottes, stijlen, kleuren of ‘user defined’ paragraaf- en regelruimtes bevatten
 - Onzichtbare opmaakopdrachten bevatten voor tabinstellingen, harde returns en sectie- en kolomonderbrekingen

De interne structuur kan impliciet of expliciet worden bepaald. Een expliciete structuur wordt aangebracht met behulp van opmaakprofielen. Een impliciete structuur wordt aangebracht door handmatig de verschillende elementen die de layout bepalen aan te passen. In de meeste gevallen echter wordt een impliciete structuur

⁴ Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 3: Bewaren van tekstdocumenten. Den Haag, ICTU, 2003, p. 12

gehanteerd. Dit maakt het bewaren van het uiterlijk moeilijker. Succesvolle migraties zijn bovendien enkel mogelijk wanneer de documenten (algemeen) goed zijn gestructureerd.⁵

B. Spreadsheets

Spreadsheets (zoals Microsoft Excel) zijn rechthoekige informatietabellen⁶, die in de meeste (overheids)organisaties gebruikt worden voor het beheren van geld, activa, tijd en een groot aantal andere gegevens.⁷ Deze bevatten één of meer werkbladen (tabs). Elk werkblad bestaat uit rijen en kolommen, waarin elke cel op het snijpunt van kolom en rij een unieke identificatie krijgt. Daarnaast kan ze ingevoegde objecten bevatten (onder meer grafieken, waarin geselecteerde celgegevens grafisch worden weergegeven). Spreadsheets kunnen aan elkaar worden gekoppeld, zodat het resultaat van de ene als invoer voor de andere kan gebruikt worden. Twee lagen kunnen onderscheiden worden: de weergave laag en de onderliggende laag die de formules en celverwijzingen bevat. Beide lagen moeten worden bewaard. Indien de formules ontbreken, kan de inhoud van een document wijzigen. Elke cel kan verschillende kenmerken hebben⁸:

- De celinhoud (de weergegeven waarde van een cel)
- Een celopmaakcategorie (bvb. een getal, valuta, datum, tijd, percentage, een breuk)
- Een formule, waarmee via één of meerdere cellen de inhoud van een andere cel wordt gecreëerd
- Een aan de cel toegevoegde opmerking
- Uiterlijke kenmerken (uitlijning, lettertype, rand, patroon)

C. Databanken (databases)

Een database is een georganiseerde verzameling van gegevens (data), bewaard op een systematische wijze, opdat een computerprogramma deze gegevens kan bevragen met het doel bepaalde antwoorden te verkrijgen. Elk *record* is meestal georganiseerd in een set van data-elementen, ten einde de ophaal- en sorteermogelijkheden te verbeteren. De gegevens die gegenereerd worden als antwoord op een zoekopdracht, worden informatie waarop beslissingen kunnen gebaseerd worden. Het programma dat gebruikt wordt voor het

⁵ Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief, 2004, p. 79

⁶ "A type of application program which manipulates numerical and string data in rows and columns of cells. The value in a cell can be calculated from a formula which can involve other cells. A value is recalculated automatically whenever a value on which it depends changes. Different cells may be displayed with different formats." Spreadsheet. In: FOLDOC, Free On-Line Dictionary Of Computing, <http://www.nightflight.com/foldoc-bin/foldoc.cgi?query=spreadsheet> , 20/02/06

⁷ Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 2: Bewaren van spreadsheets. Den Haag: ICTU, 2003, p. 17

⁸ Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 2: Bewaren van spreadsheets. Den Haag: ICTU, 2003, pp. 19-20

beheren en bevragen van databanken noemen we een *database management system* (DBMS).⁹

Een database bestaat uit drie delen:

- De database zelf (de eigenlijke inhoud)
- DBMS
- De database-applicatie of toepassing¹⁰

Voor een gegeven databank bestaat een structurele beschrijving van (het type van) data elementen (en de relaties tussen de elementen) waaruit ze is opgebouwd. Deze beschrijving wordt een schema genoemd. Er zijn verschillende manieren om een schema te organiseren, i.e. om de structuur van de database te modelleren. Die manieren worden gegevensmodellen genoemd (*database of data models*).¹¹

Diverse modellen, waarop een DBMS wordt gebaseerd, zijn mogelijk¹²:

- Relatieel: De DBMS-markt wordt gedomineerd door relationele databases. Dit model stelt informatie voor in de vorm van meervoudige verwante tabellen. Relaties (*relationships*) worden voorgesteld door het gebruik van waarden die door meerdere tabellen gedeeld worden.¹³
- Object georiënteerd: *Object Oriented Programming* is een programmeerparadigma. De idee is dat een programma kan gezien worden als een verzameling van individuele eenheden (objecten), die interageren. Elk object is in staat om boodschappen te ontvangen, data te verwerken en boodschappen te versturen naar andere objecten.¹⁴

⁹ Database. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Database> , 20/02/06

¹⁰ "Dit omvat zowel de grafische user interface en de functionaliteit die de gebruiker in staat stelt de inhoud van de database te doorzoeken en te bewerken, alsook automatisch functionerende programma's die het systeem ondersteunen bij de verwerking van in- en uitvoer. Gebruikerstoepassingen stellen gebruikers in staat om mutaties in de database uit te voeren, terwijl automatische programma's op de achtergrond bewerkingen initiëren." Testbed Digitale Bewaring, Van digitale vluchtigheid naar digitaal houvast, Deel 1: Kosten- en beslissingsmodellen/Functionele specificaties. Bewaren van databases. Den Haag: ICTU, 2003, p. 17-18

¹¹ Database. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Database> , 20/02/06

¹² Testbed Digitale Bewaring, Van digitale vluchtigheid naar digitaal houvast, Deel 1: Kosten- en beslissingsmodellen/Functionele specificaties. Bewaren van databases. Den Haag: ICTU, 2003, p. 18

¹³ Database. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Database> , 20/02/06

¹⁴ Voor een uitgebreide verklaring, zie: Object-oriented Programming. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Object_oriented , 20/02/06

- Native XML¹⁵: Dit soort database definieert een logisch model voor een XML document. Bewaren en ophalen gebeurt aan de hand van dat model. De fundamentele eenheid van (logische) opslag is een XML document. Dit type vereist geen specifieke onderliggend opslagmodel. Ze zou, bijvoorbeeld, op een hiërarchische of relationele database kunnen gebouwd worden.¹⁶
- Hiërarchisch: Gegevens worden georganiseerd in een boomtakstructuur, maar op zo een manier dat niet te veel relaties mogelijk zijn. Hierdoor kan informatie herhaald worden door gebruik te maken van een *parent/child* relatie.¹⁷
- Netwerk: Waar het hiërarchisch model gebaseerd is op een structuur, waarin elk *record* één *parent* heeft en vele *children*, laat het netwerk model elke *record* toe meerdere *parents* en *children* te hebben.¹⁸
- Associatief: In plaats van een aparte, unieke tabel te gebruiken voor de verschillende types van data, wordt in dit model een simpele, generische structuur gebruikt die alle types van data bevat. Informatie over de logische structuur wordt samen met de data in de database bewaard, waar ze door de programma's worden opgehaald (i.p.v. dat ze geprogrammeerd worden in het programma zelf).¹⁹

In de meeste gevallen worden databanken bevraagd aan de hand van een bevragingstaal, die *Structured Query Language* (SQL) heet. Deze taal werd op het einde van de jaren 1970 door IBM ontwikkeld. In 1992 werd ze verheven tot Amerikaanse standaard door the American National Standards Institute (ANSI), onder de naam SQL-92.²⁰ Met deze taal kunnen complexe vragen aan een databank worden gesteld of databanken mee worden gecreëerd.²¹ Hoewel SQL werd

¹⁵ “eXtensible Markup Language (XML) is een standaard voor het definiëren van formele markup-talen voor de representatie van gestructureerde gegevens in de vorm van platte tekst. Deze representatie is zowel machineleesbaar als leesbaar voor de mens. Met andere woorden: XML is een bepaalde manier om gegevens gestructureerd vast te leggen. Deze manier is gedefinieerd en mag iedereen gebruiken. ... XML is een vereenvoudigde vorm van SGML ... Het gaat in dit bestandsformaat meer om de *structuur* van informatie, dit in tegenstelling tot HTML, waarin het meer gaat om de *presentatie* ...” Extensible Markup Language. In: *Wikipedia, The Free Encyclopedia*, <http://nl.wikipedia.org/wiki/XML> , 22/02/06; zie ook: <http://www.w3schools.com/xml/default.asp>

¹⁶ Staken, K. Introduction to Native XML databases. 31 oktober 2001, <http://www.xml.com/pub/a/2001/10/31/nativexmlldb.html>, 22/02/06

¹⁷ Hierarchical model. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Hierarchical_model , 20/02/06

¹⁸ Network model. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Network_model , 20/02/06

¹⁹ Williams, S. The associative model of data. Lazy Software Ltd., 2002, 2nd edition, p. 3

²⁰ Kroenke, D. Database processing. Fundamentals, design, and implementation. International edition. New Jersey: Pearson Education Inc., 2006, p. 28

²¹ Cumming, A. A gentle introduction to SQL. Edinburgh: School of Computing, Napier University, 2005, <http://sqlzoo.net/w.htm>, 16/11/05

gestandaardiseerd zijn bestaan er veel variaties en uitbreidingen, die meestal producentgebonden zijn (cfr. Oracle PL/SQL).²²

Een eerste vraag die gesteld moet worden, alvorens tot archiveren over te gaan, is: wat is het digitale archiefstuk? Bij databanken zijn verschillende mogelijkheden²³:

- Het complete databasesysteem (database, DBMS en applicatie)
- Een rij opgeslagen in een tabel (*tuple*)
- Data verspreid over verschillende tabellen
- Gegevens in de database zoals die worden gerepresenteerd op het scherm door de applicatie

Complexe systemen hebben vaak de volgende kenmerken²⁴:

- Ze zijn ontwikkeld via een combinatie van een DBMS voor de database en een aparte programmeeromgeving voor de ontwikkeling van de toepassing
- Het ontwerp, de bouw en het onderhoud vereisen specialistische vaardigheden
- Conversie van de inhoud naar een ander formaat of migratie naar een nieuwe versie van hetzelfde DBMS is mogelijk

Databanken op de desktop hebben een aantal andere kenmerken²⁵:

- Het DBMS kan worden gebruikt als een volledige oplossing van de database en voor het bouwen van de toepassing
- De software is gericht op de eindgebruiker en bijgevolg makkelijker om te gebruiken
- De gebruiker kan zelf migreren of converteren

D. E-mail

Met e-mail wordt zowel het systeem langs waar berichten worden verzonden als het eigenlijke bericht bedoeld. Het systeem bestaat uit programma's, een transportmedium (bvb. netwerkvoorzieningen) en computers.²⁶ Om e-mails te kunnen verzenden en ontvangen, hebben we een eigen adres nodig (net zoals iedereen op het internet). Dit adres bestaat uit twee delen, gescheiden door het teken @. In het eerste deel wordt de eigen naam gespecificeerd, de gebruikersnaam. In het tweede deel vinden we het domein, bepaald door de domeinnaam van de

²² SQL. In: Wikipedia, *The Free Encyclopedia*, <http://en.wikipedia.org/wiki/SQL>, 16/11/05

²³ Testbed Digitale Bewaring, *Van digitale vluchtigheid naar digitaal houvast, Deel 1: Kosten- en beslissingsmodellen/Functionele specificaties. Bewaren van databases*. Den Haag: ICTU, 2003, p. 19

²⁴ Testbed Digitale Bewaring, *Idem*, p. 22; Bvb. Oracle, SQL Server, DB2.

²⁵ Testbed Digitale Bewaring, *Idem*, p. 23; bvb. Access, dBase IV

²⁶ Testbed Digitale Bewaring, *digitale vluchtigheid naar digitaal houvast, Deel 4: Bewaren van e-mail*. Den Haag, ICTU, 2003, p. 20

internetprovider. Bijvoorbeeld vincent.sterken@vlaamsparlement.be. Om het berichtenverkeer correct te laten verlopen, installeert de *provider* speciale software (*mailserversoftware*). Er bestaan twee soorten:

- SMTP (Simple Mail Transport Protocol): Deze server zorgt ervoor dat verzonden berichten naar de juiste bestemming gaan.
- POP (Post Office Protocol): Deze server neemt de berichten in ontvangst en plaatst ze in de juiste elektronische postbus²⁷.

De *mail client* (e-mailprogramma) is het programma dat gebruikt wordt om berichten te schrijven, lezen, verzenden, etc. (bvb. Outlook of, in het geval van het Vlaams Parlement, Novell Groupwise).²⁸

Een e-mail kan als een eenvoudig tekstbestand zijn gemaakt in de *editor* van de *mail client*, maar kan ook bestaan uit een complex digitaal document met bewegende beelden of geluid gekoppeld aan het bericht (bijlage = *attachment*). Een *attachment* kan gemaakt zijn met eender welk soort programma.

De manier waarop een bericht op het scherm wordt voorgesteld, wordt gegenereerd door de *mail client*. Het verzonden bestand (transmissiebestand) bestaat uit platte tekst. De structuur is altijd dezelfde²⁹:

- *Header*. Hierin vinden we informatie over het bericht zelf. Ze bevat veel meer dan de meeste e-mailapplicaties weergeven. De basis informatie bestaat uit:
 - Datum en tijdstip van ontvangst door de servers (enkel inkomende mail)
 - MIME-versie: MIME (Multipurpose Internet Mail Extensions) is een internetstandaard voor e-mail formaten³⁰
 - Bericht-ID (uniek nummer)
 - Discussieonderwerp
 - Datum discussie-index
 - Onderwerp
 - Van: hier staat het adres van de afzender
 - Aan: hier staat het adres van de bestemming
 - CC (Carbon Copy): hier kan de afzender het adres ingeven van andere personen, waardoor deze hetzelfde bericht krijgen

²⁷ De elektronische postbus is een voorbehouden *directory* op de *mailserver* van de *internetprovider*, waarin de mails blijven staan totdat ze door de ontvanger worden opgevraagd.

²⁸ Boudrez, F.; Dekeyser, H. & Van den Eynde, S. Archiveren van e-mail. DAVID Rapport 5. 2^{de} editie. Antwerpen: Stadsarchief DAVID Project, 2003, pp. 8-9

²⁹ Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 4: Bewaren van e-mail. Den Haag, ICTU, 2003, p. 20-25

³⁰ <http://www.hunnysoft.com/mime/> , 20/02/06

- BCC (Blind Carbon Copy): dit is hetzelfde als CC, behalve dat de verschillende bestemmingen de adressen van de anderen niet kunnen zien
- Aanvullende headerinformatie kan bestaan uit berichtmarkeringen, de naam van de applicatie waarmee het bericht werd gemaakt, de prioriteit, etc. De onderdelen van het bericht volgen de hoofdheader. Elk onderdeel heeft vervolgens een eigen header met de volgende informatie:
 - Content type: definieert het tekstbericht (body). Eerst het type, daarna wordt aangegeven of het als onbewerkte tekst is verzonden dan wel opgemaakt (bvb. *Text/plain*)
 - Tekenset (bvb. ISO 8859-15³¹)
 - Content-Transfer-Encoding³² (bvb. Base64³³)

Dit is belangrijk om te bewaren omdat hiermee de delen van het transmissiebestand zijn samen te stellen.

- *Body*: Het bericht zelf, inclusief eventuele bijlagen, kan data bevatten in verschillende vormen. Dit kan gaan van eenvoudige ASCII-tekst tot afbeeldingen, multimedia of uitvoerbare programmabestanden.
 - Berichten in een onbewerkte tekstindeling hebben één tekstdeel met als content type *text/plain*.
 - Berichten met een complexe opmaak kunnen uit meerdere tekstdelen of ingevoegde bestanden bestaan.
 - Berichten die in HTML³⁴ zijn opgesteld worden veelal in twee vormen weergegeven: in onbewerkte tekst of in HTML vorm. Gebruikers die niet de HTML applicatie ondersteunen kunnen dan de onbewerkte

³¹ De *American Standard Code for Information Interchange* (ASCII) is een codetabel vastgelegd door het American National Standards Institute (ANSI) met de bedoeling informatie uitwisseling tussen computers mogelijk te maken. De ASCII code (vastgelegd als ISO-646) is 7-bits, waardoor 128 combinaties mogelijk zijn. Dus, 128 lettertekens werden vastgelegd, waarvan 94 afdrukbaar zijn. De originele tabel werd gemaakt voor PC's en bevat de tekens om Westerse talen vast te leggen, waarvan verschillende nationale varianten werden gemaakt. Omwille van de beperkte mogelijkheden om met een 7-bitstabel andere talen dan het Engels vast te leggen, werd de ASCII code uitgebreid tot een 8-bitstabel. Hierdoor konden 256 tekens worden vastgelegd. Deze karakterset werd vastgelegd in de ISO/IEC-8859 standaard (soms ook wel extended ASCII genoemd), die uit verschillende tabellen bestaat die elk voor een welbepaalde taal bedoeld zijn (vb. ISO 8859-15 is bedoeld voor West-Europese talen met oa. het € teken). Boudrez, F. Standaarden voor digitale archiefdocumenten. Technisch rapport. Antwerpen: Stadsarchief DAVID Project, 2001 – 2003, p. 10-11

³² "MIME defines a set of methods for representing binary data in ASCII text format. The *content-transfer-encoding*: MIME header indicates the method that has been used." MIME. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/MIME>, 31/03/06

³³ Base64 is een manier om data te interpreteren, opdat we deze zouden kunnen verzenden via een medium dat enkel uit tekst bestaat (zoals de body van een e-mail). Een ASCII karakterset kan gebruikt worden om tekst te formatteren, echter enkel een deel daarvan zijn leesbaar op het scherm (of afdrukbaar). Een manier om de niet-leesbare karakters om te zetten naar een leesbare vorm is Base64. Melnick, C. What is Base64? 2004, <http://www.aardwulf.com/tutor/base64/base64.pdf>, 22/02/06

³⁴ Zie volgende paragraaf.

tekstinhoud zien, zonder de totale opmaak. Ze worden gedefinieerd als content type *multipart/alternative* en bevatten twee weergaven van de body.

- Bijgevoegde bestanddelen moeten door een e-mailapplicatie opnieuw worden samengesteld voor een complete weergave van de body. Dit type van berichten worden meestal gedefinieerd als content type *multipart/related*. Hierbij wordt het totale transmissiebestand als *multipart* geïdentificeerd, waarmee aan de applicatie die het bestand ontvangt wordt doorgegeven dat sommige onderdelen van het bericht moeten worden samengesteld. Elk onderdeel heeft een afzonderlijke header die aangeeft welke positie het inneemt in het opnieuw samen te stellen document. Als deze informatie verloren gaat is de authenticiteit van het bericht in gevaar.
- *Attachment(s)*: Bijlagen kunnen van vrijwel elk bestandtype zijn. De enige beperking is de grootte van de mailbox van de bestemming. Indien het bestand te groot is, worden ze geweigerd. De header content type toont met welke applicatie het ingevoegde bestand is gemaakt en welke codering werd gebruikt. Binaire bijlagen worden voor het verzenden gecodeerd in Quoted Printable³⁵ of Base64, conform de MIME-standaard. Bijgevolg moet de e-mailapplicatie de attachments eerst decoderen alvorens het volledige document kan weergegeven worden.

E. Websites

Het internet werd aanvankelijk ontwikkeld door het Amerikaanse Ministerie van Defensie (onder de naam ARPANET). Het doel was computers met elkaar te verbinden, zodat informatiebeheer en raketbesturing van op verschillende plaatsen mogelijk zou zijn.³⁶ Van hieruit werd het internet ontwikkeld. De onderliggende idee was die van *open architecture networking*, waarbij elke provider vrij de netwerk architectuur kan kiezen. Communicatie tussen deze verschillende netwerken wordt mogelijk op basis van *internetworking architecture*. Om dit mogelijk te maken werd een standaardprotocol ontwikkeld, TCP/IP (Transmission Control Protocol/Internet Protocol). Elke computer die in het netwerk werd opgenomen kreeg een uniek IP-

³⁵ Quoted Printable is een manier om de niet-leesbare tekens van een ASCII tabel om te zetten naar leesbare tekens (cfr. Base64). *Quoted-Printable Content-Transfer-Encoding*. In: *Connected: An internet encyclopedia*, <http://www.freesoft.org/CIE/RFC/1521/6.htm> , 22/02/06

³⁶ Howe, W. *A brief history of the internet. An anecdotal history of the people and communities that brought about the Internet and the Web*. 7 november 2005, <http://www.walthowe.com/navnet/history.html> , 22/02/06

adres, waaraan later een domeinnaam werd gekoppeld.³⁷ Ondertussen is het internet sterk uitgebouwd en kent het vele toepassingen, waarvan het World Wide Web (WWW) het bekendste is. Het WWW werd aanvankelijk ontwikkeld voor tekstuitwisseling, maar is ondertussen geëvolueerd tot een interactieve en dynamische client-servertoepassing. Voor het ophalen van WWW-documenten werd HTTP (Hypertext Transfer Protocol) ontwikkeld ter ondersteuning van de communicatie van hypertextbestanden, en FTP (File Transfer Protocol) voor de uitwisseling van bestanden tussen server en client.³⁸

Voor het samenstellen van WWW-documenten werd, als standaard, de HyperText Markup Language (HTML) ontwikkeld. HTML is de markuptaal, ontwikkeld door het World Wide Web Consortium, voor documenten die op het internet worden gepubliceerd, gebaseerd op de Standard Generalized Markup Language (SGML). SGML legt de structuur, de inhoud en de semantiek van documenten vast. Het is een metataal die een *markup* taal is. Specifieker, een beschrijvende *markup* taal, d.i. gewone namen (i.p.v. codes) worden gebruikt om delen van het document te karakteriseren.³⁹ Dit wordt gedaan door middel van *tags*, die de structuur en de onderlinge relatie van de elementen waaruit het bestand is opgebouwd, vastleggen. Elk gegevensveld heeft een begin- en eindtag die tussen < en > staat. Het verschil tussen SGML en HTML ligt in het feit dat SGML-tags vrij kunnen bepaald worden, terwijl HTML-tags vast liggen.⁴⁰ Een HTML-bestand kan gemaakt worden met een simpele platte tekst editor, maar eveneens met gesofisticeerde WYSIWYG *authoring tools*. Door middel van vaste *tags* wordt de tekst gestructureerd in headings, paragrafen, lijsten, hypertext links, etc.⁴¹ In HTML kunnen inhoud en opmaak samen worden bewaard. Dit gaat echter in tegen haar ontwerpregels, die aanvankelijk enkel gericht waren op het uitwisselen van wetenschappelijke teksten. Tegenwoordig wordt daarom de voorkeur gegeven aan het afzonderlijk bewaren van de beide onderdelen van het document. Door middel van Cascading Style

³⁷ Leiner, B.; Cerf, G.; Clark, D.; Kahn, R.; Kleinrock, L.; Lynch, D.; Postel, J.; Roberts, L. & Wolff, S. A brief history of the internet. 10 december 2003, <http://www.isoc.org/internet/history/brief.shtml> , 22/02/06

³⁸ Boudrez, F. & Van den Eynde, S. Archiveren van websites. Antwerpen: Stadsarchief DAVID Project, 2002, p. 8

³⁹ Het woord *markup* werd gebruikt om annotaties of andere aanduidingen (*marks*) te beschrijven die dienden om de opmaak aan te geven van een tekst wanneer ze werd uitgegeven. Bij de automatisering van tekstuele uitgaves werd de term uitgebreid naar alle soorten *markup codes*, die gebruikt worden bij de opmaak van elektronische teksten. Bij verder uitbreiding wordt met de term *markup language* de set conventies bedoeld die gebruikt wordt om tekst te encoderen. A gentle introduction to SGML. In: Sperberg-McQueen, C.M. & Burnard, L., *Guidelines for Electronic Text Encoding and Interchange (TEI P3)*, the Text Encoding Initiative, 1994, <http://www.isgmlug.org/sgmlhelp/g-index.htm> , 22/02/06

⁴⁰ Boudrez, F. Standaarden voor digitale archiefdocumenten. Technisch rapport. Antwerpen: Stadsarchief, 2001 –2003, pp. 16-18

⁴¹ <http://www.w3.org/MarkUp/> , 22/02/06

Sheets (CSS) kunnen stijl en opmaak aan het document worden toegevoegd.⁴²

Een website wordt bekeken op een client computer. Het programma dat WWW-documenten genereert is de webbrowser. Deze stelt het HTTP-verzoek samen en stuurt het door naar de server (d.i. de computer die de website host en via het web ter beschikking stelt). Op de server draaien de webserverprogramma's, modules voor scripts en uitvoerbare programma's.

Een site bestaat uit een geheel van afzonderlijke bestanden die in een bepaalde mappenstructuur worden opgeslagen. De band tussen die bestanden wordt bepaald door hyperlinks. Webpagina's kunnen op twee manieren worden gelinkt, via:

- *Absolute links*: Deze verwijzen naar en vertrekken van de root van de website. De rootaanduiding start meestal van de URL⁴³ of het IP-adres.
- *Relatieve links*: De pathaanduiding vertrekt hier van de positie van waar gelinkt wordt. Het voordeel hiervan is overdraagbaarheid, omdat zowel van een server als van bvb. een CD gelinkt kan worden.⁴⁴

Een site is een geheel van aan elkaar gelinkte webpagina's, die via het HTTP-protocol op het inter- of intranet beschikbaar zijn, en deze zijn op hun beurt weer gelinkt aan grafische elementen zoals afbeeldingen, animaties en stylesheets. De client stuurt een HTTP-verzoek naar de server, die (via een *daemon*⁴⁵) de HTML-pagina(s) en bijhorende bestanden naar de client terugstuurt. Vervolgens genereert de browser op het scherm het WWW-document. De eerste generatie websites hadden een vaste inhoud, waardoor iedereen die een bepaalde pagina bezocht hetzelfde te zien kreeg. Ondertussen hebben we een evolutie gekend, waarbij de inhoud pas na interactie met de gebruiker op het moment van het verzoek wordt samengesteld. Deze inhoud is dan afhankelijk van bepaalde gebruikersvoorkeuren, geformuleerde

⁴² Boudrez, F. Standaarden voor digitale archiefdocumenten. Technisch rapport. Antwerpen: Stadsarchief, 2001 –2003, p. 18

⁴³ Een Uniform Resource Locator (URL) is het algemeen adres van documenten op het web. Het eerste deel geeft aan welk protocol moet gebruikt worden, het tweede deel specificeert het IP-adres of de domeinnaam waar het document zich bevindt. Vb.

<http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/index.html>.

<http://www.webopedia.com/TERM/U/URL.html> , 22/02/06

⁴⁴ Absolute V.S. relative hyperlinks. Hyperlinks what are they? University of North Carolina at Wilmington, Center for Teaching Excellence, <http://www.uncw.edu/cte/how2/rvaur/> ; <http://www.mediacollege.com/internet/html/hyperlinks.html> , 22/02/06

⁴⁵ "On the Web, each server has an *HTTPD* or HyperText Transfer Protocol daemon that waits in attendance for requests to come in from the rest of the Web. A *daemon* is a program that is "an attendant power or spirit" (Webster's). It sits waiting for requests to come in and then forwards them to other processes as appropriate." Hypertext Transfer Protocol Daemon. In: *Whatis?com, The leading IT encyclopedia and learning center*, http://whatis.techtarget.com/definition/0,,sid9_gci214005,00.html , 22/02/06

zoekopdrachten of de beschikbare informatie in het *deep web*.⁴⁶ Tegenwoordig bevatten vele sites zowel statische als dynamische pagina's.⁴⁷ De interactie tussen server en client wordt mogelijk gemaakt door serverscripts, ingebed in HTML-pagina's of opgenomen in aparte bestanden. Voor het uitvoeren van scripts is de passende websserversoftware vereist. Hiervoor werkt het script in vele gevallen samen met een programma dat op de server draait. Om de server te ontlasten wordt het uitvoeren van scripts zo veel mogelijk naar de client verschoven, door bij de gevraagde pagina deze mee te zenden (ingebod in een HTML-pagina of in een afzonderlijk bestand).⁴⁸

F. Audio

In een digitale omgeving wordt geluid bewaard en gemanipuleerd als een stroom van getallen, waarbij elk nummer de luchtdruk op een gegeven tijdstip voorstelt. De te bekomen binaire data wordt gegenereerd door een microfoon gekoppeld aan een Analog to Digital Converter (ADC)⁴⁹, die de amplitude van het signaal meet. Elk getal wordt een *sample* genoemd. Het aantal *samples* per seconde wordt omschreven als de *sample rate*.⁵⁰ M.a.w. het signaal wordt bij intervallen gemeten. Het aantal metingen per tijdseenheid noemt men de frequentie en wordt uitgedrukt in kilohertz (kHz). Een waarde van 44, 1 kHz geeft aan dat per seconde 44 100 metingen worden uitgevoerd. Hoe hoger de frequentie is, dus hoe kleiner de intervallen tussen metingen, hoe preciezer het digitale signaal het analoge⁵¹ zal benaderen.⁵²

⁴⁶ “**deep web** Also known as ‘invisible web’. Refers to a vast repository of underlying content, such as documents in online databases, that general-purpose web crawlers cannot reach. Both qualitative and quantitative in difference, deep web content is estimated at 500 times that of the surface web, yet has remained mostly untapped due to the limitations of traditional search engines.” *British Library’s strategy 2005-2008 glossary*. In: *The British Library, The world’s knowledge*, <http://www.bl.uk/about/strategic/glossary.html> , 22/02/06

⁴⁷ Boudrez, F. *Archiveren van websites: een kwestie van waardering en ‘capture’*. Antwerpen: Expertisecentrum DAVID vzw, 2005, p. 7

⁴⁸ Boudrez, F. & Van den Eynde, S. *Archiveren van websites*. Antwerpen: Stadsarchief DAVID Project, 2002, pp. 10-11

⁴⁹ Verschillende manieren bestaan om een ADC te implementeren: Flash (direct conversion), successive-approximation, delta-encoded, ramp-compare (integrating, dual-slope), multi-slope, pipeline (subranging quantizer), Sigma-Delta. *Analog-to-digital converter*. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Analog_to_digital_converter , 01/03/06

⁵⁰ Elsea, P. *Basics of digital recording. Converting sound into numbers*. University of California, Santa Cruz, Division of the Arts, 1996, http://arts.ucsc.edu/EMS/Music/tech_background/TE-16/teces_16.html , 22/02/06

⁵¹ “Analoog is het werken met waarden in een continuüm zonder stappen. Een analoog signaal is een signaal dat continu variabel is, in tegenstelling tot een digitaal signaal dat slechts een beperkt aantal discrete niveaus kan aannemen.” *Analoog*. In: *Wikipedia, The Free Encyclopedia*, <http://nl.wikipedia.org/wiki/Analoog> , 23/02/06

⁵² Baaten, L. & Vandermaesen, M. *Digitaal geluidsarchief. Krachtlijnen digitalisering: Standaarden, formaten dragers*. CDAVID, Erfgoed digitaal, 2004, pp. 5-6 http://www.cdavid.be/pdf/TechnischRapport_DigitaalGeluidsarchief.pdf , 17/02/06



Bron: Boudrez, F. Standaarden voor digitale archiefdocumenten. Technisch rapport. Antwerpen: Stadsarchief, 2001 – 2003, p. 31

Vervolgens worden de metingen in numerieke digitale data omgezet, waarbij meestal voor 8, 16 of 24 bits wordt gekozen om deze voor te stellen. Dit wordt de *sampleresolutie* genoemd. Hoe hoger het aantal bits, hoe preciezer de weergave. In een 8-bits *sampleresolutie* beschikt men over 255 waarden om data voor te stellen, in een 24-bits *sampleresolutie* over 16,7 miljoen. Dit wordt aangeduid als het dynamisch bereik (*dynamic range*). Een hogere *sampleresolutie* zorgt voor een betere omzetting van lage toonsignalen.⁵³ Het theorema van Nyquist bepaald dat wanneer een analoog signaal naar een tijddiscreet signaal wordt geconverteerd, de *samplerate* minstens tweemaal zo hoog moet zijn als de hoogste in het signaal aanwezige frequentie om het origineel zonder fouten te kunnen reproduceren. De helft van de *samplerate* heet de Nyquist-frequentie. Anders gezegd: voor een foutloze reproductie na *sampling* mag het analoge signaal geen frequenties bevatten hoger dan de Nyquist-frequentie, omdat dit kan resulteren in het overlappen van signalen.⁵⁴ Het hoorbereik van een mens ligt tussen 20 Hz en 20 kHz. Dit theorema toont aan dat een *samplerate* vereist is die groter is dan 40 kHz om hoorbare frequenties te reproduceren.⁵⁵

De opname van een signaal via een ADC heeft een zekere mate van nauwkeurigheid. Deze nauwkeurigheid is afhankelijk van de foutenmarge van de conversie.⁵⁶ Deze fouten worden gemeten in LSB (least significant bit). De foutenmarge heeft twee componenten⁵⁷:

- *Quantization error*: dit wordt veroorzaakt door de eindige resolutie van elke ADC; en is dus onvermijdelijk omdat analoog geluid een oneindige resolutie heeft.

⁵³ Baaten, L. & Vandermaesen, M., *Idem*, pp. 5-7

⁵⁴ Roland-Mieszkowski, M. Introduction to digital recording techniques. Digital Recordings, 1989-2005, <http://www.digital-recordings.com/publ/pubrec.html> , 22/02/06

⁵⁵ Kondratiuk, C. Technical aspects of multimedia audio. In: Jessica Keyes (ed.), *The ultimate multimedia handbook*, New York: McGraw-Hill, 1997 (2nd ed.), p. 27.1

⁵⁶ Deze zijn: offset error, full-scale error, integrale non-lineariteit (INL), differentiele non-lineariteit, quantization error. Staller, L. Understanding analog to digital converter specifications. In: *Embedded.com*, 24/02/2005, <http://www.embedded.com/showArticle.jhtml?articleID=60403334> , 01/03/06

⁵⁷ Analog-to-digital converter. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Analog_to_digital_converter , 01/03/06

- *Non-lineariteit*: dit wordt veroorzaakt door de fysieke imperfecties van elke ADC, waardoor de output afwijkt van de lineaire functie van de input.

De specificaties van een ADC vinden we terug in de *datasheet*.⁵⁸

Om het opgenomen geluid weer af te spelen wordt de binaire data terug geconverteerd door een Digital to Analog Converter (DAC), die verbonden is aan een luidspreker. Om deze reproductie zo getrouw mogelijk te maken, moet de *samplerate* hoog genoeg zijn. Indien deze te laag is krijgen we vervormingen van het geluid. Dit noemt men *aliasing*.⁵⁹ Hier gelden, in grote lijnen, dezelfde regels als voor een ADC. Belang wordt gehecht aan de samplersresolutie, de *samplerate* en het dynamisch bereik.⁶⁰

In vele gevallen wordt *Pulse Code Modulation* (PCM), het standaard binair formaat voor ongecomprimeerde, gesampled geluidssignalen, gehanteerd. Hier wordt het analoge signaal voorgesteld aan de hand van een serie *pulses* (1) en *non-pulses* (0). Deze stroom ondervindt niet veel invloed van interferentie en ruis (*noise*). Ook hier zijn de *samplerate* en de samplersresolutie van het grootste belang. In eerste instantie wordt het geluid gefilterd, waarbij de hoogste frequenties van een signaal worden weggelaten.⁶¹ Vervolgens wordt het analoge signaal gesampled aan een constante sample frequentie (via *Pulse Amplitude Modulation*, of PAM). De volgende stap is de digitale omzetting van de opname. Dit is PCM. Elke analoge sample wordt gedecodeerd aan de hand van binaire code woorden, door middel van kwantisering (*quantization*).⁶² “Quantization is the process of converting each analog sample value into a discrete value that can be assigned a unique digital code word.”⁶³ Hier is de samplersresolutie en de *samplerate* van belang voor een correcte weergave van de opname. Tijdens de kwantiseringsfase, worden input samples toegewezen aan een kwantiseringsinterval die zo kort mogelijk bij de hoogte van haar amplitude ligt. De interval krijgt vervolgens een discrete waarde toegewezen in de vorm van een binaire code. Omdat deze intervallen op

⁵⁸ Staller, L. *Idem*, laatste paragraaf

⁵⁹ Elsea, P. Basics of digital recording. Converting sound into numbers. University of California, Santa Cruz, Division of the Arts, 1996,

http://arts.ucsc.edu/EMS/Music/tech_background/TE-16/teces_16.html , 22/02/06

⁶⁰ De meest gebruikte type DAC's zijn de volgende: Pulse Width Modulator, Oversampling DAC's (vb. Delta-Sigma), binary weighted, R2R Ladder, segmented, hybride. Digital-to-analog converter. In: *Wikipedia, The Free Encyclopedia*,

http://en.wikipedia.org/wiki/Digital_to_analog_converter , 01/03/06

⁶¹ Introduction to Pulse Code Modulation (PCM). In: *ICT Technologies, Network for Capacity Building and Knowledge Exchange in ICT Policy, Regulation and Applications*, NetTel@Africa Off-line content, 2004, <http://cbdd.wsu.edu/kewlcontent/cdoutput/TR502/page13.htm> , 23/02/06

⁶² “Het onderverdelen van de reeks waarden van een variabele in een eindig aantal niet-overlappende en niet noodzakelijkerwijs gelijke deelgebieden of intervallen.” Biemond, H. Woordenboek ICT. Arnhem: Elsevier Bedrijfsinformatie BV, 2002, p. 686

⁶³ Waveform coding techniques. Cisco Systems Inc., 18/02/2005, p. 3, http://www.cisco.com/warp/public/788/signalling/waveform_coding.pdf , 01/03/06

geregelde afstanden worden ingelast, en dus op een uniforme manier zijn gestructureerd, spreken we van uniforme kwantisering. Indien de toegewezen interval niet overeenkomt met de feitelijke hoogte van de amplitude van de input sample, wordt een fout geïntroduceerd in het PCM-proces, die kwantiseringsruis (*quantization noise*) wordt genoemd. Dit is gelijk aan de willekeurige ruis die een impact heeft op de *signal-to-noise* ratio (SNR) van het signaal van een stem. SNR meet de sterkte van een signaal ten opzichte van achtergrondgeluid (weergegeven in decibels). Hoe hoger de SNR, hoe beter de kwaliteit van de opname. kwantiseringsruis vermindert de SNR van een signaal. SNR (inclusief kwantiseringsruis) is de belangrijkste factor in het bepalen van de geluidskwaliteit. Door het gebruik van gelijke kwantiseringsniveaus bij het dynamisch bereik van het analoge signaal, hebben lage signalen een kleine SNR. Hoge signalen hebben een grote SNR. Omdat de meeste geluidssignalen laag zijn, is het inefficiënt om de kwaliteit op een hoger niveau te verbeteren. Hierom wordt uniforme veelal vervangen door een non-uniforme kwantisering (en dus door non-lineaire signalen). Dit noemen we *companding*⁶⁴ (samenstelling van *compressing* en *expanding*).⁶⁵ Het analoge signaal wordt aan de bron gecomprimeerd in logaritmische segmenten, en weer naar zijn oorspronkelijke grootte hersteld (*expanding*) alvorens te worden omgezet naar een output signaal. Door de logaritmische compressie verhoogt de kwantiseringsruis bij de input, maar wordt de SNR tijdens het dynamisch bereik constant gehouden. De gebruikte standaarden zijn μ -law (voor Noord Amerika en Japan) en A-law voor Europa.⁶⁶

Bij de digitale opname van een analogoog geluidssignaal is het de bedoeling dat zo weinig mogelijk informatie verloren gaat. Een methode die echter veel wordt toegepast is die van compressie, waarbij een bitstream kleiner wordt. Bij het archiveren moet dit echter worden afgeraden, omdat bij het reconstrueren een bijkomende stap wordt geïntroduceerd. Bovendien gaat het gebruik van *lossy compression* gepaard met informatieverlies. Het verwerken van gecomprimeerde bitstreams is complexer. Gecomprimeerde bestanden zijn bovendien kwetsbaarder dan ongecomprimeerde, waardoor een fout sneller tot onherstelbaar verlies leidt.⁶⁷

⁶⁴ “Compandor- Samentrekking van compressor en expandor. Apparatuur die het uitgaande spraakvolumegebied comprimeert en het binnenkomende spraakvolumegebied t bij een telecommunicatiesysteem.” Biemond, H. *Woordenboek ICT*. Arnhem: Elsevier Bedrijfsinformatie BV, 2002, p. 162

⁶⁵ *Waveform coding techniques*. Cisco Systems Inc., 18/02/2005, pp. 3-5, http://www.cisco.com/warp/public/788/signalling/waveform_coding.pdf , 01/03/06

⁶⁶ Smith, S.W. *The scientist and engineer’s guide to digital signal processing*. California Technical Publishing, 1997 (chapter 22, update 28/12/2001), 362-363, <http://www.dspguide.com/ch22.htm> , 01/03/06

⁶⁷ Boudrez, F. & Dekeyser, H. *Digitaal archiefbeheer in de praktijk*. Antwerpen: Stadsarchief DAVID Project, 2004, p. 69

G. Video

Een (klassieke) video camera converteert een beeld in een elektrisch signaal. Elektrische signalen zijn ééndimensionaal (i.e. ze hebben één waarde op elk gegeven tijdstip), terwijl beelden tweedimensionaal zijn met vele waarden voor alle verschillende posities binnen dat beeld. Om dit tweedimensionaal beeld om te zetten naar een ééndimensionaal elektrisch signaal, wordt dit beeld op een systematische, progressieve manier gescand (rasterscan). De afbeelding wordt door middel van horizontale lijnen gescand, waarbij de elektrische output van de scanner sensor (per lijn) de lichtintensiteit weergeeft van het beeld. Tijdens de *horizontal blanking interval*⁶⁸ (het moment dat de scanner zich terug naar het begin van de lijn beweegt), wordt de sensor uit gezet waardoor een nulwaarde als signaal wordt verzonden (*blanking-level*). Het signaal van een complete scan is een sequentie van lijn signalen, verdeeld door horizontale *blanking intervals*, die een *frame* wordt genoemd. Een belangrijke parameter in het scanproces is de *aspect ratio*. Dit is de ratio van de lengte van een horizontale, gescande lijn ten opzichte van de verticaal afgelegde afstand door alle gescande lijnen. M.a.w. de hoogte-breedte ratio van een *frame*. Om een effect van vloeiende beweging te verkrijgen, moeten vele *frames* per seconde (fsp) worden gescand.⁶⁹ De *frame rate* gaat van 6 tot 8 fsp voor oude, mechanische, tot 120 of meer fsp voor nieuwe, professionele camera's. De minimum fsp om een bewegend beeld te bekomen, is 10 fsp.⁷⁰

Er bestaan twee manieren om beeld vast te leggen:

- Bij interliniëring (*interlaced scanning*) wordt een beeld vastgelegd door middel van alternerende sets van lijnen. Eerst worden de even lijnen gescand, vervolgens de oneven, en dan weer de even, enzovoort. Eén set van even of oneven lijnen wordt omschreven als een veld. Een opeenvolgende koppeling van tegengestelde velden is dan een *frame*.
- Bij een progressieve scan wordt elke *frame* afzonderlijk geregistreerd, waarbij beide velden als identiek worden beschouwd.

Interlaced video registreert dus het dubbel aantal velden per seconde ten opzichte van progressieve video indien ze werken met hetzelfde aantal fsp.⁷¹

⁶⁸ "Blanking – *Blinderen* – (1) Het onderdrukken van het afbeelden van één of meer beeldelementen of beeldgroepen op een scherm. (2) Het gedurende een korte periode blokkeren van een beeldsignaal als de elektronenbundel van een rasterscan video monitor naar een positie wordt verplaatst om een nieuwe regel af te beelden." Biemond, H. *Woordenboek ICT*. Arnhem: Elsevier Bedrijfsinformatie BV, 2002, p. 88

⁶⁹ Morse, K. *Introduction to digital video and audio*. In: Jessica Keyes (ed.), *The ultimate multimedia handbook*, New York: McGraw-Hill, 1997 (2nd ed.), pp. 32.2-32.4

⁷⁰ *Video*. In: *Wikipedia, the free encyclopedia*, <http://en.wikipedia.org/wiki/Video>, 23/02/06

⁷¹ *Digital video*. In: *Wikipedia, the free encyclopedia*, http://en.wikipedia.org/wiki/Digital_video, 22/02/06

Bij digitale video moet een analoge golfvorm (*waveform*)⁷², die zowel in tijd als in waarde constant is, omgezet worden naar een digitaal signaal, met niet constante, discrete waarden. De amplitude wordt weergegeven als een digitaal geheel getal (met een specifiek aantal bits) en tijd als een serie van deze amplitude waarden vastgesteld tijdens gelijke tijdstippen. Het proces waarbij amplitude in discrete waarden wordt omgezet, wordt kwantisering genoemd, en tijd omzetten in discrete waarden *sampling* (d.i. ADC). Het principe van het nemen van analoge samples die vervolgens omgezet worden in digitale waarden, is dezelfde als bij audio.⁷³

Tijdens video ADC wordt het analoge signaal omgezet naar een stroom van digitale getallen (nummers) die het originele beeld voorstellen. Elk van deze getallen staat voor aanpalende punten in dat beeld die hetzelfde patroon volgen als die van de scan van de camera. Deze punten (of getallen) noemen we pixels. Tijdens de reconstructie van het beeld wordt elke pixel getoond als een klein vierkant gevuld met een kleur gebaseerd op dat nummer. Hoe realistisch de reconstructie er vervolgens uitziet, is afhankelijk van het aantal pixels en het gebruikte kleurschema⁷⁴ (vb. YUV, RGB⁷⁵).

De kwaliteit van het beeld wordt (net zoals bij geluid) bepaald door de sample rate, sampleresolutie en het dynamisch bereik. De maximale resolutie van een apparaat hangt samen met het aantal lichtsensoren dat het bevat. In principe kan geen beeld gemaakt worden met een hogere resolutie dan de maximale resolutie. In de praktijk is het wel mogelijk door *interpolatie*. Artificiële pixels worden dan toegevoegd door waarden af te leiden uit de omliggende pixels. Bij beelden staat het dynamische bereik voor het verschil tussen de zwakste en de sterkste lichtintensiteit die een camera (of scanner) kan waarnemen.⁷⁶

⁷² “Waveform means the shape and form of a signal, such as the wave moving across the surface of water, or the vibration of a plucked string. ... the shape of a graph of the varying quantity against time or distance.” *Waveform*. In: *Wikipedia, the free encyclopedia*, <http://en.wikipedia.org/wiki/Waveform> ; “a usually graphic presentation of a wave that indicates its characteristics (as frequency and amplitude)” *Waveform*. In: *Merriam-Webster Online Dictionary*, <http://www.m-w.com/dictionary/waveform> , 23/02/06; “Waveform - *Golfvorm* - Een grafische weergave van het verloop van een signaal in de tijd.” Biemond, H. *Woordenboek ICT*. Arnhem: Elsevier Bedrijfsinformatie BV, 2002, p. 916

⁷³ Morse, K. *Introduction to digital video and audio*. In: Jessica Keyes (ed.), *The ultimate multimedia handbook*, New York: McGraw-Hill, 1997 (2nd ed.), pp. 32.8-32.9

⁷⁴ Morse, K. *Idem*, pp. 32.12-32.14

⁷⁵ YUV is het kleurschema dat in het PAL televisiecodeersysteem (de norm in het grootste deel van Europa) wordt gebruikt. In het YUV model wordt de kleurenruimte bepaald in termen van een helderheidscomponent (Y) en twee kleurcomponenten (U en V). RGB (Red, Green, Blue) is een kleurcodering, waarbij kleur wordt uitgedrukt als een combinatie van drie primaire kleuren, uitgaande van additieve kleurmenging. *YUV*. In: *Wikipedia, the free encyclopedia*, <http://en.wikipedia.org/wiki/YUV> ; *RGB Color Model*. In: *Wikipedia, the free encyclopedia*, http://en.wikipedia.org/wiki/RGB_color_model , 24/02/06

⁷⁶ Vandermaessen, M. *Rapport. Digitaal beeldarchief. Aanbevelingen voor het realiseren van beeldbanken, aan de hand van 'lessons learnt' en 'best practices'*. Antwerpen: Stadsarchief, cDAVID, 2005, pp. 40-45

3. Inventarisatie van software en formaten, met aanbevelingen voor bewaring

Alvorens te weten hoe gearchiveerd moet worden, is het noodzakelijk te weten met wat gewerkt wordt. Een inventarisatie van de verschillende systemen en de daaruit voortvloeiende documenten is bijgevolg aangewezen. De enige die een overzicht kan voorleggen van de gebruikte producten, zijn de archiefvormers en de ICT-afdeling. Tijdens verschillende interviews werd nagegaan hoe het ICT-plaatje eruitziet. Op de volgende bladzijden worden de bevindingen samengevat. In het eerste deel van dit hoofdstuk vinden we de enquête zelf, vervolgens worden de bevindingen besproken.

3.1 De bevraging (ICT enquête)

VLAAMS PARLEMENT



DIRECTIE INFORMATIE
Archief en Biografische Dienst

VOORONDERZOEK DIGITALE ARCHIVERING
BEVRAGING VAN ICT
(aangevuld door: IPEA-enquête¹)

KADER

Het ruimere kader waarbinnen ICT bevraged wordt is de uitbouw van een digitaal depot.

Het begrip **digitaal depot** of **E-depot** is gedefinieerd als ‘geheel van apparatuur, programmatuur, procedures, methoden, kennis en vaardigheden waarmee het Archief en Biografische Dienst van het Vlaams Parlement in staat is zijn digitale informatie te beheren en beschikbaar te stellen’.

De bedoeling van het E-depot is zijn gebruikers te voorzien in de langdurige en betrouwbare toegang tot de digitale bronnen die het beheert. Dit dient op een zodanige wijze plaats te vinden dat de archiefbescheiden kunnen worden teruggevonden, worden gelezen en in authentieke en begrijpbare vorm worden geproduceerd.

Met deze bevraging wordt een eerste aanzet gegeven tot het vanuit archivalisch perspectief in kaart brengen van de ICT-infrastructuur en

¹ Innovative Platform for Electronic Archiving, Het IPEA-project is hét project in Vlaanderen dat zich bezig houdt met het digitaal archiveren en uitwisselen van beeldgegevens (bewegend beeld). In die zin zijn ze complementair met het DAVID-project (Digitale Archivering in Vlaamse Instellingen Diensten) dat zich nog niet met het bewegend beeld had ingelaten. Partners in het project zijn - naast de Universiteit van Gent- de grote spelers op het vlak van broadcasting (Vlaamse Mediamaatschappij, VRT en Vidohouse). Op termijn worden ook de andere sectoren betrokken: cultuur, overheid, onderwijs. Omwille van copyright issues kan de enquête die ter beschikking hebben gesteld, niet worden opgenomen.

technische randvoorwaarden waarbinnen de digitale objecten tot stand komen.

Het selectieplan voor de archiefdocumenten enerzijds en de ICT omgeving anderzijds vormen de belangrijkste uitgangspunten bij het uitwerken van het projectplan 'digitale archivering'. De scope van dit plan is het zoeken naar en uitwerken van een digitale en duurzame archiveringsoplossing voor het Vlaams Parlement. Dit plan kan slechts gerealiseerd worden in samenwerking met de Directie Informatica.

Na een reeks algemene vragen met betrekking tot de algemene inrichting van de ICT-infrastructuur, worden de vragen specifiek en richten ze zich op de technische omgeving waarbinnen de diverse digitale objecten tot stand komen.

Hilde Van Ongevalle
Archivaris

Vincent Sterken
Stagiair Archivistiek & Hedendaags Documentbeheer (VUB)

Brussel, 31 maart 2006

I. Algemeen

- 1.1. Hoe ziet het ICT-plaatje er uit? Kan er een schematische voorstelling gegeven worden van de huidige ICT-infrastructuur? Storage (evt. externe storage), network-management, datamanagement (back up), information management, applicaties, beveiliging, controle.
- 1.2. Wordt er op het niveau van de organisatie en het beheer van processen gewerkt volgens bepaalde standaarden, normen, interne richtlijnen?
- 1.3. Zijn er algemene richtlijnen met betrekking tot veiligheid en authenticiteit/integriteit van het databeheer? (Bepaalde normen die gevolgd worden?)
- 1.4. Bestaat een risico-analyse, een rampenplan?
- 1.5. Bestaat een inventaris van de huidige hard- en software?
- 1.6. Heeft men reeds een visie ontwikkeld op het vlak van digitale archivering?

II. Specifiek

2.1. Kantoordocumenten

2.1.1. Server

- Naam
- Type
- Operating System
- Versie

2.1.2. Applicaties (welke nu ondersteund worden, naam en doelstelling)

- tekstdocumenten
- spreadsheets
 - naam
 - versie
 - doelstelling
 - formaat
 - volume

2.1.3. Opslag

Waar(op) worden de geproduceerde documenten opgeslagen? Welke opslagmedia wordt gebruikt? Hoe ver kan men terug gaan?

2.1.4. Historiek

Applicaties die vroeger ondersteund werden, logboek van uitgevoerde conversies en migraties.

2.1.5. Authenticiteit

- Zijn er garanties ingebouwd in het huidige systeem om aan te tonen dat de documenten de documenten zijn die ze beweren te zijn? (vb. digitale handtekening, statuut van de PDF's op de site [categorie gedigitaliseerde, huidige PDF's], gebruik van metadata [context, technisch, bewaring], beveiliging, ...).
- Zijn er regels met betrekking tot het vastleggen van metadata mbt de technische omgeving, bewaarmetadata?

2.1.6. Systeem van Back-up?

2.1.7. Wat denkt de Directie Informatica over de uitbouw van een documentair structuurplan? (Opbouw van een digitaal klassement aan de hand van een mappenstructuur).

2.2. E-mail

2.2.1. Server

- Naam
- Type
- Operating System
- Versie

2.2.2. Van welke programmatuur wordt gebruik gemaakt (welke mailserversoftware)?

2.2.3. Hoe verloopt de opslag van de huidige mails (op het niveau van de Directie Informatica)?

- Houdt ICT alle mails bij?
- In welk formaat wordt opgeslagen?
- Wat houdt men bij (volledige mail zoals de gebruikers hem zien met bijlagen, complete transmissiebestand, ...)
- Wordt de structuur bewaard? (Structuur zijnde het logische verband tussen de elementen en het e-mailbericht).
- Wordt het gedrag bewaard; is interactie nog mogelijk? (vb. activeren van een hyperlink?)

2.2.4. Authenticiteit

Zijn er garanties ingebouwd in het huidige systeem om aan te tonen dat de documenten de documenten zijn die ze beweren te zijn?

2.2.5. Historiek

Heeft men steeds met het bestaande e-mailplatform gewerkt?
Werden conversies of migraties of andere bewaarstrategieën uitgevoerd?

2.2.6. Om welke redenen wordt eind dit jaar een nieuwe e-mailomgeving uitgebouwd? In welk stadium zit dit project?

2.2.7. Wat gebeurt er met de oude mails?

2.2.8. Kan het Archief en de Biografische Dienst nog aandachtspunten overmaken.

Zo niet: hoe werd de archiveringsfunctie geconcipieerd?

- wordt een gebruiksvriendelijke integratie van mail en kantoordocumenten voorzien?
- integratie van mail en bijlage?
- waar en hoe verloopt de opslag? (mail en bijlagen in origineel formaat, waar kunnen de gebruikers hun mails opslaan, hoe verloopt de opslag op het niveau van ICT?...)
- welke metadata kunnen worden gearchiveerd? (in het bijzonder context aan de hand van headergegevens: links met dossier, directie, werkproces, bijlagen, ...)

2.2.9. Bestaat er een e-mail-policy?

2.2.10. Systeem van back-up?

2.3. DBMS – Databanken – Gebruikersapplicaties

2.3.1. Welke databanken worden gebruikt? En over welk type(s) gaat het (relationeel, hiërarchisch, netwerk, ...)?

2.3.2. Welke DataBase Management Systemen? Besturingssystemen?

2.3.3. Welke gebruikersapplicaties draaien op de DBMS? In welke taal wordt gecommuniceerd (SQL bvb. Versie)?

2.3.4. Welke directie(s) is (zijn) betrokken?

2.3.5. Databank

Waar worden de gegevens van de databank bewaard? In welk formaat? Opslagmedium?

Is de logische structuur van de databank bewaard? (vb. ERD-Entity Relationship Diagram of het XML-schema).

Worden gegevens over het gebruik in de vorm van metadata vastgelegd (bvb. logboek)? Zo ja, welke gegevens worden opgeslagen en op welke manier?

Wordt voorzien in beveiliging (vb. restricties qua toegankelijkheid, inhoud, ...)?

Bestaat er een overzicht van wie welke inhoud aanlevert?

2.3.6. Gebruikersapplicatie

Welke output genereert de databank en waar wordt die opgeslagen? In welk formaat? Garanties voor authenticiteit?

2.3.7. Bestaan er koppelingen met andere databanken?

2.3.8. Historiek

Werden er reeds conversies, migraties, ... uitgevoerd? Of koos men voor andere bewaarstrategieën? (vb behouden van de oude apparatuur en software).

2.4. Website

2.4.1. Wanneer werd de huidige website in gebruik gesteld?

2.4.2. Server

- Naam
- Type
- Operating System

2.4.3. Webserverconfiguratie

- welke webserversoftware?
- welke bijbehorende modules?
- welke webbrowser?

2.4.4. Welke soort scripts worden op de server uitgevoerd?

2.4.5. Welke toepassingen, databanken, systemen zijn verbonden met en toegankelijk via de website?

2.4.6. Bestaat er een overzicht van wie welke informatie via het web aanlevert?

2.4.7. Opslag

Wat wordt op het niveau van ICT opgeslagen? Waar? Opslagmedium?

2.4.8. Werden er reeds bewaaroperaties uitgevoerd?

2.4.9. Met welke bestandsformaten (en versies) wordt gewerkt?

- HTML (versie)
- Tekstbestanden
- audio
- afbeeldingen
- video
- uitvoerbare programma's
- clientscripts

2.4.10. Wanneer kan de archivering van de oude website worden ingepland?

2.4.11. Hoe werd de archiveringsfunctie geconcipieerd?

2.5. Intranet

2.5.1. Wanneer werd het huidige intranet in gebruik gesteld?

2.5.2. Server

- Naam
- Type
- Operating System

2.5.3. Configuratie

- welke software?
- welke bijbehorende modules?

2.5.4. Welke toepassingen, databanken, systemen zijn verbonden en toegankelijk via de website?

2.5.5. Wat wordt op het niveau van ICT opgeslagen? Waar?

2.5.6. Werden er reeds bewaaroperaties uitgevoerd?

2.5.7. Met welke bestandsformaten (en versies) wordt gewerkt?

- HTML (versie)
- Tekstbestanden
- Audio
- Afbeeldingen
- Video
- uitvoerbare programma's
- clientscripts

2.5.8. Hoe werd de archiveringsfunctie geconcipieerd?

2.6. Audio

2.6.1. Server

- Naam
- Type
- Operating System

2.6.2. Van welke applicatie wordt gebruik gemaakt? Welk type? (DAC/ACD)

2.6.3. In welke formaten wordt opgeslagen? Opslagmedium?

2.6.4. Wat is de sampleresolutie?

2.6.5. Wat is de samplerate?

2.6.6. Hoeveel kanalen bevat de opname, 1 (mono), of 2 (stereo)?

2.6.7. Wordt gebruik gemaakt van compressie? Zo ja, welk type?

2.6.8. Hoe verloopt opslag op niveau van ICT? Opslagmedium?

2.6.9. Bewaring

Werden reeds migraties of conversies uitgevoerd? Andere bewaarstrategieën?

2.6.10. Beveiliging

Hoe werkt de beveiliging? In welke mate is de integriteit van de bestanden gegarandeerd?

2.7. Video

2.7.1. Hoe worden de beelden aangemaakt?

2.7.2. In welke formaten wordt opgeslagen? Opslagmedium?

2.7.3. Wordt gebruik gemaakt van compressie? Zo ja, welk type?

2.7.4. Bewaring

Werden reeds migraties of conversies uitgevoerd? Andere bewaarstrategieën?

2.7.5. Beveiliging

Hoe werkt de beveiliging? In welke mate is de integriteit van de bestanden gegarandeerd?

3.2 Analyse

Wegens veiligheidsredenen heeft de Directie Informatica gevraagd bepaalde bevindingen en details te weerhouden.

A. Algemeen

Algemeen kan gesteld worden dat vanuit de Directie Informatie tot voor kort niet veel aandacht werd besteed aan de problematiek van digitale archivering. De nadruk van hun opdracht ligt uiteraard niet op bewaring, maar op het onderhoud van de technische infrastructuur. Hieruit vloeit een visie voort die gericht is op de korte termijn. Simpel gesteld, het netwerk moet *nu* werken, de databanken moeten *nu* beschikbaar zijn, indien problemen zich stellen moeten die *nu* opgelost worden. Het is uiteraard niet aan ICT om de archiveringsfunctie uit te denken en te ontwerpen (voor iemand denkt dat ze hun werk niet doen). Dat is een taak voor de archiefdienst.

Tot nu werd het geheel van virtuele toepassingen niet in een vaste structuur gegoten. De verschillende directies werken met systemen die ze zelf hebben gekozen. In geval van problemen wordt gewerkt met *ad hoc* oplossingen. Op het niveau van beveiliging, authenticiteit en integriteit van het databeheer bestaan een aantal richtlijnen die voor een aantal garanties zorgen. Om dataverlies tegen te gaan wordt een rampenplan ontwikkeld. De server zal ontdubbeld worden en op twee verschillende verdiepingen fysiek geplaatst. Als beveiliging voor bvb. system failures worden backups gemaakt, die op losse media in een kluis worden bewaard. Die backups zijn niet toegankelijk voor de gebruikers. Ze volgen wel de aangebrachte veranderingen met een afstand van het origineel van maximum twee uur. Op dit moment staat de replica echter fysiek dicht bij het origineel. In geval van wateroverlast, brand of een ander probleem zit de kans erin dat beide beschadigd worden. Hierom zullen de verschillende servers op verschillende plaatsen worden bewaard. In het geval van problemen kan snel overgeschakeld worden op het stand-by systeem. In principe zou het volledige systeem binnen maximum zeven minuten *up and running* moeten zijn, met nagenoeg geen verlies aan gegevens. Verschillende mogelijkheden werden geopperd bij de uitvoering van het rampenplan. Zo is het mogelijk een data center op te zetten. Dit een computerruimte, een fysieke plaats waar servers, netwerken en ander computermateriaal kunnen geplaatst en gecontroleerd worden.² Algemeen genomen vinden we hier ook systemen van klimaatcontrole (air conditioning, brandbeveiliging, etc.), backup elektriciteitsvoorzieningen en internetconnecties, en een hoge mate van beveiliging.³ Een andere optie zou zijn te werken met clusters. Een computer cluster is een groep van computers die aan elkaar gekoppeld zijn. Ze werken zo nauw samen dat ze kunnen worden bekeken als één computer. In de meest gevallen (niet altijd) zijn ze met elkaar verbonden via een snelle LAN (Local Area

² <http://www.fastpcnet.net/datacenter.html>, 16/07/06

³ Data center, In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Datacenter>, 19/07/06

Network) verbinding.⁴ De verwerking gebeurt dan op twee systemen in verschillende clusters. Het doel is *load balancing*. De verwerking van gegevens wordt dan op gelijke wijze verspreid over meerdere computers, waardoor de capaciteit wordt verhoogd.⁵ Indien één systeem stilvalt, neemt het andere het werk over. Dit verhoogt wel de *workload* op het overgebleven systeem.

B. Kantoordocumenten

Qua kantoordocumenten worden gewoon de typische elementen uit het MS Office pakket gebruikt, d.i. Word, Excel, PowerPoint, en occasioneel Access. De documenten die hieruit voortvloeien, worden veelal op de lokale harde schijf opgeslagen. Een aantal worden nog geplaatst op de centrale server, helaas is deze door een gebrek aan regels totaal onoverzichtelijk geworden. De archiefdienst heeft een tussentijdse richtlijn voor de overdrachten aan de Directie ICT overgemaakt. Zo nu en dan worden CD's overgedragen aan de archiefdienst, zonder verdere uitleg bij welke dossiers de gegevens behoren. De archiefdienst heeft een tussentijdse richtlijn voor de overdrachten aan de Directie ICT overgemaakt.

Op het niveau van ICT wordt het geheel van alle gegevens weliswaar op een Raid 5 Hard Disk en op backup-schijven bewaard. Via deze reservekopieën kunnen gegevens worden opgevraagd tot twee maanden voordien. In principe zijn alle opgeslagen gegevens *downward compatible*. Dit dient echter enkel voor het oplossen van problemen op de korte termijn in het geval van een crashend systeem, maar voor de uitbouw van een e-depot totaal ongeschikt.

Garanties dat documenten de documenten zijn die ze beweren te zijn, zijn er niet echt in het huidige systeem. Er is voorzien in de mogelijkheid te ondertekenen met een digitale handtekening, geverifieerd met een eID. Dit is de elektronische Identiteitskaart, die tegen 2008 in het bezit zou moeten zijn van alle Belgen.⁶ De kaart heeft meerdere toepassingsmogelijkheden, waaronder het certifiëren van documenten en e-mails.⁷ Microsoft voorziet hier zelfs expliciet in.⁸ Hier wordt echter zelden gebruik van gemaakt.

Op het niveau van de metadata kan gesteld worden dat het systeem zelf een aantal technische gegevens genereert. De mogelijkheid zou moeten bestaan om de gebruikers zelf een aantal gegevens in te laten vullen. Dit wordt echter nooit gedaan.

⁴ Computer cluster. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Computer_cluster, 19/07/06

⁵ Load balancing. In: *Webopedia*, http://www.webopedia.com/TERM/L/load_balancing.html, 19/07/06

⁶ Fedict, *Elektronische identiteitskaart: eID*. Februari 2006, <http://www.belgium.be/eportal/application?languageParameter=nl&pageid=contentPage&docId=6418>, 19/07/06

⁷ Fedict, *eID gebruiken*. <http://eid.belgium.be/nl/navigation/documents/39763.html>, 19/07/06

⁸ <http://www.microsoft.com/belux/nl/office/eid/word.aspx>

De essentiële elementen van dit type van documenten zijn vrij eenvoudig te bepalen.⁹ In eerste instantie zijn de inhoudelijke gegevens het belangrijkste, samen met een duidelijke bepaalde structuur. Zowel voor tekstdocumenten als voor presentaties is het mogelijk om, ten laatste voor de opname in het digitaal depot, de bestanden om te zetten naar Portable Document Format (PDF). PDF werd in 1993 door de softwareproducent Adobe op de markt gebracht. Hoewel de eigendomsrechten van één producent zijn, zijn de specificaties vrij gegeven¹⁰, waardoor het nu een open *de facto* standaard is.¹¹ Het grote voordeel van PDF-documenten ligt in het feit dat het een bestandsformaat is met een universele bestandsindeling (i.e. resolutie-, programma- en platformonafhankelijk).¹² In andere woorden, “PDF-bestanden zien er exact hetzelfde uit als de originele documenten en behouden informatie van het bronbestand – tekst, tekeningen, 3D, afbeeldingen in kleur, foto’s en zelfs bedrijfslogica – ongeacht de toepassing waarin ze zijn gemaakt” en “kan op elk platform worden weergegeven en afgedrukt.”¹³ Alle gegevens die nodig zijn om dit te kunnen bewerkstelligen worden in het PDF-bestand zelf bewaard. Mogelijkheden tot het bewaren van metadata wordt geboden door het *Extensible Metadata Platform* (XMP). Met XMP, gebaseerd op specificaties van het *World Wide Web Consortium*,¹⁴ kunnen context gegevens ingebed worden in het bestand.¹⁵ Om een document te openen is Acrobat Reader vereist, die gratis kan gedownload worden op de website van Adobe. Overduidelijk ideaal voor archiveringsdoeleinden. Of ook niet. De voordelen die reguliere versies bieden op het vlak van gebruiksvriendelijkheid, stellen naar archivering toe een aantal problemen. De aanmaak en opslag kan op heel gevarieerde wijze gebeuren. Doordat de specificatie bekend is, zijn ook verschillende programma’s in omloop om PDF-bestanden aan te maken (bvb. Adobe Writer, pdf995, Easy PDF, eDocPrinter PDF Pro, etc.¹⁶). Bij de aanmaak kunnen de verschillende eigenschappen van een document aangepast worden aan de wensen van de gebruiker (inclusief in meerder of mindere mate compresseren). Opslag kan op gestructureerde, ongestructureerde en

⁹ Ik beperk mij tot tekstbestanden, spreadsheets en presentaties (cfr. PowerPoint). Programma’s zoals Access worden behandeld bij de bespreking van databanken.

¹⁰ En zijn hier te vinden:

http://partners.adobe.com/public/developer/pdf/index_reference.html

¹¹ Boudrez, F. *Standaarden voor digitale archiefdocumenten*. Technisch rapport. Antwerpen: Stadsarchief, 2001 –2003, p. 21

¹² *Portable Document Format*. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Portable_Document_Format, 20/07/06

¹³ <http://www.adobe.com/nl/products/acrobat/adobepdf.html>

¹⁴ *The World Wide Web Consortium* (W3C) is een internationaal consortium waar leden (organisaties), eigen personeel en het grote publiek samenwerken om webstandaarden te ontwikkelen. “W3C’s mission is: to lead the World Wide Web to its full potential by developing protocols and guidelines that ensure long-term growth for the Web.” Jacobs, I. *About the World Wide Web Consortium (W3C)*. 13 July 2006, <http://www.w3.org/Consortium/>, 21/07/06

¹⁵ <http://www.adobe.com/products/xmp/>

¹⁶ En dit is nog maar een kleine greep uit de eerste Google-pagina na een simpele zoekopdracht.

getagde (cfr. HTML-tags) wijze. Getagde documenten zijn te verkiezen boven (on)gestructureerde, omdat dit het mogelijk maakt dat andere applicaties dan Acrobat Reader voor een correcte weergave kunnen zorgen. En zelfs met vrijgegeven specificaties, blijft het formaat eigendom van één enkele producent, wat altijd een risico inhoudt.¹⁷

Omwille van de bovenstaande redenen werd in 2002 een comité van afgevaardigden van de overheid, het bedrijfsleven en de academische wereld opgericht onder auspiciën van *the Association for Information and Image Management (AIIM)*¹⁸ en *the Association for Suppliers of Printing, Publishing and Converting Technologies (NPES)*¹⁹. Het doel van dit comité was “to specify the use of PDF in a manner that is specifically geared to long-term management and use”, die de naam Portable Document Format/ Archive, kortweg PDF/A, meekreeg.²⁰ In navolging van het gedane werk, werd een PDF/A Joint Working Group opgericht door ISO. Het resultaat kwam in 2005 onder de vorm van de eerste van een nieuwe familie standaarden voor document management: ISO 19005-1.²¹ PDF/A is een subset van de PDF Reference 1.4 die beter geschikt is voor archivalische doeleinden. De bedoeling is de *look-and-feel* van elektronische documenten te bewaren en toekomstige toegankelijkheid te ondersteunen. Dit, door te zorgen voor een kader voor het inkapselen van metadata en de logische structuur en semantische eigenschappen te definiëren.²² Het gebruik van PDF/A door het Vlaams Parlement kan enkel aan te raden zijn. De technische voordelen zijn hierboven aangehaald. Door de status van ISO-standaard bestaat een bepaalde zekerheid dat ondersteuning mogelijk blijft. Het gebruik van een standaard, indien duidelijk kenbaar gemaakt, kan enkel bijdragen aan een positieve evaluatie als betrouwbaar digitaal depot. Het enige nadeel (geldig voor alle ISO normen) ligt in het feit dat deze specificatie betalend is, en dus iets minder *open format* is dan de andere versies.²³

Een andere optie kunnen we vinden in de OpenOffice.org suite.²⁴ OpenOffice is een gemeenschap die als doel heeft “to create, ..., the leading international office suite that will run on all major platforms and provide access to all functionality and data through open-component based APIs

¹⁷ Boudrez, F. Standaarden voor digitale archiefdocumenten. Technisch rapport. Antwerpen: Stadsarchief, 2001 –2003, pp. 22-23

¹⁸ <http://www.aiim.org/standards.asp?ID=25013>, 12/02/06

¹⁹ <http://www.npes.org/standards/toolspdfa.html>, 12/02/06

²⁰ LeFurgy, W. PDF/A : Developing a file format for long-term preservation. In: *RLG Diginews*, Volume 7, Number 6, 15 december 2003, http://www.rlg.org/preserv/diginews/v7_n6_feature1.html, 20/06/06

²¹ ISO 19005-1: 2005, Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF (PDF/A)

²² PDF/A Joint Working Group, ISO 19005-1: 2005. PDF/A-1. Frequently Asked Questions (FAQs). July 10, 2006, p. 2, http://www.aiim.org/documents/standards/19005-1_FAQ.pdf, 20/07/06

²³ Stoehr, T. Un format PDF pour l'archivage a été normalisé, mais est-ce un format ouvert ? 6 novembre 2005, <http://formats-ouverts.org/blog/2005/11/06/604-un-format-pdf-pour-l-archivage-a-ete-normalise-mais-est-ce-un-format-ouvert->, 20/07/06

²⁴ <http://www.openoffice.org/>

and an XML-based file format.”²⁵ Ze biedt platformonafhankelijke applicaties aan voor kantoordocumenten als een *open-source* project.²⁶ Als basis voor haar documenten, wordt het *Open Document Format* (ODF of OpenDocument) van de *Organization for the Advancement of Structured Information Standards*, OASIS (niet te verwarren met OAIS), gehanteerd.²⁷ OASIS is een internationale, non-profit organisatie die sinds 1993 een drijvende kracht tracht te zijn achter de ontwikkeling, het samensmelten en de ingebruikneming van e-business standaarden.²⁸ OpenDocument is een rechtenvrij bestandsformaat gebaseerd op XML, geschikt voor tekstbestanden, spreadsheets, grafieken en grafische documenten.²⁹ Het gehele OpenOffice pakket kan gratis worden afgehaald van het internet. Microsoft documenten kunnen geopend worden in de verschillende applicaties en vervolgens opgeslagen worden in ODF. Gezien de code bekend is, moet het ook mogelijk zijn via een *tool* deze conversie rechtstreeks uit te voeren alvorens te archiveren.

Binnen de Linux gemeenschap is ODF al langer een geaccepteerde tegenhanger voor MS Office formaten. Daarbuiten begint de specificatie zijn plaats op te eisen. In België en Denemarken zal tegen 2008 alle officiële gegevensuitwisseling door de overheid met ODF worden gedaan.³⁰ Banco de Brasil heeft al haar volledige XP omgeving vervuild voor een OpenOffice omgeving.³¹ Google heeft zet zich achter het project en heeft zich aangesloten bij de *ODF Alliance*.³² Google Writely is een online teksverwerkingsprogramma gebaseerd op de ODF standaard. Als tegenreactie probeert Microsoft haar eigen Open XML formaat te promoten. Hoewel ze niet zal dienen als basis voor Office 2007, was er wel

²⁵ Mission Statement, <http://about.openoffice.org/index.html>, 15/03/06

²⁶ Open-source is computerprogrammatuur waarvan de broncode bekend is, met als doel aanpassingen te maken om een snellere vooruitgang te boeken. Wat dit nu exact inhoudt, werd omschreven door verschillende groepen, nl. the Open Source Initiative (OSI), the Free Software Foundation en Debian Free Software. OpenOffice hanteert de definitie opgesteld door the Open Source Initiative (<http://www.opensource.org/docs/definition.php>). OSI vat het als volgt samen: “Open source promotes software reliability and quality by supporting independent peer review and rapid evolution of source code. To be OSI certified, the software must be distributed under a license that guarantees the right to read, redistribute, modify, and use the software freely.” <http://www.opensource.org/advocacy/faq.php>, 20/07/06

²⁷ OASIS Open Document Format for Office Applications (OpenDocument) TC, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office, 20/07/06

²⁸ <http://www.oasis-open.org/who/>

²⁹ Cover, R. *Open Document Format for Office Applications (OpenDocument) v 1.0 an OASIS standard*. In: *XML Coverpages*, May 23, 2005,

<http://xml.coverpages.org/OpenDocumentV10-Standard.html>, 20/07/06

³⁰ Deckmyn, D. *Belgian government chooses OpenDocument*. In: *CNET News.com*, June 23, 2006, http://news.com.com/Belgian+government+chooses+OpenDocument/2100-7344_3-6087275.html?tag=nefd.top, 20/07/06; *Denmark’s resolution on open standards-updated*. In: *Groklaw*, June 3, 2006,

<http://www.groklaw.net/article.php?story=20060602210610876>, 20/07/06

³¹ *BB completes XP migration to open source*. In: *TMC Net*, May 30, 2006, <http://www.tmcnet.com/usubmit/2006/05/30/1662778.htm>, 20/07/06

³² ODF Alliance is een ngo, die het gebruik van ODF promoot opdat de publieke sector meer controle over de eigen *records* zou hebben. <http://www.odfalliance.org/>

sprake van steun aan de ontwikkeling van externe plug-ins waarmee MS Office ODF-bestanden zou kunnen lezen en opslaan.³³

OpenDocument bevat alle kenmerken om een geschikt archiveringsformaat te zijn: open, gedocumenteerd, onafhankelijk en de verspreiding neemt gestaag toe.

C. E-mail

Op dit moment wordt nog gebruik gemaakt van Novell GroupWise 6 als e-mailsysteem. Dit is een softwarepakket van Novell dat een berichten- en agendasytem biedt voor personen en bedrijven. Het omvat e-mail, een agenda, taak- en documentmanagement, een adresboek, notities, *instant messaging* en een beperkte set van archiveringsmogelijkheden. Agenda's, taken, adressen en notities kunnen met verschillende personen gedeeld worden, waardoor bepaalde afspraken makkelijker kunnen gemaakt worden. Ondertussen zijn versies 6.5 en 7 uitgekomen.³⁴ Het systeem werd in 2001 op de markt gebracht³⁵ en niet lang daarna in gebruik genomen op het Parlement. Stilaan wordt nu een fase ingegaan die *end-of-life* wordt genoemd. Deze verkoopsterm geeft aan dat een producent een aantal dingen niet meer zal doen (marketing, verkoop, promotie) of de ondersteuning afbouwen van een bepaald product.³⁶ Op de website van Novell is nu al geen informatie meer beschikbaar over versie 6, wel nog voor 6.5.

Op het einde van dit jaar wordt een nieuwe e-mailomgeving uitgebouwd omwille van die reden. Bovendien wordt nu te veel op de centrale servers bewaard. Normaal gezien zal meer op andere servers worden opgeslagen (hoewel dit nog niet vaststaat). De mails zullen centraal beheerd worden.

Er bestaat een externe server, maar die fungeert enkel als doorgeefluik voor de parlementariërs naar de eigen mail. Eens afgehaald, worden ze lokaal bewaard. Niet alle mails worden dus door IT bijgehouden. Voor de veiligheid worden wel dagelijks backups genomen van het hele systeem op een vastgesteld uur. Dit is een *snapshot*, die enkel gericht is op de korte termijn.

Nu kan men enkel in het gebouw van het Vlaams Parlement mails nagekeken worden. Een webinterface zal worden geïntroduceerd waardoor dit ook extern kan gebeuren. De oude mails zullen vervolgens geconverteerd worden naar het nieuwe systeem.

³³ Lai, E. Google backs OpenDocument format. In: *Computerworld*, July 11, 2006, http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=9001713&bcsi_scan_747D1649BA44C3E8=3s8ko7306n12R7CcILyLwoAAABzvpcB&bcsi_scan_filename=article.do, 20/07/06

³⁴ <http://www.novell.com/products/groupwise/>

³⁵ <http://www.novell.com/news/press/archive/2001/03/pr01031.html>, 19/07/06

³⁶ End-of-life. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/End-of-life>, 19/07/06

Berichten zijn in principe degelijk beveiligd. Eerst moet aangelogd worden, zowel lokaal als op het netwerk. Zonder paswoord kan men moeilijk binnen geraken. Een degelijke beveiliging van het systeem is één van de verkoopsargumenten van Novell. De CERT Coordination Center (Carnegie Mellon University) geeft voor de laatste versie aan dat GroupWise een hogere veiligheid waarborgt dan andere netwerken zoals MS Exchange of Lotus Notes.³⁷

De archiveringsfunctie die nu aanwezig is, laat toe mails op te slaan in een *directory* naar keuze. Een onderscheid wordt gemaakt tussen archiveren en opslaan. Ze kunnen worden bewaard in het formaat van het mailsysteem. Bij gebruik van de functie *opslaan als* wordt de mogelijkheid geboden de verschillende onderdelen van een bericht apart op te slaan. Het MIME gedeelte kan worden bewaard in een simpel *text* formaat en kan dus door eender welke tekstverwerker worden geopend, in een MS omgeving veelal door *kladblok*. De extensie in de naam is .882, die feitelijk onbestaand is (behalve voor Novell misschien). Door deze te veranderen naar .txt is het probleem wel opgelost. De naamgeving blijft echter voor elk bestand simpelweg Mime, waardoor deze bestanden enkel per één kunnen worden bewaard. Vervolgens moet dan handmatig een onderwerp worden gegeven. Naast de pure MIME gegevens, wordt het volledige bericht ook weergegeven in HTML-notatie.

Bijlagen blijven in het oorspronkelijke formaat en kunnen als dusdanig bewaard worden. Het bericht zelf wordt in de meeste gevallen als een Word-bestand opgeslagen, maar soms ook als HTML en een enkele keer als PDF. Als titel van het document wordt voor de bijlagen de oorspronkelijke naam en voor het bericht de onderwerpstitel overgenomen. In de Word- en PDF-bestanden worden een aantal basisgegevens mee vermeld: het adres van de verzender en de ontvanger, datum van verzenden en adres(sen) die in kopie staan (CC). Dit ontbreekt echter bij de HTML-files.

Qua uitzicht zijn de berichten in hun originele omgeving vrij eenvoudig, *plain text* om meer precies te zijn. De *significant properties* van de berichten zijn bijgevolg de eigenlijke inhoud, gegevens over de verzender en ontvanger, datum van ontvangst (en eventueel van verzenden), personen die in kopie staan en de bijlagen. Tijdens de dynamische fase kunnen mails in het oorspronkelijke formaat worden bewaard, omwille van het gebruiksgemak. Vanaf de (semi-) statische fase kan ze worden omgezet naar een AIP. Die kan dan mogelijk uit een XML-bestand bestaan, waaraan bepaalde metadata worden gekoppeld. In een DTD (Document Type Definition) of XML Schema wordt de structuur van de archiefdocumenten of -bestanden vastgelegd. Voor MS Outlook is een *tool* ter beschikking die werd ontwikkeld door het DAVID-project en Testbed

³⁷ Novell Workgroup, *A collaboration solution that supports your business goals. Business White Paper*. Novell, 2006, p. 5, <http://www.novell.com/collateral/4621449/4621449.pdf>, 20/07/06

Digitale Bewaring.³⁸ Het bericht ter beschikking stellen, als DIP, kan bijvoorbeeld als PDF/A-bestand.

Naar aanleiding van de uitbouw van een nieuwe e-mailomgeving, werd een richtlijn door de archiefdienst overgemaakt aan ICT, waarvan de belangrijkste punten zijn:

- “Integratie van e-mail met andere digitale documenten moet mogelijk zijn (bvb. opslag in de digitale mappenstructuur op de gemeenschappelijke serverschijf)
- Metadata: mogelijkheid tot het linken van een e-mail-bericht aan het grotere geheel waarop het mailbericht betrekking heeft (bvb. de ‘dossiercode’)
- Metadata: link met bijlagen moet bewaard blijven
- Metadata-context: de essentiële headergegevens moeten bewaard blijven
- Inhoud van mail (incl. bijgevoegde bestanden) + bijlagen moeten kunnen bewaard worden
- Migratie naar een geschikt archiveringsformaat moet mogelijk zijn. Voor de mail en zijn bijlagen is dit niet noodzakelijk hetzelfde. Voor het eigenlijke mailbericht heeft XML ingang gevonden
- Eens gearchiveerd moet een goede ontsluiting mogelijk zijn. Verschillende opties of combinaties kunnen mogelijk zijn
- Het is belangrijk dat de digitale archiveringsstrategie gebruikersvriendelijk en zoveel als mogelijk automatisch verloopt”³⁹

Op dit moment lijkt het gebruik van XML de beste mogelijkheden te bieden op het vlak van duurzame bewaring. De *eXtensible Markup Language* is het resultaat van een samenwerkingsverband tussen de academische en industriële wereld. De ontwikkeling startte in 1996 en werd in '98 door het W3C als *recommendation* ⁴⁰ vastgelegd. Het gaat hier eigenlijk om een gesimplificeerde subset van de *Standard Generalized Markup Language* (SGML), die behoort tot dezelfde familie als HTML.⁴¹ SGML (dus ook XML) is een meta taal voor de beschrijving van een breed scala van data en kan eveneens de data zelf bevatten. Deze markuptaal heeft als doel het delen van data via verschillende systemen te

³⁸ Boudrez, F. Digitaal archiveren van archiefdocumenten. Casus: e-mail. Antwerpen : Stadsarchief Antwerpen, Onuitgegeven lezing, 6 maart 2006; Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 4: Bewaren van e-mail. Den Haag, ICTU, 2003, pp. 92-94

³⁹ Directie Informatie, Archief en Biografische Dienst, Aandachtspunten voor e-mailarchivering (beknopt). Interne nota, 3 april 2006

⁴⁰ De specificaties zijn vrij te downloaden. Bray, T.; Paoli, J.; Sperberg-McQueen, C.M.; Maler, E.; Yergeau, F. (eds.) Extensible Markup Language (XML) 1.0 (Third Edition). W3C Recommendation, 4 February 2004, <http://www.w3.org/TR/REC-xml/>, 16/03/06

⁴¹ SGML is een zeer sterk uitgewerkte markuptaal. Het is ook een zeer complexe taal, die veel plaats in beslag neemt. De SGML specificatie is 600 pagina's lang, die van XML 26. Dit bemoeilijkt de uitwisselingsmogelijkheden en beperkt het gebruiksgemak. Hierdoor is SGML nooit echt populair geworden. Boudrez, F. <XML/> en digitaal archiveren. Antwerpen: Stadsarchief, 2002, p. 2

vergemakkelijken. Informatie wordt beschreven met menselijk leesbare tekst⁴² in de vorm van een boomstructuur. De formele syntax staat vast, waardoor alle software die XML kan verwerken op zijn minst de relatieve plaatsing van informatie kan lezen en begrijpen. Via een *Document Type Definition* of een XML Schema kunnen bepaalde beperkingen bijgevoegd worden. Op die manier kan iedereen de taal aanpassen naar zijn eigen noden en wensen.⁴³ Opmaak en inhoud zijn van elkaar gescheiden. De inhoud wordt in het XML-bestand bewaard, de opmaak kan in een *Cascading Style Sheet* (CSS) of een XSL bestand worden vastgelegd.

XML heeft vanuit archivalisch oogpunt een aantal voordelen. Het is een publieke *de facto* standaard, goedgekeurd door een internationale standaardiseringsinstantie, zijnde W3C. Ze is platform- en softwareonafhankelijk en licentievrij. Omdat de gebruiker zijn eigen tags en attributen definieert, kunnen deze niet verouderd raken door versiewijzigingen. XML-data kunnen in een andere structuur of outputformaat worden getoond. Voor de omzetting is enkel een stylesheet vereist.⁴⁴ Omdat alles in universele, tekstuele gegevens worden geëncodeerd, blijven de gegevens in menselijk leesbare vorm. Ze blijven bovendien makkelijk te interpreteren door het zelfbeschrijvende karakter van de tags en attributen. Zoekmogelijkheden zijn vrij uitgebreid. Er kan gezocht via *full text scans*, maar ook op basis van de structuur, metadata en semantische tags. De structuur wordt op een expliciete en formele wijze bewaard door de hiërarchische relatie van de markuptags. Bijkomende informatie kan gevonden worden in een DTD of XML Schema.⁴⁵ Binnen één document kunnen bovendien zowel sterk gestructureerde informatie, zoals een e-mailheader, als ongestructureerde data, zoals de e-mailbody, worden bijgehouden.⁴⁶

Er zijn ook een nadelen aan XML verbonden. Door het gebruik van tekstuele encoding (in plaats van binaire) nemen XML-bestanden vrij veel plaats in beslag. Om de performantie te verhogen wordt bij netwerkuitwisselingen veelal compressie toegepast. Door de hoge redundantie veroorzaakt door het terugkomen van dezelfde tags, kunnen hoge compressieratio's worden bereikt. Deze zijn evenwel lossless. Het lezen van en het omzetten naar XML neemt tijd in beslag. Het is een uitbreidbare taal, waardoor een wildgroei aan verschijningsvormen zich

⁴² Gebaseerd op Unicode (cfr. ASCII), waardoor zo goed als alle gekende talen kunnen worden opgenomen.

⁴³ Extensible Markup Language. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/XML>, 16/03/06

⁴⁴ Een voorbeeld voor een XML stylesheet voor e-mails: Boudrez, F.; Dekeyser, H. & Van den Eynde, S. Archiveren van e-mail. DAVID Rapport 5. 2^{de} editie. Antwerpen: Stadsarchief DAVID Project, 2003, 94-96

⁴⁵ Voorbeelden van een XML DTD of Schema voor e-mails zijn hier te vinden: Boudrez, F.; Dekeyser, H. & Van den Eynde, S. Archiveren van e-mail. DAVID Rapport 5. 2^{de} editie. Antwerpen: Stadsarchief DAVID Project, 2003, pp. 90-93; Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 4: Bewaren van e-mail. Den Haag, ICTU, 2003, pp. 95-101

⁴⁶ Boudrez, F. <XML/> en digitaal archiveren. Antwerpen: Stadsarchief, 2002, p. 10-12

kan voordoen.⁴⁷ Dit laatste valt daarentegen op te lossen voor het Parlement door de gebruikte vorm duidelijk te definiëren in een Schema.

XML als bewaarstrategie wordt door het Nederlandse *Testbed Digitale Bewaring*⁴⁸ en het Antwerpse DAVID-project⁴⁹ naar voren geschoven als de beste methode voor e-mailarchivering op dit moment. De bestanden kunnen naar een XML-bestand worden omgezet of er kan gewerkt worden met een XML-wrapper. Bij deze strategie wordt het te bewaren bestand (of bestanden indien bijlagen worden toegevoegd) in het oorspronkelijke en/of geconverteerde formaat bewaard. Daar rond wordt een inhoudsbeschrijvende wrapper geplaatst. Door te werken met verschillende lagen kan in die wrapper verschillende soorten metadata worden geplaatst: technische gegevens, beschrijvingen van het object, toegankelijkheidsvoorwaarden, de historiek van de bewaaracties, etc.⁵⁰ Aan de hand van de toegevoegde metadata worden gerichte zoekacties mogelijk en worden nieuwe bescheiden gemakkelijker ingevoegd in een inventaris.

D. Databanken

In het Vlaams Parlement worden verschillende types van databanken gebruikt. Alle productiedatabanken zijn op basis van Oracle: de centrale databank (Proteus), de ombudsdienst, het Kinderrechten-commisariaat, Human Resources.

In 1970 publiceerde Dr. Edgar Codd het ophefmakende artikel *A relational model of data for large shared data banks*,⁵¹ waarmee hij de basis legde voor het relationele databasemodel. Na een moeilijke start kenden zijn ideeën een sterke opmars en vandaag is dit het dominante model op het vlak van databaseprogrammatuur.⁵² In 1979 bracht Relational Software, Inc. de eerste commerciële relationele databank op de markt: Oracle V.2. Later werd de naam van het bedrijf veranderd naar haar belangrijkste product. Als verkoopsargument haalde (en dit is nog steeds waar) aan dat ze op meerdere platformen (70 vandaag) kan gedraaid worden, data uitwisselbaar zijn tussen die platformen en standaard SQL hanteert zoals ze werd bepaald door ANSI.⁵³

⁴⁷ Boudrez, F. *Idem*, p. 13

⁴⁸ Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 4: Bewaren van e-mail. Den Haag, ICTU, 2003, pp. 55-60

⁴⁹ Boudrez, F. E-mailarchieven. E-mails: hoe bewaren en goed archiveren? Antwerpen: Stadsarchief DAVID Project, 2003, pp. 19-20

⁵⁰ Testbed Digitale Bewaring, *Ibidem*, pp. 37-39

⁵¹ Codd, E.F. A relational model for large shared data banks. In: *Communications of the ACM*, Vol. 13, No 6, Association for Computing Machinery Inc., June 1970, pp. 377-387, zie ook: <http://www.acm.org/classics/nov95/toc.html>, 19/03/06

⁵² Committee on Innovations in Computing and Communications: Lessons from History, Funding a revolution. Government support for computing research. Washington: National Academy Press, 1999, <http://www.nap.edu/readingroom/books/far/ch6.html>, 20/06/07

⁵³ Greenwald, R.; Stackowiack, R. & Stern, R. Oracle essentials: Oracle9i, Oracle8i & Oracle8 (2nd edition). US: O'Reilly & Associates, 2001, <http://www.developer.com/db/article.php/1582621>, 07/04/07

In 1995 werd gestart met de uitbouw van het geïntegreerde informatiesysteem van het Vlaams Parlement onder de naam Vlaris (Vlaamse Raad InformatieSysteem). Het hoofddoel was in ondersteuning te voorzien van de decreetgevende taken. Ze bestaat uit verschillende met elkaar verbonden deelsystemen, nl. het Documentair Fonds (parlementaire documenten vanaf 1995) en de modules die verband houden met Procedurebeheer, nl. Agenda, Samenstelling en Personeel, Registers en Statistiek. Het deelsysteem Procedurebeheer geeft de stand van een bepaald parlementair initiatief in de behandelingsprocedure vanaf de indiening tot de afhandeling.⁵⁴ Als opvolger werd vanaf 1997 gewerkt aan Proteus. Hiermee werd een aanzienlijke functionele uitbreiding (ook naar niet-decreetgevende diensten toe) beoogd. Het Documentair Fonds werd in het begin van 1997 vervangen door een nieuwe databank die alle stukken in PDF-formaat bevat. Deze is volledig raadpleegbaar via het internet. Vanaf dan werd ook begonnen aan een personendatabank met biografische gegevens van de Vlaamse volksvertegenwoordigers.⁵⁵ Na een re-engineering van een eerste versie werd Proteus de opvolger van Vlaris en ingeschakeld als geautomatiseerd informatiesysteem, dat als centrale databank de procedurele gegevens over de parlementaire werking bevat. “Die gegevens worden, in de geest van het “Glazen Huis”, via de website van het Vlaams Parlement ter beschikking gesteld aan de Vlaamse bevolking.” Bijkomend werd een inspanning geleverd om alle parlementaire documenten sinds 1971 te digitaliseren en vervolgens werden deze raadpleegbaar gemaakt via de website.⁵⁶

Van CIPAL wordt gebruik gemaakt van System Z voor de loonberekening en Aphrodite voor Human Resources. CIPAL is een dienstverlenende, intergemeentelijke vereniging, die automatiseringsprojecten uitwerkt voor “administratieve, technische en beheerstoepassingen, voornamelijk in de publieke sector (lokale, regionale, gewestelijke en federale overheidsdiensten).”⁵⁷

De boekhouding hanteert SAP. SAP AG (*Systeme, Anwendungen und Produkte in der Datenverarbeitung*) is een van oorsprong Duitse multinational (°1972) die applicatiesoftware maakt gericht op *real-time business processing*. Het bedrijf ontwikkelt *software solutions* en verkoopt vervolgens licenties hierop aan ondernemingen (zowel multinationals als KMO's) en overheidsinstellingen. Sinds de jaren negentig wordt gewerkt binnen het client-server concept en ligt de nadruk op uniforme grafische interfaces, consistent gebruik van relationele databanken en de

⁵⁴ Van Esbroeck, K. (ed.) Legislatuurverslag van het Vlaams Parlement, 1995-1999. Brussel: Externe Relaties Vlaams Parlement, 1999, p. 136

⁵⁵ Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 1997-1998. Brussel: Externe Relaties Vlaams Parlement, 1998, pp. 96-97

⁵⁶ Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 1999 en 1999-2000. Brussel: Externe Relaties Vlaams Parlement, 2000, p. 109

⁵⁷ <http://www.cipal.be/>

mogelijkheid te draaien op meerdere platformen.⁵⁸ De historische boekhouding wordt nog onderhouden in een Oracle databank. Voor de huidige boekhouding wordt, zoals gezegd, met een systeem van SAP gewerkt. Daar onder echter draait nog een specifieke Oracle databank die enkel kan bevraagd worden door middel van SAP.

De individuele gebruiker zal zich behelpen met een desktoptoepassing zoals MS Access. Dit is een relationele DBMS, met als onderliggende basis Jet Database Engine, die voornamelijk gebruikt wordt door kleine ondernemingen en hobbyisten voor *ad hoc* systemen voor datamanipulatie.⁵⁹ Access kan met SQL bevraagd worden. Er zijn wel een aantal verschillen met standaard SQL (sommige functies van SQL-92 werken niet in Access).

Het grote voordeel van Access als desktoptoepassing, ligt in het feit dat men zonder voorkennis een simpele database kan uitwerken aan de hand van wizards⁶⁰ (men zou de applicatie evengoed *database for dummies* kunnen genoemd hebben). Tegelijkertijd is dit het grootste nadeel. Microsoft heeft de onhebbelijke gewoonte *intelligent agents* voor van alles en nog wat toe te voegen aan haar programma's. Dit is zeer praktisch voor de gemiddelde gebruiker van een pc. De meeste mensen zijn geen volleerde informatici, dus soms is het zeer handig om een beetje hulp te krijgen. In de meeste gevallen is het ook beter dat de essentiële bestanddelen van een applicatie worden afgeschermd. Voor IT-professionals kan dit zeer irritant zijn omdat de voornoemde *agents* sommige dingen in hun plaats wil doen, wat ze soms niet willen dat het dat doet. Microsoft hanteert hierbij de gevoelsmatige logica van de gemiddelde eindgebruiker, wat soms lijnrecht staat tegenover de formele logica toegepast door informatici. Bovendien is voor de uitbouw van grote databanken en gebruik in netwerken Access niet bijzonder geschikt. In de ontwikkeling van systemen binnen een client-serverstructuur wordt meer vertrouwen gesteld in producten zoals Oracle, DB2, Microsoft SQL Server, PostgreSQL en MySQL.⁶¹ Binnen het Parlement wordt het serieuzere programmeer werk met een welbepaalde variant van Linux gedaan.

In het Vlaams Parlement worden de bestanden bewaard op de centrale harde schijf, in het formaat van de DBMS zelf. Enkel de parlementaire stukken worden op een andere server opgeslagen dan de databank (hierop bevinden zich alleen pointers). Elke avond wordt een exportformaat gegenereerd. Deze dient als backup voor een snelle recovery. Daarnaast wordt ook een *disc-to-disc* backup gemaakt voor een langzame recovery. Veelal gebeurt het op magnetische banden, maar soms ook op CD of DVD.

⁵⁸ SAP history: From start-up software vendor to global market leader. <http://www.sap.com/company/history.epx>, 20/07/06; SAP AG. In: *Wikipedia, Die freie Enzyklopädie*, http://de.wikipedia.org/wiki/SAP_AG, 20/07/06; SAP: Profile for SAP Aktiengesell ADS. In: *Yahoo! Fincance*, <http://finance.yahoo.com/q/pr?s=SAP>, 20/07/06

⁵⁹ Microsoft Access. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Microsoft_Access, 20/07/06

⁶⁰ Wizards zijn hulpjes, verder uitwerken.

⁶¹ Microsoft Access. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Microsoft_Access, 20/07/06

In een later fase zal een externe schijvenkabinet geïnstalleerd worden op een grote centrale server die data overzet via een SAN-kabel.⁶² De documenten die bewaard worden op de centrale server, blijven in het oorspronkelijk formaat. Bij iedere upgrade van Oracle, worden de gegevens wel mee geconverteerd naar de laatste versie. De gegevens die werden gegenereerd door een oudere DBMS, werden gemigreerd naar het Oracle-formaat. Gegevens over het gebruik van de databanken (in de vorm van een logboek bijvoorbeeld) werden niet bijgehouden.

Samenvattend kan gezegd worden dat de gebruikte databanken relationeel van aard zijn en bevraagd kunnen worden aan de hand van SQL. In eerste instantie moet nog, in overeenstemming met de archiefvormers en de Directie Informatica vastgesteld worden wat de essentiële eigenschappen zijn van de verschillende databanken. Een simpele oplossing kan hier nog niet voorgesteld worden. Testbed Digitale Bewaring⁶³ stelt opnieuw het gebruik van XML voor. De verschillende onderzochte tools waren in staat om verschillende types database om te zetten. Een omzetting van alle gegevenstabellen via één, eenvoudige handeling was echter niet mogelijk. Om die reden hebben ze zelf een tool ontwikkeld op basis van Java om dit te bewerkstelligen. Ze vermelden ook SIARD, *Software Independant Archiving from Relational Databases*. Dit project heeft een tool ontwikkeld voor de archivering van relationele databanken. "SIARD bestaat uit een geavanceerde verzameling toepassingen waarmee: precies kan worden aangegeven welke componenten van de database moeten worden geconverteerd, diverse metagegevens kunnen worden ingevoerd en een conversie kan worden uitgevoerd naar een groep bestanden die bestaat uit XML, SQL3 en de platte inhoud van de database."⁶⁴ Dit Zwitserse project wil een platformonafhankelijk en gebruiksvriendelijk product aanbieden.⁶⁵ De metadata wordt naar XML omgezet, de structuur naar SQL3 DDL, de tabelgegevens als een *flat file*.⁶⁶ De eigenlijke tool is voorlopig nog niet op de markt gebracht.

⁶² "Storage Area Network (SAN) is a high-speed subnetwork of shared storage devices. A storage device is a machine that contains nothing but a disk or disks for storing data." SAN. In: *Webopedia*, 2006, <http://www.webopedia.com/TERM/S/SAN.html>, 20/07/06; voor een uitgebreide introductie, zie: Tate, J.; Lucchese, F.; Moore, R. *Introduction to Storage Area Networks*. IBM Redbooks, 2004, <http://www.redbooks.ibm.com/abstracts/sg245470.html>, 20/07/06

⁶³ Op het vlak van archivering zijn zij de enige voor zover ik weet die een uitgebreid overzicht geven van databanken.

⁶⁴ Testbed Digitale Bewaring, *Van digitale vluchtigheid naar digitaal houvast, Deel 1: Kosten- en beslissingsmodellen/Functionele specificaties. Bewaren van databases*. Den Haag: ICTU, 2003, p. 50-51

⁶⁵ Ross, S.; Donnelly, M. & Dobрева, M. *Emerging technologies for the cultural and scientific heritage sector*. DigiCULT Technology Watch Report 2, February 2004, p. 57 <http://www.digicult.info/downloads/html/1082026370/1082026370.html>, 16/07/06

⁶⁶ Heuscher, S. *SIARD under the hood. Technical aspects of SIARD*. Presentation at the Urbino Erpanet Congress, 10 april 2003, www.erpanet.org/events/2003/bern/presentations/WS%20Bern%20Presentation%20Heuscher.ppt, 16/07/06

Testbed geeft aan hoe een database het beste kan bewaard worden.⁶⁷ Vijf elementen worden aangegeven: de originele database, XML-overzichtsbestand, de toepassing, bewaarlogbestand en metadata:

- *Originele database*: zij raden aan om het originele bestand mee te archiveren, om een maximale flexibiliteit te bieden naar toekomstige bewaarstrategieën toe.
- *XML-overzichtsbestand*: dit bestaat uit een overzicht van de aanwezige tabellen, de onderlinge relaties tussen die tabellen in een DTD of XML-Schema, de inhoud (XML inhoudsbestanden) en de structuur van de tabellen en views (DTD of XML-Schema).
- *De toepassing*: deze hoeft in principe niet mee gearcheveerd te worden, tenzij volledige functionaliteit als essentiële eigenschap wordt gezien. Dan moet evenwel gekozen worden voor emulatie als strategie. Wat wel wordt bewaard, is de systeemdocumentatie en de gebruikershandleiding. Dit omwille van informatie over functionaliteiten, de gebruikte queries en de gebruikersinterface.
- *Bewaarlogbestand*: hierin wordt alle informatie bijgehouden over de ondernomen bewaaractiviteiten, en eventueel over bewaar- en toegangseisen. Dit bestand wordt aangemaakt bij de omzetting naar XML.
- *Metadata*: een metadata moet worden verzameld en opgeslagen op het moment dat het document wordt aangemaakt.

Concreet betekent dit dat XML als bewaarstrategie een reële mogelijkheid biedt. Sinds de publicatie van de aangehaalde Testbed-publicatie zijn alweer enkele jaren verstreken. Een marktonderzoek zou moeten uitgevoerd worden naar de beschikbaarheid van gebruiksvriendelijke migratietools.⁶⁸

E. Websites

De eerste versie van de website werd in november 1998 in werking gesteld. Sindsdien zijn verschillende *reengineerings* uitgevoerd. Deze werden op het niveau van de directieraad beslist en gaan over periodes van twee tot drie jaar. Het betrof dan voornamelijk veranderingen qua interface (gemakkelijkere opzoekingen, ander lay-out, ...). De eerste site had als doelstelling efficiënt, sober en stijlvol te zijn. De officiële publicaties werden geleidelijk ter beschikking gesteld online, in het kader van het Proteusproject.⁶⁹ Sinds de openstelling op 21 januari 1999 werd de site permanent *up-to-date* gehouden. Naar aanleiding van de parlementsverkiezingen van dat jaar werd ze grondig geactualiseerd. Later

⁶⁷ Testbed Digitale Bewaring, *Idem*, pp. 52-55

⁶⁸ Een simpele Google search levert al duizenden hits op. Bvb.

http://www.stylusstudio.com/db_to_xml_mapper.html, <http://www.programurl.com/aiob-xml-database-converter-and-aiob-xml-editor.htm>,
<http://www.xmldatabases.org/radio/xmlDatabases/projects/FreeDBtoXML/>

⁶⁹ Van Esbroeck, K. (ed.) *Jaarverslag van het Vlaams Parlement, 1997-1998*. Brussel: Externe Relaties Vlaams Parlement, 1998, pp. 95-96

werden nog een aantal kleinere aanpassingen uitgevoerd aan de bestaande structuur, het *design* en het concept. Een aanpassing werd in 2000 doorgevoerd omwille van de link tussen de site en Proteus.⁷⁰ Aan het projectteam re-engineering website werd de opdracht gegeven om: alle internettoepassingen gelijkvormig en transparant te maken (gelijke vorm statische en dynamische pagina's), voor Proteus een gebruiksvriendelijke interface te ontwikkelen, een nieuwe lay-out te ontwerpen, het beheer van de website, het extranet en het intranet op elkaar af te stemmen.⁷¹ Op 5 december 2001 werd de vernieuwde site officieel voorgesteld ter gelegenheid van de viering van het dertigjarig bestaan van het Vlaams Parlement. Het resultaat liet zich zien in een verbeterde functionele structuur in een eigentijdse lay-out⁷², die verder werd geoptimaliseerd tijdens de daarop volgende jaren.⁷³

Aan het Parlement zijn verschillende websites verbonden. Als uithangbord naar de "buitenwereld" is de externe website te vinden op <http://www.vlaamsparlement.be/>. Deze bestaat voornamelijk uit statische HTML-pagina's gebaseerd op een Oracle-programmering. Qua webprogrammering wordt gewerkt met Oracle Application Server. Volgens de *salespitch* biedt ze een "comprehensive solution for developing, integrating, and deploying your enterprise's applications, portals, and Web services"⁷⁴ aan. Ze laat gebruikers toe met een webbrowser HTML-pagina's (statische inhoud) en data van Oracle databanken (dynamische inhoud) te raadplegen. Door middel van plug-ins kan PL/SQL code, Java, Perl of een andere programmeertaal code uitgevoerd worden.⁷⁵ Naast de statische inhoud, vinden we een zoekmachine, waarmee alle stukken, vragen en interpellaties, schriftelijke vragen, agenda's, verslagen en handelingen vanaf 1995 kunnen opgezocht worden. Via de zoekfunctie wordt een verbinding gemaakt met de centrale databank. Documenten worden vervolgens gepresenteerd in PDF, Word of HTML. Andere documentatie wordt toegevoegd op een statische pagina in dezelfde bestandsformaten. Zo vinden we onder meer de 11 juli-toespraak van aftredend voorzitter Norbert De Batselier en een PDF-versie van het biografisch lexicon terug. Afbeeldingen van oa. het gebouw, politici, doorklik icoontjes, etc. zijn eveneens terug te vinden. Elke woensdag wordt de plenaire vergadering gehouden in de koepelzaal. Deze kan rechtstreeks

⁷⁰ Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 1999 en 1999-2000. Brussel: Externe Relaties Vlaams Parlement, 2000, p. 109

⁷¹ Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2000-2001. Brussel: Externe Relaties Vlaams Parlement, 2001, p. 95

⁷² Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2001-2002. Brussel: Externe Relaties Vlaams Parlement, 2002, p. 115

⁷³ Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2002-2003. Brussel: Externe Relaties Vlaams Parlement, 2003, p. 111-112 en Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2004 en 2004-2005. Brussel: Externe Relaties Vlaams Parlement, 2005, p. 130-132

⁷⁴ <http://www.oracle.com/appserver/index.html>

⁷⁵ Naudé, F. Oracle Internet/Web Application Server (iAS) FAQ. Version 2.82, 2001, <http://www.oraFAQ.com/faqweb.htm#WHAT>, 22/03/06

worden mee gevolgd op de site. De bewaring hiervan gebeurt op een ander niveau (hier kom ik later op terug).

Voor de medewerkers wordt ook een intranet ter beschikking gesteld (<https://intranet.vlaamsparlement.be/DIS/>). De interne site wordt gepresenteerd in een andere, meer functionele lay-out. Ook hier worden de documenten, die terug te vinden zijn op de externe site, ter beschikking gesteld. Hier bovenop vinden we bijkomende bestanden. Onder de hoofding *Parlementair werk* bijvoorbeeld, kunnen een aantal hulpmiddelen worden geraadpleegd: reglementen en protocollen in een word-bestand, het huishoudelijk reglement van de Directieraad, een link naar de federale databank wetgeving (http://www.juridat.be/cgi_loi/wetgeving.pl), het beleidsplan van de voorzitter, etc. Onder *Documentatie* zien we een link naar Van Dale online, sjablonen voor nota's en PowerPoint presentaties, logo's van het Vlaams Parlement, etc. Apart is er nog een intern telefoonboek, met een apart overzicht in Excel. Op het vlak van archivering zijn nog geen concrete plannen uitgewerkt voor het intranet. De aanwezige documenten zullen bewaard worden, eventueel wordt nog een *walk-through* video gemaakt, maar dit staat nog niet vast.

En dan zijn er nog de afzonderlijke sites van de Paraparlamentaire Instellingen (PPI's), namelijk het Kinderrechtencommissariaat (<http://www.kinderrechtencommissariaat.be/>), Vlaamse Ombudsdienst (<http://www.vlaamseombudsdienst.be/>), Vlaams Instituut voor Wetenschappelijke en Technologisch aspectenonderzoek (viWTA, <http://www.viwta.be/>). Deze instellingen hebben een duidelijke band met het Vlaams Parlement, maar werken volledig autonoom. Ze zijn zelf verantwoordelijk voor het onderhoud van hun archief, dus ook voor de eigen website. Delen van sommige sites worden weliswaar gehost op de server van het Parlement. ICT biedt een ondersteunende functie. Maar meer ook niet.

Qua webprogrammering wordt gewerkt met Oracle Application Server. Volgens de *salespitch* biedt ze een “comprehensive solution for developing, integrating, and deploying your enterprise’s applications, portals, and Web services”⁷⁶ aan. Ze laat gebruikers toe met een webbrowser HTML-pagina's (statische inhoud) en data van Oracle databanken (dynamische inhoud). Door middel van plug-ins PL/SQL code, Java, Perl of een andere programmeertaal code uitgevoerd worden.⁷⁷ Voor de *deployment* van programmatuur wordt gebruik gemaakt van OC4J's. *Oracle Containers For J2EE* is een gebruiksvriendelijke (i.e. voor *web developers*), en snelle J2EE container.⁷⁸ *Java 2 Enterprise Edition (J2EE)*⁷⁹ is een ontwikkelomgeving

⁷⁶ <http://www.oracle.com/appserver/index.html>

⁷⁷ Naudé, F. *Oracle Internet/Web Application Server (iAS) FAQ*. Version 2.82, 2001, <http://www.orafaq.com/faqweb.htm#WHAT>, 22/03/06

⁷⁸ Panda, D. *Oracle Containers for J2EE (OC4J)*. In: *OnJava.com, The independent source for enterprise Java*, 2002, <http://www.onjava.com/pub/a/onjava/2002/01/16/oracle.html>, 15/02/06

⁷⁹ <http://java.sun.com/javaee/>

van Sun Microsystems, die een component-gebaseerde aanpak voor het ontwerpen, ontwikkelen, samenstellen en gebruiken van bedrijfsapplicaties biedt. Ze verzorgt onder meer een ondersteuning voor webservices dmv een geïntegreerde data-uitwisseling op basis van XML.⁸⁰

Alvorens aan websitearchivering te doen moeten op het vlak van procedures en beleid eerst een aantal vragen worden beantwoord. Samengevat: WAT wordt HOE door WIE op welk tijdstip (WANNEER) gearchiveerd?⁸¹

De website maakt deel uit van een informatiesysteem. Daarom moet eerst vastgesteld worden welke delen van de site wel of niet in aanmerking komen voor archivering in overleg met de webmaster en de archiefvormers. Vragen die aan hen werden voorgelegd zijn⁸²:

- welke doelstellingen lagen/liggen aan de basis voor het uitbouwen van de site?
- wat is het profiel (unieke informatie, statisch, dynamisch, interactief)?
- wat is de visibiliteit (doelpubliek, aantal bezoekers)?
- welke risico's loopt men wanneer de website of een bepaald onderdeel niet wordt gearchiveerd?

Voor de archivering van een site komen meerdere onderdelen in aanmerking⁸³:

- *webpagina's*: het uitgangspunt is een zo getrouw mogelijke reconstructie van het geheel van de site. Hierbij moet rekening worden gehouden met context, inhoud, structuur, *look-and-feel* en functionaliteiten (hyperlinks, animaties, ...). Een archiefdocument heeft een vast karakter. Voor een website is dat minder evident, omwille van de mogelijke verschillen in inhoud en tijd (snelle opeenvolging van updates), karakter (afhankelijk van gebruikersvoorkeuren of -rechten), uitzicht (afhankelijk van voorkeuren, rechten, browser, ...)
- *deep web*: dit is een verzamelterm voor alle inhoud beschikbaar via een site die wordt beheerd in content management systemen, databanken, documentbeheerssystemen of op fileservers. De inhoud hiervan kan zeer uiteenlopend zijn en wordt beter bewaard op het niveau van die toepassingen met een verwijzing naar de site
- de neerslag van transacties en handelingen: via een webformulier kunnen gebruikers gegevens doorsturen naar de webserver of

⁸⁰ Java 2 Enterprise Edition. In: *Wikipedia, De Vrije Encyclopedie*, <http://nl.wikipedia.org/wiki/J2EE>, 15/02/06

⁸¹ Boudrez, F. Archiveren van websites: een kwestie van waardering en 'capture'. Antwerpen: Expertisecentrum DAVID vzw, 2005, p. 6

⁸² Archivering van de website, interne nota als voorbereiding op de vergadering van 10/11/2005

⁸³ Boudrez, F. *Idem*, pp. 6-9

bepaalde handelingen doorvoeren.⁸⁴ De neerslag hiervan wordt meestal in een databank bewaard, waardoor bewaring op dat niveau kan worden uitgevoerd

- metadata: de verantwoordelijkheid voor het bijhouden en verzamelen hiervan wordt gedeeld door de webbeheerders, de archiefvormers en de archivaris. Van bij het ontwerp zou de nodige documentatie moeten bijgehouden worden, die wordt aangevuld tijdens de verdere levensloop van de site (notatie van belangrijke veranderingen). Bij de archivering wordt deze informatie mee in het depot opgenomen

Er zijn verschillende manieren om webpagina's te archiveren.⁸⁵ Een eerste methode bestaat erin om de broncode te bewaren. De bronbestanden worden dan in hun oorspronkelijke formaat overgenomen van de webserver. Dit is technisch gemakkelijk, maar enkel bruikbaar indien de pagina's in (X)HTML beschikbaar zijn van op de server, reconstrueerbaar zijn zonder webservertechnologie, aan elkaar gelinkt zijn via relatieve pathaanduidingen (idem voor *stylesheets*). In essentie is dit een back up methode en dus niet echt geschikt voor archivering.

Een tweede manier is het archiveren van een *snapshot*. Aan de hand van een webharvester⁸⁶ wordt een lokale "foto" gemaakt door alle pagina's en bijhorende grafische elementen naar een harde schijf te kopiëren. De HTML-pagina's zoals ze te zien zijn in een webbrowser worden bewaard.

Elke unieke webpagina kan volautomatisch gearchiveerd worden telkens wanneer ze voor de eerste keer wordt aangevraagd. Op die manier kunnen alle mogelijke versies van pagina's (gegenereerd op basis van alle mogelijke instellingen) bewaard worden. Op die manier kan de gearchiveerde site niet in zijn geheel worden geraadpleegd (enkel de individuele pagina's) en worden niet opgevraagde pagina's niet mee opgenomen in het archief.

Om de originele *look-and-feel* vast te leggen kan een surfessie gefilmd worden aan de hand van een screenrecorder.

Tijdens een vergadering, gehouden op 21 juni 2006 werden afspraken gemaakt over de te volgen bewaarstrategie. Voor de archivering van de oude site werden de volgende zaken afgesproken:

⁸⁴ Cfr. het voornemen van het Vlaams Parlement om alle Schriftelijke Vragen volledig digitaal te maken in de toekomst. Een effectieve regeling staat nog niet vast, maar één mogelijkheid is het inter- of intranet hiervoor te gebruiken. De Batselier, N. Een slagvaardig Vlaams Parlement, 2004 – 2009. 3^e Beleidsplan. Februari 2005, p. 27, <https://intranet.vlaamsparlement.be/docs/portlet/p615dp26lijn487.doc>, 12/07/06

⁸⁵ Boudrez, F. Archiveren van websites: een kwestie van waardering en 'capture'. Antwerpen: Expertisecentrum DAVID vzw, 2005, pp. 10-14

⁸⁶ Zie onder meer: Arms, W. ; Adkins, R. ; Ammen, C. & Hayes, A. Collecting and preserving the web: The Minerva Prototype. In: *RLG Diginews, Volume 5, Number 2*, 15 april 2001, <http://www.rlg.org/legacy/preserv/diginews/diginews5-2.html#feature1> , 16/02/06 en Hakala, J. Collecting and preserving the web: Developing and testing the NEDLIB Harvester. In: *RLG Diginews, Volume 5, Number 2*, 15 april 2001, <http://www.rlg.org/legacy/preserv/diginews/diginews5-2.html#feature2> , 16/02/06

- “voorafgaand aan de lancering van de nieuwe website en dus eind dit jaar, bezorgt ICT aan het Archief de bronbestanden (webpagina’s in HTML) van de oude site zoals ze op dat ogenblik op de webserver staan (unzipped)
- in geval bepaalde bronbestanden het risico lopen niet meer op termijn leesbaar te zijn (in het geval van niet gestandaardiseerde bestanden) zal een omzetting naar een archiveringsformaat onderzocht worden en noodzakelijk zijn
- de nog bestaande bronbestanden van de vroegere versies van de website worden nagekeken; verschillende van deze bestanden zijn niet leesbaar
- voor wat betreft de ‘deep-web-inhoud’ en dus de documenten die zijn beheerd en aangeleverd vanuit andere systemen (in het bijzonder de databank Proteus): deze worden afzonderlijk en op het niveau van die toepassingen gearcheveerd
- om vast te leggen hoe de oude website zich presenteerde in zijn oorspronkelijke omgeving zal eind dit jaar en met behulp van een screenrecorder een surfsessie gefilmd worden (de frequentie van de screenshots wordt als een videobestand bewaard en de verfilmde website zal bezocht kunnen worden met de webbrowser waarvoor hij is gebouwd)”

Voor de archivering van de nieuwe site geldt het volgende:

- “ICT houdt er technisch rekening mee dat in het kader van de openbaarheid van bestuur al de versies van de webinformatie die wordt aangeleverd aangaande aanbestedingen en vacatures, bewaard moeten worden en gearcheveerd gedurende de in dit kader vast te leggen juridische bewaartermijn (te bespreken met de juridische dienst); ICT lost dit op met zipfiles van de website en zal hierover communiceren
- ICT bespreekt met het archief de te implementeren metadatavelden (in de mate van het mogelijke zullen metagegevens automatisch gegenereerd worden) ...
- Archiefoverdracht van de bronbestanden zal gebeuren op ‘strategische ogenblikken’ vb. wanneer de nieuwe versie gelanceerd wordt
- Het archiveren en filmen van een surfsessie zal eveneens plaats vinden op strategische ogenblikken (bedoeling is de impressie van de *look and feel* te archiveren)
- De ‘deep-web-inhoud’ wordt gearcheveerd op het niveau van de respectievelijke toepassingen (vb. Proteus)
- In functie van een goede archivering van de bronbestanden: alleen de bestanden die in de online versie worden gebruikt, blijven op de webserver staan”⁸⁷

⁸⁷ Directie Informatie, Archief en Biografische Dienst, Archivering oude en nieuwe website: bewaarstrategie & planning. Verslag van de vergadering met ICT van 21 juni 2006

In essentie, worden dus enkel de statische HTML-pagina's bewaard (samen met een surfsessie voor de *look-and-feel*). Deze pagina's kunnen in hetzelfde formaat worden opgeslagen, aangezien het hier om een internationaal aanvaarde standaard gaat, waarvan de broncode gekend is en waarvoor men enkel een standaard webbrowsers nodig heeft. Deze webpagina's worden (op het niveau van ICT) in principe browseronafhankelijk geschreven en alleszins getest op de courante browsers gebruikt in het Vlaams Parlement, d.i. Internet Explorer en Mozilla Firefox. Aangezien Explorer eigendom is van één producent, kan beter Firefox mee gearhiveerd worden. Firefox is een *open source*, grafische webbrowsers, ontwikkeld door de Mozilla Corporation en vele vrijwilligers. Ze is geschikt voor meerdere platformen, ondersteunt open formaten en volgt de W3C standaard beter dan Explorer.⁸⁸

F. Audio

Geluidsopnames worden gemaakt van zowel de plenaire als de commissie vergaderingen die in het Parlement worden gehouden. Voor de opnames wordt gebruik gemaakt van Arbor Logger & Fetcher.⁸⁹ Arbor is een systeem integrator en applicatie software provider voor professionele audio en video.⁹⁰ Via een *web-based* architectuur worden opnames ter beschikking gesteld door het gebruik van een webbrowsers op een pc aangesloten op het interne netwerk. Opnames kunnen gemaakt worden met een lage kwaliteit (Windows Media Format), maar ook met een hoge kwaliteit (MPEG Laag 2 of 3 en lineaire PCM). De lage kwaliteitsopname wordt altijd gebruikt om via het netwerk te luisteren, om een *overload* op de bandbreedte te vermijden.⁹¹

In het Vlaams Parlement worden opnames gemaakt in het Windows Media Audio 9 (WMA9) formaat, met lossy compressie, en in WAV, met lossless compressie. WMA is een bestandsformaat voor geluidsopnames, eigendom van Microsoft, dat gebruik maakt van compressie. Opnames worden doorgaans gemaakt aan 44.1 kHz en 16 bits, in overeenkomst met de standaard voor CD's.⁹² Het WAV (of Wave) bestandsformaat werd gezamenlijk ontwikkeld door Microsoft en IBM, met als doel geluidsopnames te bewaren en af te spelen op pc's. Hoewel het mogelijk is deze bestanden te comprimeren aan de hand van de gebruikte codec, wordt dit zelden gedaan. In de meeste gevallen wordt gebruik gemaakt van het PCM formaat, wat lossless bewaring inhoudt. Het gevolg hiervan is dat WAV-bestanden nogal groot qua formaat zijn.⁹³

⁸⁸ <http://www.mozilla.com/firefox/>

⁸⁹ <http://www.arbor-audio.com/>

⁹⁰ Company profile Arbor, <http://www.arbor-audio.com/index.php?action=company>, 20/05/06

⁹¹ LogDepot Audio Specification, <http://www.arbor-audio.com/index.php?action=product&productid=2&view=specification>, 20/05/06

⁹² Windows Media Codecs, <http://www.microsoft.com/windows/windowsmedia/forpros/codecs/audio.aspx>, 20/07/06

⁹³ WAV. In: Wikipedia, The Free Encyclopedia, <http://en.wikipedia.org/wiki/Wav>, 20/07/06

De opnames (mono) van de vergaderingen worden gemaakt aan 44 kHz, met een samplerate van 32 kbps voor WMA9 en 64 kbps voor WAV op een Windows 2000 server. Voor WMA wordt een lossy compressie toegepast, voor WAV wordt gebruik gemaakt van ADPCM. *Differential Pulse-Code Modulation* (DPCM) vermindert de vereiste bits per sample met 25% door het toevoegen van een bepaalde encoding. *Adaptive PCM* (ADPCM) is een variant hierop die de vereiste bandbreedte voor een gegeven *signal-to-noise ratio* vermindert.⁹⁴ De WMA-bestanden worden op de server ongeveer 55 dagen bewaard, WAV 5 dagen. Archivering wordt manueel door Mr. Van Paemel van de archiefdienst uitgevoerd. Aangezien enkel de lage kwaliteitsopnames ter beschikking worden gesteld via het netwerk, worden WMA-bestanden geëxtraheerd.

Op dit moment omvat het archief 322 gigabyte (Gb) audio-opnames die frequent door het Algemeen Secretariaat worden geraadpleegd. 150 Gb opnames plenaire- en commissievergaderingen voor de periode 12 maart 1996 tot 23 mei 2003 staan op magneto-optische schijven (Sony, Magneto Optical Disc). Het beluisteren vormt een probleem omdat het te gebruiken machine systeem van Efrat Future Technology in België niet meer wordt ondersteund. Binnen het Parlement zijn enkel nog twee personen die er ooit mee hebben gewerkt. Een project werd door de Directie Informatica opgestart om deze banden te digitaliseren en over te zetten naar een Windows Media-omgeving. Over de huidige kwaliteit van die banden zijn geen gegevens beschikbaar. De registratiegegevens aangaande dit archief werden opgenomen in een Access-bestand en geven informatie over het soort vergadering, zittingsjaar, datum van de vergadering, begin- en einduur, nummer van de schijf, nummer audio-systeem, beschreven zijde. Sinds 29 april 2003 worden de opnames digitaal op de server geplaatst en veiligheidskopieën op CD-R gebrand. De bestanden op deze 172 Gb worden geordend volgens een mappenstructuur per vergaderweek en vervolgens op basis van de agenda van de commissies en plenaire vergaderingen onderverdeeld.

Een echt goede archiefstandaard voor audio-bestanden bestaat op dit moment niet. Van de aanwezige formaten komt Wave het meeste in de buurt. Hoewel het formaat eigendom is van Microsoft en IBM, werd bijkomende informatie vrijgegeven over de structuur⁹⁵ en data specificaties⁹⁶ voor multimedia gebruikt in een Windows-omgeving.⁹⁷ Wave kan beschouwd worden als een *de facto* standaard, omdat het merendeel van de computers op de markt de Microsoft OS ondersteunen. Hiernaast

⁹⁴ Pulse-code modulation. In : *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/ADPCM>

⁹⁵ Waveformatex. http://msdn.microsoft.com/library/default.asp?url=/library/en-us/multimed/htm/_win32_waveformatex_str.asp, 26/07/06

⁹⁶ IBM Corporation & Microsoft Corporation, Multimedia Programming Interface and Data Specifications 1.0, August 1991, <http://www.kk.ij4u.or.jp/~kondo/wave/mpidata.txt> of http://www.tactilemedia.com/info/MCI_Control_Info.html, 26/07/06

⁹⁷ Zie ook: The National Digital Information Infrastructure and Preservation Program, WAVE Audio File Format, Library of Congress, June 2005, <http://www.digitalpreservation.gov/formats/fdd/fdd000001.shtml>, 26/07/06

kan het formaat op de meeste audio-apparaten en software afgespeeld worden. Als nadeel wordt aangehaald dat dit formaat veel plaats in beslag neemt. De prijs van computergeheugen blijft echter verminderen, terwijl de capaciteit enkel toeneemt (we zijn op 50 jaar tijd van kilobytes naar petabytes gegaan⁹⁸). Het blijft te bezien natuurlijk hoeveel geld vrijgemaakt kan worden, maar omwille van “de torenhoge” kosten hoeft het niet gelaten te worden. Een andere optie is de toepassing van de *Moving Pictures Expert Group* (MPEG) standaard. MPEG is een ISO werkgroep, opgericht in 1988 met als opdracht compressiestandaarden voor video, audio en een combinatie van beide te creëren. Het resultaat is een familie van open standaarden: MPEG-1, 2, 4, 7 en 21. Hier kom ik later op terug.⁹⁹ Voor MPEG-1 Audio bestaan drie compressieschema's elk met een performantie (Layer 1 tot 3). MPEG-1 Audio Layer 3, kortweg MP3, levert de hoogste kwaliteit en compressieratio. Het is ook de meest complexe variant. Achteraan een MP3-bestand kan een ID3-tag toegevoegd worden. Hierin kunnen gegevens over titel, artiest, album, genre, jaar en commentaar toegevoegd worden op het moment van encoderen of achteraf met een ID3-editor. De tag bevat metadata in de vorm van ASCII-karakters. MPEG 2-Audio is een uitbreiding op zijn voorganger. Tijdens de testfase werd vastgesteld dat grotere compressieratio's konden bereikt worden door het gebruik van andere algoritmes. Na onderzoek, werd in 1997 de MPEG-2 Advanced Audio Coding (AAC)¹⁰⁰ standaard uitgebracht, bedoeld als opvolger van MP3 omwille van de hogere compressieratio met minder kwaliteitsverlies.¹⁰¹

G. Video

Elke woensdag vindt de plenaire vergadering plaats. In een *inhouse* regiekamer worden hiervan audiovisuele opnames gemaakt (uitbesteed aan Videohouse¹⁰²), wat een jaarlijkse aanwinst van ongeveer 720 uur beeldmateriaal oplevert. Dit gebeurt aan de hand van de *capture* van een digitale *feed* via SDI¹⁰³ van beeldregie (PAL: de Europese standaard voor kleurentelevisie-uitzendingen¹⁰⁴). Hiervoor wordt gebruik gemaakt van Videum, een audiovisuele capture card van het bedrijf Videum¹⁰⁵, die beelden opslaat in IYUV Uncompressed YUV 4:2. YUV is de kleurenruimte

⁹⁸ Lawrence Livermore National Laboratory [From kilobytes to petabytes in 50 years](#).

University of California, 2003, <http://www.llnl.gov/50science/computation.html>, 26/07/06

⁹⁹ Boudrez, F. [Standaarden voor digitale archiefdocumenten](#). Technisch rapport. Antwerpen: Stadsarchief, 2001–2003, p. 28

¹⁰⁰ ISO/IEC 13818-7:1997, Information technology -- Generic coding of moving pictures and associated audio information -- Part 7: Advanced Audio Coding (AAC)

¹⁰¹ Boudrez, F. *Ibidem*, p. 34

¹⁰² Videohouse biedt geïntegreerde audiovisuele oplossingen aan voor bedrijven, productiehuisen en overheden. <http://www.videohouse.be/>

¹⁰³ Serial Digital Interface (SDI) is een standaard ontwikkeld door the Society of Motion Picture and Television Engineers (SMPTE), die beschrijft hoe ongecomprimeerde, digitale video data moet getransporteerd worden via videokabels. Voor meer info, zie: <http://www.smpte.org/>

¹⁰⁴ Phase Alternating Line (PAL). [PAL](#). In: *Webopedia*, <http://www.webopedia.com/TERM/P/PAL.html>, 09/04/06

¹⁰⁵ <http://www.winnov.com/videum.htm>

die wordt gebruikt door PAL¹⁰⁶ (voor een beschrijving van kleuren schema's, zie Deel II, hoofdstuk 2).

Deze beelden worden gratis ter beschikking gesteld aan interne medewerkers (sporadische opvraging) en de media (berichtgeving in het nieuws). De bewaring van ouder materiaal gebeurt op digitale Betacam SX-tapes (enkel toegankelijk voor archiefmedewerkers), nieuw materiaal lokaal op harde schijf in WMV9-formaat (enkel toegankelijk voor ICT). De gearchiveerde tapes kunnen op dit moment niet door de archiefdienst worden afgespeeld en zijn bijgevolg niet ontsloten. Betacam SX werd in 1996 geïntroduceerd door Sony als digitale opvolger van de analoge SP tape. Videobeelden worden opgeslagen in MPEG 4:2:2, samen met 48kHz 16 bit PCM audio. De S tapes hebben een opslagcapaciteit van 62 minuten, L tapes tot 194 minuten. Hoewel het formaat nog wordt gebruikt door sommige producties, werden de afspeelmachines uit de handel genomen.¹⁰⁷ Windows Media Video (WMV) is een video codec die werd ontwikkeld door Microsoft. De technologie werd aanvankelijk ontwikkeld als een producentgebonden, *low-bitrate streaming* applicatie. In 2003 echter werd een specificatie van versie 9 voorgelegd aan the Society of Motion Picture and Television Engineers (SMPTE) voor standaardisatie.¹⁰⁸ In 2006 werd de standaard officieel erkend als SMPTE-421M.¹⁰⁹

Optimaal worden de beelden opgeslagen in een niet gecomprimeerd formaat. Aangezien video-opnames nogal veel plaats innemen en het bijgevolg vrij duur is (het Vlaams Parlement wil werken “op basis van de principes van effectiviteit, efficiëntie en zuinigheid”¹¹⁰), zal dit niet mogelijk zijn. Als archiveringsstandaard zijn dan een aantal mogelijkheden. Een eerste mogelijkheid ligt in de aannahme van MPEG-Video. Zoals al vermeld werd, behoort de MPEG-familie tot de groep van ISO-standaarden. MPEG-4¹¹¹ dient voor de bundeling van multimediabestanden, interactieve afbeeldingen en digitale TV binnen netwerktoepassingen. Deze standaard werkt voornamelijk met lage *samplerates* en *data-encoding* en splitst de verschillende onderdelen van een multimediatoepassing in afzonderlijke objecten op. Als aanvulling werd MPEG-7 ontwikkeld.¹¹² Deze ondersteunt MPEG-4 door een specifieke en gestandaardiseerde tool aan te bieden voor de beschrijving van metadata als gestructureerde informatie.¹¹³ Als

¹⁰⁶ Voor een meer technische beschrijving, zie: Poynton, C. *Video engineering*. January 2003, <http://www.poynton.com/Poynton-video-eng.html>, 09/04/06

¹⁰⁷ *Betacam*. In: *Answers.com*, <http://www.answers.com/topic/digital-betacam>, 09/04/06

¹⁰⁸ *Windows Media Video*. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Windows_Media_Video, 09/05/06

¹⁰⁹ Girod, C., *SMPTE releases VC-1 standard*. SMPTE Press Release, April 3, 2006, http://www.smppte.org/news/press%5Freleases/003_06.cfm, 09/05/06

¹¹⁰ Missieverklaring van het Vlaams Parlement, <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/HoeWerktHetVlaamsParlement/Missie/MissieVerklaring.html>, 14/07/07

¹¹¹ MPEG-4: ISO/IEC 14496:1998, Information technology -- Coding of audio-visual objects

¹¹² ISO/IEC TR 15938-8:2002, Information technology -- Multimedia content description interface -- Part 8: Extraction and use of MPEG-7 descriptions

¹¹³ Boudrez, F. *Standaarden voor digitale archiefdocumenten*. Technisch rapport. Antwerpen: Stadsarchief, 2001 –2003, p. 29

verdere aanvulling werd MPEG-21¹¹⁴ gepubliceerd. Deze standaard “legt vast hoe computers en software met de verschillende soorten informatie moeten omgaan, om ze op een zinvolle manier samen te gebruiken. Het laat bijvoorbeeld toe om verschillende soorten digitale data synchroon te verwerken.”¹¹⁵ De visie voor MPEG-21 is “to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities.”¹¹⁶

In principe is alles wat door Microsoft wordt ontwikkeld niet aan te raden voor archivering, omwille van de gebondenheid aan één producent. Door de aanneming van SMPTE van WMV9 als internationale standaard, is dit formaat echter perfect bruikbaar als oplossing voor het Vlaams Parlement. Bovendien is het argument dat Microsoft aanhaalde om dit product te promoten ook hier valabel: WMV is efficiënter dan MPEG-4 voor het encoderen van video, waardoor meer informatie op één drager kan bewaard worden.¹¹⁷

3.3 Draggers

Om digitale archiefdocumenten te kunnen bewaren moeten ze uiteraard ergens op worden opgeslagen. Het gebruik van duurzame dragers is aan te raden. Ze moeten in staat zijn gegevens op lange termijn te bevatten en mogen niet snel in fysiek verval geraken. De levensduur wordt doorgaans onderzocht door bij tests het verouderingsproces te versnellen terwijl het aantal fouten wordt gemeten. Vervolgens worden dan voorspellingen gemaakt over de duur van het verouderingsproces wanneer de drager in goede materiële staat wordt bewaard. Hierbij wordt rekening gehouden met het opsporings- en foutencorrectiemechanisme van het opslagmedium. Elk medium heeft zijn eigen systeem, dat tot op zekere hoogte fouten kan verbeteren. Eens de bovengrens hiervan wordt overschreden, worden de opgeslagen bestanden onleesbaar. Regelmatige kwaliteitscontroles zijn dus een *must*.¹¹⁸

In het Vlaams Parlement wordt gebruik gemaakt van zowel magnetische (zoals de Betacam-tapes) als optische dragers (zoals CD-R). Voor magnetisch dragers worden diskettes en (lokale) harde schijven als niet duurzaam aanzien. Na verloop van tijd neemt het magnetisme af, waardoor uiteindelijk

¹¹⁴ ISO/IEC TR 21000-1:2004, Information technology -- Multimedia framework (MPEG-21) -- Part 1: Vision, Technologies and Strategy

¹¹⁵ Stroeykens, S. *MPEG-21 schept orde in de multimedia-chaos*. In: *De Standaard*, 21 oktober 2002, p. 17

¹¹⁶ Timmerer, C. *Information technology – MPEG-21 multimedia framework*. Department of Information Technology, University of Klagenfurt, January 2005, <http://mpeg-21.itec.uni-klu.ac.at/cocoon/mpeg21/>, 09/05/06

¹¹⁷ *WMV*. In: *Audio Video Affair*, <http://www.audio-video-affair.com/wmv.html>, 09/05/06

¹¹⁸ Boudrez, F. & Dekeyser, H. *Digitaal archiefbeheer in de praktijk*. Antwerpen: Stadsarchief, 2004, p. 65-66

de bewaarde gegevens gewoon verdwijnen. Bovendien is deze technologie in onbruik aan het geraken. Voor vele media in deze categorie is al geen afspeelapparatuur op de markt.¹¹⁹ Qua optische media bestaan ook geen garanties op de lange termijn. Op basis van eigen onderzoek heeft het VRT-archief vastgesteld dat lineaire audio-CD's al na vijf jaar onherroepelijke fouten beginnen te vertonen. Voor data-CD's waarbij compressie werd toegepast, was dit al het geval na twee jaar. De kwaliteit van beschrijfbaar CD's bleek sterk af te hangen van de plaats van productie. Zelfs bij gerenommeerde merken moet worden nagekeken worden in welke specifiek fabriek de schijven werden geproduceerd.¹²⁰ Iedereen kijkt nu natuurlijk naar de evolutie van DVD's. In de voorgaande zin is het woord 'evolutie' van belang. In tegenstelling tot de andere opslagmedia zijn nog geen vaste standaarden voor handen. Heel de technologie staat nog niet op punt, bovendien geregeld worden nieuwe soorten op de markt gegoid.¹²¹

De veiligste manier van bewaren is op één of meerdere servers. Servers hebben het voordeel dat ze van kortbij worden gevolgd door ICT (en eventueel de archivaris). Fouten worden op die manier sneller opgemerkt en gecorrigeerd. De leesbaarheid van de fysieke dragers blijft gegarandeerd, omdat het niet waarschijnlijk is dat alle verschillende onderdelen van een server (al dan niet binnen een netwerk) tegelijkertijd fouten zullen vertonen. Zelfs indien het systeem crasht kan worden voorzien in een backup-systeem, waarmee dat probleem binnen afzienbare tijd kan worden verholpen. Een blijvende inspanning wordt wel verwacht van de Dienst Informatica. Tegenkantingen van die kant lijken mij echter niet waarschijnlijk eenmaal het digitaal depot op het vlak van technische ondersteuning bij hun takenpakket wordt gevoegd (al was het maar omdat die opdracht geformaliseerd zal worden op de hogere echelons).

¹¹⁹ Boudrez, F. Magnetische dragers voor het archief. Antwerpen: Stadsarchief, 2002

¹²⁰ Steyaert, C. Rondleiding in de beeld- en geluidsarchieven van de VRT. Onuitgegeven lezing, 20 maart 2006; Volgens hun onderzoek worden de CD's van de hoogste kwaliteit geproduceerd door het bedrijf Taiyo Yuden. <http://www.t-yuden.com/recordablemedia/index.cfm>

¹²¹ Zie onder meer: Next-generation DVD battle begins. In: *BBC News UK*, <http://news.bbc.co.uk/1/hi/technology/4921784.stm>, 19 April 2006, 09/05/06

3.4 Rol van de digitale handtekening

Documenten worden ondertekend met een handtekening om verschillende redenen¹²²:

- *bewijskracht geven*: wanneer de handtekening een duidelijke distinctieve aard heeft (uitgesproken en expressief is) kan ze aan de ondertekenaar worden toegeschreven
- *ceremoniële functie*: de handtekening zorgt voor een plechtig moment, wat helpt om slecht overwogen engagementen tegen te gaan
- *goedkeuring*: de ondertekenaar erkent de inhoud en dus de wettelijke uitwerking daarvan
- *(logistieke) efficiëntie*: de handtekening zorgt voor transactionele klaarheid, door geldig te blijven ook bij afwezigheid van de ondertekenaar

Met een elektronische handtekening wordt hetzelfde beoogd. Niet te verwarren met een digitale handtekening. Een elektronische handtekening is de computervariant van een ‘analoge’ handtekening in de zin van “an electronic sound, symbol, or process, attached to or logically associated with a contract or other record and executed or adopted by a person with the intent to sign the record.”¹²³ Het omvat alle identificatie mechanismen die een elektronische methode gebruiken om partijen te binden (bvb. de pincode van een bankkaart, paswoorden, het toevoegen van de eigen naam onderaan een e-mail). Eigenlijk is de digitale handtekening één van de mogelijke vormen van een elektronische handtekening.

Een digitale handtekening is een methode voor het bevestigen van de juistheid van digitale informatie aan de hand van technieken van cryptografie, die de ontvanger toelaat de identiteit en de autoriteit van de ondertekenaar/ zender en/of de integriteit van het document na te gaan. Op internationaal en lokaal niveau bestaan ondertussen wetgevingen die digitale handtekeningen even bindend en universeel willen maken als hun handgeschreven varianten.¹²⁴

¹²² Scheelings, F. Statisch en dynamisch archief van administraties en bedrijven. Brussel: VUB, Onuitgegeven hoorcollege, 2005

¹²³ Electronic signature. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Electronic_signature, 29/03/06

¹²⁴ Wat? In: *DigitaleHandtekening.be*, <http://www.digitalehandtekening.be/what.cfm>, 27/07/06; Wetgeving op Europees niveau: Directive 1999/93/EC of the European Parliament and of the council of 13 December 1999 on a Community framework for electronic signatures. In: *Official Journal of the European Communities*, 19 January 2001; België: Wet tot invoering van het gebruik van telecommunicatiemiddelen en van de elektronische handtekening in de gerechtelijke en buitengerechtelijke procedure, 20 oktober 2000 (BS 22 december 2000); Verenigde Staten: Public Law 106-229 – June 30, 2000, Electronic signatures in global and national commerce Act

Voor we ons buigen over de verschillende manieren om een digitaal object te versleutelen, beginnen we best bij het begin. Een digitale handtekening wordt gemaakt aan de hand van cryptografie. Het is de studie van technieken en toepassingen die afhangen van het bestaan van een moeilijk probleem. Of in andere woorden, het maken of begrijpen van “verborgen woorden” (of is dat iets te cryptisch?). Het Van Dale woordenboek stelt simpelweg “de kunst van geheimschrift te schrijven.”¹²⁵ De Merriam-Webster Online Dictionary voegt hier nog “the enciphering and deciphering of messages in secret code or cipher; *also* : the computerized encoding and decoding of information” aan toe.¹²⁶ Encryptie is de transformatie van data in een vorm waar het zo onmogelijk mogelijk is gemaakt om nog te kunnen lezen zonder specifieke kennis. Decryptie is de tegenstelling hiervan.¹²⁷ Het encoderingsproces kan oa. op basis van *hashing*. Met een hashingcode wordt aan elke spatie, cijfer en andere lettertekens een getal toegekend. Hiermee wordt voor een bepaald document een unieke berekening gemaakt, d.i. de hashingcode. Door de uitkomst van die berekening te controleren wordt nagegaan of het document onveranderd is gebleven.¹²⁸

De effectieve toepassing kan op verschillende manieren¹²⁹:

- *symmetrische encryptie*: één sleutel wordt ter beschikking gesteld aan zowel de gebruiker als de archivaris. Bij het ontsleutelen gebruiken ze beide dezelfde sleutel voor het openen van het bestand.
- *asymmetrische encryptie*: één sleutel wordt ter beschikking gesteld aan de gebruiker, een andere aan de archivaris. Ook hier zijn beide sleutels nodig voor het openen van het bestand.
- *Public Key Infrastructure (PKI)*: er wordt gewerkt met een openbare sleutel en een geheime sleutel. Een hashingcode wordt berekend met de geheime sleutel en met het document. Het resultaat is een digitale handtekening die aan het document wordt toegevoegd. Om de integriteit en authenticiteit na te gaan wordt dan een berekening uitgevoerd met het document, de handtekening en de openbare sleutel. Indien het resultaat klopt volgens een vooraf opgestelde wiskundige relatie is de handtekening echt. Indien niet is de handtekening frauduleus of werd het document veranderd. Er kan ook gewerkt worden met drie sleutels. Een openbare sleutel op de

¹²⁵ *Cryptografie*. In: G. Geerts. & H. Heestermans (eds.), *van Dale. Groot woordenboek der Nederlandse taal. a-i*, Utrecht-Antwerpen: Van Dale Lexicografie bv, 1992 (Twaalfde uitgave), p. 587, kol. 2

¹²⁶ *Cryptography*. In: *Merriam-Webster Online Dictionary*, <http://www.m-w.com/dictionary/cryptography>

¹²⁷ RSA Laboratories, *RSA Laboratories' Frequently Asked Questions about today's cryptography, Version 4.1*. RSA Security Inc., 2000, <http://www.rsasecurity.com/rsalabs/node.asp?id=2157>, 29/03/06

¹²⁸ RSA Laboratories, *Ibidem*, <http://www.rsasecurity.com/rsalabs/node.asp?id=2176>, 29/03/06

¹²⁹ Scheelings, F. *Statisch en dynamisch archief van administraties en bedrijven*. Brussel: VUB, Onuitgegeven hoorcollege, 2005; voor een technisch overzicht, zie: Menezes, A.; van Oorschot, P. & Vanstone, S. *Handbook of applied cryptography*. Boca Raton: CRC Press, 1996

certificatie (hiervoor is samenwerking met een certificatie-autoriteit nodig¹³⁰), een openbare sleutel van de instelling op de hashingcode en een priv sleutel van de archivaris op een onveranderlijke hashingcode. De gebruiker kan de twee hashinggetallen vergelijken, terwijl het archiefsysteem de code berekent.

Hoewel het gebruik van een digitale handtekening een zeer bruikbaar middel is om op korte termijn documenten te valideren, is het van archiefwetenschappelijk standpunt niet direct aangewezen om hiermee de authenticiteit van de bewaarde bestanden trachten aan te tonen. "De authenticatieverklaring reikt niet verder dan het bewijs dat iets/iemand in het bezit van de private sleutel van de ondertekenaar het document ondertekende. Met betrekking tot de integriteit bewijst een digitale handtekening enkel dat de bits van het verstuurde document intact zijn. Met digitale handtekeningen worden evenmin onrechtmatige wijzigingen of manipulaties voorkomen."¹³¹ Om authentieke en integere archiefdocumenten te archiveren is het belangrijker om essenti le identificerende metadata te registreren, archiefdocumenten te beschrijven en maatregelen te ondernemen om de betrouwbaarheid van de opgenomen stukken te garanderen.¹³²

Omdat digitale handtekeningen vooral geschikt zijn voor validatie op korte termijn (vlak na ontvangst), brengt het gebruik hiervan weer andere problemen met zich mee op lange termijn. De software die gebruikt wordt om digitale handtekeningen te maken en te controleren zijn even tijdsgebonden als die van de archiefdocumenten zelf. De gebruikte methodes zullen na verloop van tijd verouderen of uit de mode geraken. Het wordt dan noodzakelijk om ze te migreren. Door het wijzigen van de bitstream wordt de digitale handtekening vervolgens onbruikbaar als authenticatiemethode. Een mogelijke oplossing zou zijn bij migratie het document opnieuw te laten ondertekenen. Dit kan echter problemen opleveren indien de oorspronkelijke ondertekenaar dit weigert of overleden is. In plaats van te hertekenen zou men ook het migratieproces zelf kunnen laten certifi ren. Binnen een PKI-structuur moet bovendien het toegekende certificaat haar bewijskracht ontlenen aan een externe organisatie. Daardoor wordt een bepaalde afhankelijkheid van een derde partij gecre erd. Een aantal initiatieven (oa. door de European Electronic Signature Standardisation Initiative, EESSI en Victorian Electronic Records Strategy, VERS) werden ondernomen om de mogelijkheid de *validation chain* mee te archiveren te onderzoeken. Hierbij worden ook alle elementen van de PKI nodig voor verificatie op lange termijn bewaard.¹³³

¹³⁰ Bvb. <https://www.pkioverheid.nl/> voor Nederland, voor een overzicht, zie: <http://www.pki-page.org/>

¹³¹ Boudrez, F. *Digitale handtekeningen en archiefdocumenten*. Antwerpen: Expertisecentrum DAVID vzw, 2005, p. 6

¹³² Boudrez, F. *Idem*, p. 6

¹³³ Boudrez, F. *Digitale handtekeningen en archiefdocumenten*. Antwerpen: Expertisecentrum DAVID vzw, 2005, pp. 7-13

Een andere mogelijkheid bestaat erin direct na validatie de gegevens hierover in de metadata te registeren. Vervolgens zou dan de oorspronkelijke handtekening kunnen verwijderd worden, aangezien ze toch overbodig is geworden. Hiertegen zijn wel een aantal juridische (verplichte bewaring van de oorspronkelijke handtekening) en archiefwetenschappelijke (de handtekening als extrinsiek element van de documentaire vorm) bezwaren.¹³⁴

Het gebruik van een digitale handtekening zorgt alleszins voor een bijkomende moeilijkheid voor de digitale archivering. Welke de beste methode is moet nog uitgewezen worden.

De digitale handtekening heeft voor het stellen van parlementaire vragen en antwoorden ingang gevonden op het Vlaams Parlement. Het gebruik ervan wordt vooralsnog sporadisch toegepast.

¹³⁴ Boudrez, F. *Idem*, pp. 10-11

Besluit

“Knowledge comes from crafted bindings and pages, ..., not ones and zeros.”¹

De wereld waarin archiefstukken worden aangemaakt, is de laatste 15 jaar (en bij uitbreiding 50 jaar) sterk veranderd. Voor de verschillende soorten van bescheiden waren de mechanische eigenschappen in meer of mindere mate gekend. Voor zowel het fysieke als het intellectuele beheer waren een aantal methodes voor handen die hun nut al hadden bewezen en waarop werd verder geborduurd. Door de toenemende informatisering zijn een aantal van de aanvaarde concepten op de helling komen te staan. Voor het intellectuele beheer is het niet meer voldoende om pas in te grijpen vanaf de statische fase van een document. Fysiek is het niet voldoende om enkel de drager in een goede materiële en geordende staat te bewaren. De archiefwereld heeft ondertussen niet stilgezeten gelukkig. Bestaande concepten werden verder verfijnd en aangepast (cfr. het *records continuum* principe, het OAI-model). Technische oplossingen werden gezocht en, hoewel ze nog verre van perfect zijn, gevonden. Hoewel onze manier van leven en werken een grote invloed ondervindt van deze vierde informatierevolutie², is de opdracht van de archivaris in essentie niet veranderd: opnemen, bewaren, ordenen en klasseren, ontsluiten, ter beschikking stellen en valoriseren. Enkel de middelen zijn veranderd.³

Het grote verschil vinden we terug in het feit dat we nu met data moeten werken die eigenlijk niet bestaan in de fysieke zin. Het virtuele karakter van digitale archiefdocumenten brengt met zich mee dat de opgeslagen gegevens niet met het blote oog kunnen waargenomen worden. Bepaalde hard- en software zijn vereist om een gegeven bestand te kunnen openen en een menselijk leesbare presentatie te genereren. Hier gaan een aantal specifieke problemen mee gepaard. Beide zijn onderhevig aan veroudering. Door de snelle evolutie in technologie, is het nagenoeg onmogelijk om archiefdocumenten in de oorspronkelijke vorm te bewaren. Authenticiteitseisen, zoals ze worden gesteld aan papieren stukken, zijn niet meer volledig geldig in een digitale omgeving. De idee van ‘origineel’ verliest op een bepaald niveau zijn betekenis, omdat digitale kopieën volstrekt identiek zijn. Aan de andere kant kan een archiefdocument op verschillende manieren worden voorgesteld. Hoewel de presentatie en de bitstream verschillen van het oorspronkelijke stuk, kunnen de inhoudelijke gegevens ongewijzigd zijn. Op het vlak van context, moet buiten de informatie betreffende de plaats van het stuk in het groter geheel (verhouding tot andere stukken, plaats in een dossier of reeks, relatie met de archiefvormer,

¹ A. S. Head in *Buffy the Vampire Slayer*, Season Seven, Episode 17

² De voorgaande informatierevoluties: uitvinding van de taal, het schrift, de boekdrukkunst

³ En de nood aan selectie door de enorme hoeveelheid informatie die sinds de Tweede Wereldoorlog geproduceerd wordt en alleen maar exponentieel toeneemt.

...), ook technische gegevens worden opgenomen om grotere garanties voor de leesbaarheid naar de toekomst te bieden. Eén van de grootste problemen die gepaard gaat met de technische ontwikkeling, ligt besloten in één van de grootste voordelen: evolutie. Een snelle evolutie, waarbij informatici en andere wetenschappers continue naar nieuwe mogelijkheden zoeken, zorgt ervoor dat onze maatschappij aan een razend tempo vooruitgang boekt op het vlak van de geneeskunde, organisatie, ruimtevaart, communicatie, ... Door deze snelle vooruitgang is er echter geen tijd om terug te kijken. De aangereikte middelen moeten nu werken. Zijn er hier en daar bepaalde foutjes in de software, dan is dat niet zo erg want volgend jaar staat de verbeterde opvolger al klaar. Ondertussen wordt wel gewerkt met technologie die sneller wordt vergeten dan de hits van Britney Spears die archiefdocumenten opleveren. Om toch te kunnen zorgen voor een rustig oog in de storm, hebben verschillende organisaties en internationale organen het op zich genomen te zorgen voor een aantal standaarden en *best practices* op het vlak van archiefbeschrijvingen, bewaarstrategieën en technische formaten (voor applicaties, dragers, archiefbeheerssystemen, transferprotocols, ...).

Tijdens mijn stage aan het Vlaams Parlement werd in eerste instantie van mij verwacht de literatuur betreffende deze problematiek na te gaan en op een rijtje te zetten. In een tweede fase werd de concrete technische omgeving waarbinnen wordt gewerkt in kaart te brengen. De bedoeling was hierop aansluitend een aantal concrete aanbevelingen te formuleren, die toegepast zouden kunnen worden door het te vormen digitaal depot.

Op het vlak van de mogelijke conserveringsstrategieën die op dit moment voor handen zijn, kan best de algemene trend gevolgd worden. Niet omdat het de algemene trend is, maar omdat het op dit moment de beste garanties biedt voor een duurzame bewaring. Migratie houdt in dat de te bewaren stukken worden overgezet van het oorspronkelijke formaat naar een geschikt archiveringsformaat. De meest gebruikte IT toepassingen zijn door de band genomen eigendom van een bepaalde producent. Voor het gebruik van de software worden een aantal licenties toegekend. Dit brengt met zich mee dat het archief afhankelijk is van een externe speler wat betreft de rechten van de gebruikte techniek. Door te migreren naar een open standaard met voldoende marktpenetratie, blijft de ondersteuning langer gegarandeerd. Zelfs bij het wegvallen van een externe ondersteuning, kan aan de hand van publiekelijk verkrijgbare broncodes naar bepaalde oplossingen gezocht worden op het moment dat het nodig is. Hoewel emulatie een veelbelovende techniek is, is het een technisch complexe zaak, waarvan de lange-termijn-garanties nog niet zijn bewezen. Lange termijn in de virtuele wereld is een relatief begrip natuurlijk, maar door met open, gedocumenteerde formaten te werken, worden de mogelijkheden open gehouden. Om die mogelijkheden zo groot mogelijk te maken, valt het aan te bevelen om de gemigreerde samen met de oorspronkelijke bitstream te bewaren. Op die manier kan (zolang de oorspronkelijk software nog voor handen is) bijkomende migraties worden uitgevoerd aan de hand van de originele samenstelling. Hierdoor vermindert de kans op cumulatieve fouten door de omzetting. Een andere optie die zo

blijft bestaan, is emulatie. Indien in de toekomst toch wordt besloten de originele omgeving te reconstrueren, dan kan dat. Een interessante piste die hiermee verband houdt is Cedars' *migration on request*. Het digitale archiefdocument wordt volgens die technologie pas gemigreerd naar een geschikt voorstellingsformaat (DIP) op het moment van aanvraag. Dit zou mogelijk moeten zijn vanaf de originele of de gemigreerde bitstream. Op die manier wordt vermeden dat op periodieke wijze bepaalde klassen van documenten opnieuw moeten gemigreerd worden, omdat ook dat formaat te sterk is verouderd en de bewaring in het gedrang komt. Enkel de tool die het feitelijke migratieproces uitvoert moet op geregelde tijdstippen aan een *update* worden onderworpen. Voorlopig lijkt een klassieke migratie naar een open, gestandaardiseerd formaat de aangewezen optie.

Wat ons bij het gebruik van standaarden brengt. De ontwikkeling van een standaard is een langzaam proces, dat niet noodzakelijk de marktevolutie bijhoudt. Maar het volgen van de allernieuwste trends en gadgets is niet de bedoeling. Een standaard moet vooral zo volledig mogelijk, kwaliteitsvol en geschikt voor de archiefwetenschap zijn.

Een eerste model om aan degelijk archiefbeheer te doen, is het *Open Archival Information System* (OAIS). Dit model biedt een aantal manieren om het proces van opname, bewaring en ter beschikking stellen te organiseren. Het is een soort van *best practice* handleiding waarop het archiefbeheer kan gebaseerd worden. De aangereikte concepten vormen eveneens de basis voor een ander model, dat van *trusted digital repositories*. Een betrouwbaar digitaal depot ontwikkelen komt de procedurele en administratieve verantwoordelijkheid van de archivaris en de instelling ten goede en zal het valoriseren de opgenomen archieven vergemakkelijken. Het uiteindelijke doel van een archief is het vergroten van de effectiviteit van een organisatie en een middel voor maatschappelijk en historisch onderzoek aan te bieden. Anders gesteld, "indien de archivaris niet dient, dient hij/zij tot niets."⁴

Om de efficiëntie van het archief- en documentbeheer nog te vergroten, kan gebruikt gemaakt worden van een *electronic document and records management system* (EDRMS). Voor de ontwikkeling ervan of als kwaliteitscontrole zijn eveneens aanknopingspunten. Belangrijke concepten hieromtrent worden aangereikt door de internationale standaard ISO 15489. Ondertussen zijn de ideeën die hierin zijn opgenomen een basis geworden om naar terug te grijpen. Ter aanvulling voor het uit te voeren onderzoek naar de organisatorische vereisten kunnen we ons beroepen op het Australische DIRKS-model. Voor de technische specificaties kan verwezen worden naar de Europese standaard MoReq (en de Nederlandse variant ReMANO).

Voor de beschrijvingen gaat een voorkeur uit naar de internationaal goedgekeurde ISAD-norm. De archivaris van het Vlaams Parlement acht het echter wenselijk om één metadatamodel te hanteren als basis voor de

⁴ Janssens, G. & Put, E. Geschiedenis, principes en terminologie van de archivistiek. Syllabus. Brussel: VUB-KUL-UG-UA, 2005-2006, p. 107

verschillende soorten van archiefbescheiden, zowel de papieren als de digitale variant. Dit lijkt logisch, omdat zo een manier van zoeken ter beschikking kan worden gesteld die gelijk is voor het hele bestand. Onderzoek zal gedaan worden naar de haalbaarheid van deze piste. Aangezien een online archiefloket zal worden ontwikkeld, dat kan worden geraadpleegd via de website, kan op voorhand aangegeven welke soort van resultaten bepaalde zoekacties zullen opleveren. Naar de klanten van het archief toe kan een gebruiksvriendelijke oplossing worden aangeboden. Een metadata standaard als Dublin Core heeft natuurlijk ook zijn voordelen, maar het gebruik van één manier van werken vermijdt een afweging op basis van twee maten en gewichten.

Op applicatieniveau lijkt het mij aangeraden om een vaste set van formaten te hanteren. Ten laatste voor opname in het digitaal depot kunnen de eigendoms- en platformgebonden formaten worden gemigreerd naar een geschikte archiveringsstandaard. De voorkeur gaat uit naar open, gedocumenteerde formaten, en in de mate van het mogelijke naar internationaal aanvaarde standaarden. De besproken applicaties behandelen deze die gebruikt worden bij de dagelijkse werking van het Vlaams Parlement: kantoordocumenten (tekstverwerkingsbestanden, spreadsheets, presentaties), e-mail, databanken, websites, audio en video. Dit zijn mijn bevindingen:

- *kantoordocumenten*: afhankelijk van de visie op de essentiële elementen kunnen verschillende formaten aangeraden worden. PDF/A is als internationale zelfbeschrijvende, standaard zeer geschikt voor tekstbestanden en presentaties. Indien de wiskundige formules niet van belang zijn, kunnen spreadsheets ook naar PDF geëxporteerd worden. Een andere optie vinden we in het OpenDocument Format van de *open source community* OpenOffice.org. In dit formaat worden bestanden bewaard op een XML-gebaseerde manier, waarvan de broncode open en gedocumenteerd is. Bij het openen worden ze zelfs aangeboden in een visuele presentatie die nauw aanleunt bij de gangbare office pakketten. Het is wel nodig voor een DIP te zorgen die enkel kan gelezen worden.
- *e-mail*: XML wordt naar voren geschoven als beste bewaarstrategie voor e-mail berichten. Het is hier wel van belang dat samen met het bericht de relevante headergegevens worden bewaard. De band met de bijlagen (die in een aangepast formaat worden bewaard) mag niet verloren gaan.
- *databanken*: ook hier wordt XML aangegeven. Een XML DTD of Schema moet worden toegevoegd om de relaties tussen de tabellen, en de inhoud en de structuur van de tabellen en views te bewaren. Informatie over functionaliteiten, de gebruikte queries, de gebruikersinterface, de ondernomen bewaaractiviteiten, en eventueel de bewaar- en toegangseisen mogen niet vergeten worden
- *website(s)*: in onderling overleg werd al vastgelegd dat de statische HTML-pagina's zullen bewaard worden, samen met een video van een surfessie om een impressie van de *look-and-feel* te behouden.

Belangrijk hier is het behoud van de interne links tussen de pagina's op basis van relatieve pathaanduidingen. Aanvullend zou documentatie moeten opgemaakt worden, waarin de opbouw van de site wordt beschreven. De archiefdocumenten op het niveau van de *deep web* toepassingen worden afzonderlijk gearchiveerd op het niveau van de documentbeheerssystemen die ze genereren.

- *Audio*: nu wordt gewerkt met bestanden op basis van het WAV en WMA formaat. Enkel de WMA bestanden worden opgeslagen door de archiefdienst. Het gebruikte systeem echter laat toe om MPEG bestanden aan te maken, wat een internationale standaard is. Dus een eenvoudige omschakeling zou al voldoende moeten zijn om een degelijk archiveringsformaat te bekomen. Indien mogelijk zouden de WAV-bestanden ook moeten bewaard worden, omwille van het feit dat werken zonder compressie de voorkeur geniet. Op die manier wordt bovendien het risico gespreid. De documenten in WAV kunnen als AIP bewaard worden, terwijl een versie in MP3 bvb. als DIP kan worden aangeboden.
- *Video*: de opname van digitale video wordt nu opgeslagen in WMV9. Sinds kort werd een variant hierop als internationale standaard vastgelegd. Deze zou bijgevolg perfect gebruikt kunnen worden als archiveringsformaat. Een andere mogelijkheid vinden we op nieuwe bij de MPEG-familie.

Het Vlaams Parlement is nu bezig aan de opbouw van een digitaal depot. Een aantal voorwaarden om een betrouwbaar digitaal depot te worden zijn al ingevuld, maar verre van allemaal. Steun van het beleid werd al formeel bekrachtigd. Medewerking bij het organisatorische en procedurele onderzoek is daardoor al verzekerd.⁵ Eenmaal dit proces is afgerond kunnen de finale specificaties voor het systeem worden opgesteld en kan een EDRMS ontwikkeld of aangekocht worden. Het moeilijkste gedeelte van deze opdracht begint echter pas daarna. Naast de uitbouw van een softwarearchitectuur zijn - om op een degelijke manier aan digitale archivering te doen - degelijke archiefdocumenten nodig. Dit betekent ondermeer dat gewerkt zal moeten worden met een van hogerhand opgelegde klasseringsstructuur, een vaste wijze van benamingen en contentieus ingevulde metadata. Naar alle waarschijnlijkheid is daar een algemene mentaliteitswijziging voor nodig. Nadat mensen al lang op een bepaalde manier werken, zijn ze meestal niet al te happig op veranderingen die hun op eerste zicht alleen maar meer werk opleveren. Het personeel van het Vlaams Parlement zal goed ingelicht en overtuigd van het nut moeten worden. Zolang de ideeën betreffende goed documentbeheer, die in verband staan met goed archiefbeheer, geen ingang hebben gevonden, is het werk nog niet afgerond. Dit kan op één of twee jaar lukken, of het kan een werk van lange adem zijn. *Only time will tell.*

⁵ en kan in het slechtste geval afgedwongen worden, maar liever niet natuurlijk.

Bibliografie

Algemeen

Chen, S. The paradox of digital preservation. University of Missouri-Columbia, 2001, <http://www.gseis.ucla.edu/us-interpares/pdf/ParadoxOfDigitalPreservation.pdf> , 17/02/06

Cunningham, A. Ensuring essential evidence – Changing archival and records management practices in the electronic recordkeeping era. In: *Provenance, the web magazine*, Vol. 2, N° 2, Spring 1997, <http://www.netpac.com/provenance/vol2no2/features/evidence.htm>, 17/02/06

Curtin, J. Understanding society through its records, <http://john.curtin.edu.au/society/index.html>

Curtin, J. Australian contributions to recordkeeping. In: John Curtin, *Understanding society through its records*, [2003], <http://john.curtin.edu.au/society/australia/index.html>, 11/07/06

Document Management: Integrated System for Scientific Organisations, <http://www.meteo.be/DISSCO/>

Doctorow, C. Metacrap: Putting the torch to seven straw-men of the meta-utopia. 26 august 2001, <http://www.well.com/~doctorow/metacrap.htm#0> , 17/02/06

DPC, The handbook. Digital Preservation Coalition, <http://www.dpconline.org/graphics/handbook/index.html> , 16/03/06

Ellis, J.A. (ed.) Selected essays in electronic recordkeeping in Australia. Victoria: Australian Society of Archivists Inc., 2000

Galloway, P. Preservation of digital objects. In: Blaise Cronin (ed.), *Annual Review of Information Science and Technology*, Volume 38, Medford: Information Today Inc., 2004, pp. 549-590

Janssens, G. & Vermote, M. Het archiefwezen in België. Berchem: VVDAB, 2004, <http://www.fondspascaldecroos.org/organisaties/opleidingen/researchjournalistiek/hennekam/archiefwezenBelgie.PDF> , 05/04/06

Janssens, G. & Put, E. Geschiedenis, principes en terminologie van de archivistiek. Syllabus. Brussel: VUB-KUL-UG-UA, 2005-2006

Ketelaar, E. Tacit narratives: the meanings of archives. In: Thomassen, T., Looper, B., Kloosterman, J. (eds.), *Toegang. Ontwikkelingen in de ontsluiting van archieven*, 's-Gravenhage: Stichting Archiefpublicaties, Jaarboek 2001, pp. 68-75

Levy, D. Heroic measures: Reflections on the possibility and purpose of digital preservation. In: Witten, I.; Akscyn, R.; Shipman III, F. (eds.), *Proceedings of the Third ACM Conference on Digital Libraries*, Pittsburgh, Pennsylvania, United States, June 23-26, New York: ACM Press, 1998, pp.152-161

Position statement on microfilm preservation. ICON: International Coalition on Newspapers, June 15, 2006, <http://icon.crl.edu/icon/statement.htm>, 21/07/06

Scheelings, F. Statisch en dynamisch archief van administraties en bedrijven. Brussel: VUB, Onuitgegeven hoorcollege, 2005

Stuckey, S. Forward. In: Judith A. Ellis (ed.), *Selected essays in electronic recordkeeping in Australia*, Victoria: Australian Society of Archivists Inc., 2000, p. (v)

<http://www.taskforce-archieven.nl/>

Toms, E.G. & Campbell, D.G. Genre as interface metaphor: Exploiting form and function in digital environments. In: *32nd Hawaii International Conference on System Sciences*, Volume 2, IEEE Computer Society, 1999, pp. 2008-2025; zie ook: <http://doi.ieeecomputersociety.org/10.1109/HICSS.1999.772652> , 13/03/06

<http://memory.loc.gov/ammem/about/techIn.html>

Moore, R. Knowledge-based persistent archives. Technical Report. San Diego Supercomputer Center, 2001, <http://www.sdsc.edu/TR/TR-2001-07.doc.pdf>, 15/03/06

Preserv Project, Preserv bibliography on digital preservation. Preservation Eprints Services, <http://preserv.eprints.org/Preserv-bibliography.html>, 12/04/06

Ramirez, R. Concept: Society. Chapter 5: Stakeholder analysis and conflict management. Canada: IDRC Books, 2005, http://www.idrc.ca/en/ev-27971-201-1-DO_TOPIC.html#to205ftn01, 18/07/06

United Nations Educational, Scientific and Cultural Organization, Charter on the preservation of the digital heritage. UNESCO, http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html, 15/12/05

United Nations Educational, Scientific and Cultural Organization, Guidelines for the preservation of digital heritage. <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>, 15/12/05

http://www.unesco.nl/main_6-3.php

Van Driel, M. Traditie en digitaliteit: toegankelijkheid volgens het Proeftuinproject Archiefvormersoverzicht. In: Horsman, P.J., Ketelaar, F. C. J., Thomassen, T. H. P. M., *Context. Interpretatiekaders in de archivistiek*, 's-Gravenhage: Stichting Archiefpublicaties, Jaarboek 2000, pp. 215-243

Webb, C. e.a. Guidelines for the preservation of digital heritage. UNESCO, Information Society Division, National Library of Australia, 2003, <http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>, 13/04/06

Your data at risk. Why you should be worried about preserving electronic records. The National Council on Archives, <http://www.ncaonline.org.uk/materials/yourdataatrisk.pdf>, 13/03/06

Vlaams Parlement

Archiefreglement van het Vlaams Parlement, <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/Pdf/20042005/Archief/Archiefreglement.pdf>, 20/11/05

Archivering van de website, interne nota als voorbereiding op de vergadering van 10/11/2005

De Batselier, N. Een slagvaardig Vlaams Parlement, 2004 – 2009. 3^e Beleidsplan. Februari 2005, <https://intranet.vlaamsparlement.be/docs/portlet/p615dp26lijn487.doc>, 12/07/06

Directie Informatie, Archief en Biografische Dienst, Nota van de Directieraad aan het Uitgebreid Bureau. Opdracht van het archief in een modern Parlement. 17 november 2005

Directie Informatie, Archief en Biografische Dienst, Aandachtspunten voor e-mailarchivering (beknopt). Interne nota, 3 april 2006

Directie Informatie, Archief en Biografische Dienst, Archivering oude en nieuwe website: bewaarstrategie e& planning. Verslag van de vergadering met ICT van 21 juni 2006

Goossens, M. Dertig jaar Vlaams Parlement : historiek en dynamiek van een parlementaire instelling 1971-2001-2002, Kapellen: Pelckmans, 2002

Missieverklaring van het Vlaams Parlement, <http://www.vlaamsparlement.be/vpWeb/p3app/htmlpages/vp/HoeWerktHetVlaamsParlement/Missie/MissieVerklaring.html>, 14/07/07

Notulen van de vergadering Vast Bureau, 19 december 2005

Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 1997-1998. Brussel: Externe Relaties Vlaams Parlement, 1998

Van Esbroeck, K. (ed.) Legislatuurverslag van het Vlaams Parlement, 1995-1999. Brussel: Externe Relaties Vlaams Parlement, 1999

Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 1999 en 1999-2000. Brussel: Externe Relaties Vlaams Parlement, 2000

Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2000-2001. Brussel: Externe Relaties Vlaams Parlement, 2001

Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2001-2002. Brussel: Externe Relaties Vlaams Parlement, 2002

Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2002-2003. Brussel: Externe Relaties Vlaams Parlement, 2003

Van Esbroeck, K. (ed.) Jaarverslag van het Vlaams Parlement, 2004 en 2004-2005. Brussel: Externe Relaties Vlaams Parlement, 2005

Van Esbroeck, M. & Goossens, M. Het Vlaams Parlement. Tielt: Uitgeverij Lannoo, 2003

Van Mechelen, D. De Vlaamse Begroting 2006. Brussel: Vlaamse Regering, 2006, http://www2.vlaanderen.be/ned/sites/financien/Info/Vlaamse_begroting_2006.pdf, 14/07/06

Van Ongevalle, H. Digitaal depot Vlaams Parlement. Conceptnota voor een plan van aanpak. Onuitgegeven, Mei 2006

Preserveringsstrategieën

Aschenbrenner, A. Long-term preservation of digital material – Building an archive to preserve digital cultural heritage. [2001?], http://www.ifs.tuwien.ac.at/~aola/publications/thesis-ando/Long_Term_Preservation.html, 05/04/06

Coy, W. Perspektiven der Langzeitarchivierung multimedialer Objekte. Frankfurt am Main: Nestor c/o Die Deutsche Bibliothek – Berlin: Humboldt-Universität zu Berlin, 2006, <http://nbn-resolving.de/urn:nbn:de:0008-20051214015>, 16/03/06

Hedstrom, M. Digital preservation: problems and prospects. In: Koichi Tabata, Shigeo Sugimoto, Tetsuo Sakaguchi (eds.), *Digital Libraries Journal*, N° 20, March 2001, http://www.dl.slis.tsukuba.ac.jp/DLjournal/No_20/1-hedstrom/1-hedstrom.html, 06/04/06

Linden, J.; Martin, S.; Masters, R. & Parker, R. The large-scale archival storage of digital objects. Digital Preservation Coalition, Technology Watch Series Report 04-03, February 2005, <http://www.dpconline.org/docs/dpctw04-03.pdf>, 15/12/05

Lorie, R. A. A project on preservation of digital data. In: RLG DigiNews, Vol.5, Nr. 3, June 2001, <http://www.rlg.org/legacy/preserv/diginews/diginews5-3.html#feature2>, 04/04/06

Mellor, P.; Wheatley, P.; Sergeant, D. Migration on request, a practical technique for preservation. Leeds: University of Leeds CAMiLEON Project, 2002, <http://www.si.umich.edu/CAMiLEON/reports/migreq.pdf>, 06/04/06

Moore, R.; Baru, C.; Gupta, A.; Ludaescher, B.; Marciano, R. & Rajasekar, A. Collection-based long-term preservation. San Diego: Supercomputer Center, 1999, zie ook: <http://www.sdsc.edu/NARA/Publications/nara.pdf>

Moore, R. Final Report for the research project on application of distributed object computation testbed technologies to archival preservation and access requirements. San Diego: Supercomputer Center, 2001, zie ook : <http://www.sdsc.edu/TR/TR-2001-08.doc.pdf>

National Science Foundation. Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century. Pre-Publication Draft May 26, 2005, http://www.nsf.gov/nsb/documents/2005/LLDDC_report.pdf, 05/12/05

PREMIS Working Group. Implementing Preservation Repositories for Digital Materials: Current Practice and Emerging Trends in the Cultural Heritage Community. OCLC – RLG, Report September, 2004, <http://www.oclc.org/research/projects/pmwg/surveyreport.pdf>, 05/12/05

PREMIS Working Group. Data dictionary of for preservation metadata. OCLC – RLG, Final Report May 2005, <http://www.oclc.org/research/projects/pmwg/premis-final.pdf> , 05/12/05

Records Continuum Research Group, Create once, use many times – the clever use of metadata in egovernment and ebusiness processes in networked environments. Monash University, <http://www.sims.monash.edu/research/rcrg/research/crm/index.html>, 15/12/05

Rothenberg, J. Avoiding technological quicksand: Finding a viable technical foundation for digital preservation. Washington: Council on Library and

Information Resources, 1998; <http://www.clir.org/pubs/reports/rothenberg/contents.html>, 05/04/06

Thibodeau, K. Preservation and migration of electronic records: the state of the issue. NARA, Electronic Records Archives, <http://www.archives.gov/era/papers/preservation.html#top> , 15/03/06

Standaarden

<http://www.adobe.com/nl/products/acrobat/adobepdf.html>

Adobe, PDF as a standard for archiving. White Paper. 2002, http://www.adobe.com/enterprise/pdfs/pdfarchiving.pdf?bcsi_scan_747D1649BA44C3E8=0&bcsi_scan_filename=pdfarchiving.pdf, 16/05/06

<http://www.aiim.org/standards.asp?ID=25013>

ANSI INCITS 27-1987 (R1989), File structure and labeling of magnetic tape for information interchange

ANSI/NISO, The Dublin Core Metadata Element Set. Bethesda: NISO Press, 2001, <http://www.niso.org/standards/resources/Z39-85.pdf>, 29/06/06

Aschenbrenner, A. & Kaiser, M. White paper on digital repositories. reUSE! Digital Master Files of Printed Materials, March 2005, http://www.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf, 12/04/06

Barry, R., K. Development of the Encoded Archival Description DTD. EAD, December 2002, <http://www.loc.gov/ead/eaddev.html> , 06/12/05

Blake, R. Electronic Records Management – the role of TNA. UK: The National Archives, 2003, <http://www.dpconline.org/graphics/events/presentations/pdf/RichardBlakepres.pdf> , 13/03/06

Borghof, U. e.a. Vergleich bestehender Archivierungssysteme. Frankfurt am Main: Nestor c/o Die Deutsche Bibliothek – München: Universität der Bundeswehr München, 2005, <http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:0008-20050117016> , 16/03/06

Bradley, K. ASPR sustainability issues discussion paper. Australian Partnership for Sustainable Repositories, National Library of Australia, January 2005, http://www.apsr.edu.au/documents/APSR_Sustainability_Issues_Paper.pdf?bcsi_scan_747D1649BA44C3E8=0&bcsi_scan_filename=APSR_Sustainability_Issues_Paper.pdf, 12/04/06

Bradley, M. ISO 15489: A practical guide. In: *eGov Monitor*, 16 June 2005, <http://www.egovmonitor.com/node/1401>, 04/05/06

Brübach, N. International standard for archives and records management ISO 15489. In: *Proceedings of the DLM-Forum 2002, @access and preservation of electronic information: best practices and solutions*, Luxemburg: European Communities, Barcelona, 6-8 May 2002, pp. 443-448

Cain, P. MoReq: The standard of the future? In: *Information Management Journal*, March/April 2003, http://www.findarticles.com/p/articles/mi_qa3937/is_200303/ai_n9184315, 30/06/06

Campbell, L.; Blinco, K. & Mason, J. Repository management and implementation. A white paper for alt-i-lab. Australia: DEST, UK: JISC-CETIS, 2004, http://www.jisc.ac.uk/uploaded_documents/Altilab04-repositories.pdf, 12/04/06

The Cedars Project Team & UKOLN, Metadata for digital preservation: The Cedars Project Outline Specification. Draft for public consultation. CURL Exemplars in Digital Archives, March 2000, <http://www.leeds.ac.uk/cedars/MD-STR~5.pdf>, 13/03/06

The Cedars Project Team, The Cedars Project Report. CURL Exemplars in Digital Archives, April 1998-March 2001, <http://www.leeds.ac.uk/cedars/pubconf/papers/projectReports/CedarsProjectReportToMar01.pdf>, 13/03/06

The Cedars Project Team, Cedars guide to: Digital preservation strategies. CURL Exemplars in Digital Archives, April 2002, <http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html>, 13/03/06

Committee on descriptive standards, ISAD (G): General Standard International Archival Description. Stockholm: ICA, 2000, zie ook: [http://www.icacds.org.uk/eng/ISAD\(G\).pdf](http://www.icacds.org.uk/eng/ISAD(G).pdf)

Committee on descriptive standards, ISAAR (CPF). International Standard Archival Authority Record For Corporate Bodies, Persons and Families. Second edition. Cranberra: ICA, 2004, zie ook: http://www.icacds.org.uk/eng/isaar2ndedn-e_3_1.pdf

Configuration Management Team, Configuration Management Plan (CMP v2.3). Draft. NARA, Electronic Records Archives, Program Management Office, July 2004, <http://www.archives.gov/era/pdf/configuration-management-plan.pdf>, 15/03/06

Consultative Committee for Space Data Systems, Reference model for an Open Archival Information System (OAIS). Blue Book. Washington: CCSDS, National Aeronautics and Space Administration, 2002, <http://public.ccsds.org/publications/archive/650x0b1.pdf>, 08/03/06

Papers referencing OAIS <http://nost.gsfc.nasa.gov/isoas/referencing.html>

Consultative Committee for Space Data Systems, Producer-Archive interface methodology abstract standard. Blue Book. Washington: CCSDS, National Aeronautics and Space Administration, 2004, <http://public.ccsds.org/publications/archive/651x0b1.pdf>

Cover, R. (ed.) Technology reports. Encoded Archival Description (EAD). In: *Cover Pages, Online resource for markup language technologies*, 2002, <http://xml.coverpages.org/ead.html>, 29/06/06

Crockett, M. & Foster, J. Using ISO 15489 as an audit tool. In: *Information Management Journal*, July/August 2004, http://findarticles.com/p/articles/mi_qa3937/is_200407/ai_n9455717, 26/05/06

Cumming, K. Evolution and requirements of ISO 15489. Continuum-Create and Maintain, Archives New Zealand, 24 June 2003, <http://www.archives.govt.nz/continuum/dls/pdfs/katec-03-10.pdf>, 04/05/06

Dempsey, L. & Heery, R. Specification for resource description methods Part 1: A review of metadata: a survey of current resource description formats. Bath: The UK Office for Library and Information Networking (UKOLN), The University of Bath, 1999, <http://www.ukoln.ac.uk/metadata/desire/overview/>, 29/06/06

Development of the Encoded Archival Description DTD. Encoded Archival Description, Version 2002, <http://www.loc.gov/ead/eaddev.html>, 29/06/06

<http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=117&categorie=4>, 04/05/06

Digital preservation management: Implementing short-term strategies for long-term problems. Tutorial. <http://www.library.cornell.edu/iris/tutorial/dpm/index.html>, 12/04/06

Digital preservation management: Implementing short-term strategies for long-term problems. Survey of institutional readiness. <http://www.library.cornell.edu/iris/tutorial/dpm/foundation/tdr/readiness.pdf>, 13/04/06

DoD 5015.2-STD. Design criteria standards for electronic records management. Washington: Department of Defense, Washington Headquarters Services, 2002, http://www.dtic.mil/whs/directives/corres/pdf/50152std_061902/p50152s.pdf, 09/03/06

Dublin Core Metadata Initiative, <http://dublincore.org/>

Dublin Core. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Dublin_Core, 29/06/06

Encoded Archival Description, Version 2002, Official Site:
<http://www.loc.gov/ead/>, 29/06/06

ERPANET, Trusted digital repositories for cultural heritage. Final Report. Trusted repositories for preserving cultural heritage. ERPANET Workshop Report, Rome, 17-19 November 2003, <http://www.erpanet.org/events/2003/rome/RomeFinalReport.pdf>, 12/04/06

Fresko, M. & Waldron, M. Model Requirements for the management of electronic records. MoReq Specification. Interchange of Data between Administrations, Prepared by Cornwell Management Consultants plc, October 2002, <http://www.cornwell.co.uk/moreq.html>, 06/11/05

Gable, J. Everything you wanted to know about DoD 50515.2. In: *Information Management Journal*, November/ December 2002, http://www.findarticles.com/p/articles/mi_qa3937/is_200211/ai_n9164040, 30/06/06

Gladney, H.M. Critique: Attributes of a trusted digital repository. October 2001, <http://home.pacbell.net/hgladney/CritiqueTrusted.htm>, 12/04/06

Gladney, H.M. Digital Document Quarterly (DDQ). 2002 - ?, <http://home.pacbell.net/hgladney/ddq.htm>

Hedstrom, M. & Lee, C. Significant properties of digital objects: definitions, applications, implications. In: *Proceedings of the DLM-Forum 2002, @ccess and preservation of electronic information: best practices and solutions*, Luxemburg: European Communities, Barcelona, 6-8 May 2002, pp. 218-227

Heery, R. & Anderson, S. Digital repositories review. UK: UKOLN, The Arts and Humanities Data Service (AHDS), 2005, http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf, 12/04/06

Hokke, E.; Horsman, P., Nelissen, M. & Vanneste, W. ISAD(G)Algemene Internationale Norm voor Archivistisch Beschrijven. Vertaling van de tweede uitgave. Antwerpen-Leuven-Amsterdam, 2004, zie ook: <http://www.ica.org/biblio/ISAD2NL.pdf>

Hofman, H. Review: Some comments on Preservation Metadata and the OAIS Model. In: *Digicult, A newsletter on digital culture*, Issue 2, October 2002, http://www.digicult.info/downloads/digicult_info2.pdf, 13/03/06

Holdsworth, D. & Wheatley, P. Emulation, preservation and abstraction. Leeds: University of Leeds CAMiLEON Project, 2001, <http://cedarsgw.leeds.ac.uk/CAMiLEON//dh/ep5.html>, 06/12/05

Holdsworth, D. Architecture of CEDARS demonstrator. CURL Exemplars in Digital Archives, July 2001, <http://www.personal.leeds.ac.uk/~eclhdh/cedars/architecture.html> , 13/03/06

Horsman, P.J.; van Bussel, G.J. & Waalwijk, H. Softwarespecificaties voor Records Management Applicaties voor de Nederlandse Overheid. Remano 2004. Amsterdam: Archiefschool, maart 2004; Zie ook: http://www.archiefschool.nl/docs/ReMANO_2004.pdf

Integrated Computer Engineering of Directorate of American Systems Corporation, Quality management plan (QMP v2.6). NARA, Electronic Records Archives, Program Management Office, December 2004, <http://www.archives.gov/era/pdf/quality-management-plan.pdf> , 15/03/06

Integrated Computer Engineering of Directorate of American Systems Corporation, Requirements Management Plan (RQM) (Tomp version 2.0, Task 4.3.10). NARA, Electronic Records, Archives, Program Management Office, August 2003, <http://www.archives.gov/era/pdf/requirement-management-plan.pdf> , 15/03/06

International Organization for Standardization, New ISO standard for managing business records. 5 February 2002, <http://www.iso.org/iso/en/commcentre/pressreleases/archives/2002/Ref814.html>, 04/05/06

International Organization for Standardization. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/ISO>, 04/05/06

International Organization for Standardization, ISO in brief. Genève: ISO, 2005, <http://www.iso.org/iso/en/aboutiso/isoinbrief/isoinbrief.html>, 04/05/06

InterPARES 1 Project Team, The long-term preservation of authentic electronic records: Findings of the InterPARES Project. <http://www.interpares.org/book/index.cfm>, 20/11/05

InterPARES 2 Project, International Research on Permanent Authentic Records in Electronic Systems, http://www.interpares.org/ip2/ip2_index.cfm, 20/11/05

ISO/IEC 9529-1 : 1989, Information Processing Systems - Data Interchange on 90 mm (3.5 in) Flexible Disk Cartridges Using Modified Frequency Modulation Recording at 15 916 ftprad on 80 tracks on Each Side - Part 1: Dimensional, physical and magnetic characteristics

ISO/IEC 9529-2 : 1989, Information Processing Systems - Data Interchange on 90 mm (3.5 in) Flexible Disk Cartridges Using Modified Frequency Modulation Recording at 15 916 ftprad on 80 tracks on Each Side - Part 2: Track Format

ISO/IEC 13818-7:1997, Information technology -- Generic coding of moving pictures and associated audio information -- Part 7: Advanced Audio Coding (AAC)

MPEG-4: ISO/IEC 14496:1998, Information technology -- Coding of audio-visual objects

ISO/IEC DIS 9660: 1999 (E), Information processing – Volume and file structure of CD-ROM for Information Interchange, http://www.yadagio.com/public/standards/iso_cdromr/tocont.htm, 11/07/06

ISO 15489-1: 2001 (E), Information and documentation – Records Management –Part 1: General

ISO/TR 15489 –2: 2001 (E), Information and documentation – Records Management –Part 2: Guidelines

ISO/IEC TR 15938-8:2002, Information technology -- Multimedia content description interface -- Part 8: Extraction and use of MPEG-7 descriptions

ISO 15836:2003, Information and documentation – The Dublin Core metadata element set

ISO/TS 23081-1: 2004 (E), Information and documentation – Records Management Processes – Metadata for guidelines – Part1: Principles

ISO 19005-1: 2005, Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF (PDF/A)

ISO 15489 – the essentials. White Paper. Data Capture Solutions, www.datacapture.co.uk/pdf/ISO-15489-White-Paper.pdf , 13/03/06

JaJa, J.; McCall, F.; Smorul, M.; Moore, R.; Chadduck, R. Digital archiving and long term preservation: An early experience with grid and digital library technologies. NARA Papers, 2004, <http://www.archives.gov/era/papers/thic-04.html> , 15/03/06

Jantz, R. & Giarlo, M. Digital preservation. Architecture and technology for trusted digital repositories. In: *D-Lib Magazine*, Volume 11, Number 6, June 2005, <http://www.dlib.org/dlib/june05/jantz/06jantz.html>, 12/04/06

JCDL, Workshop on digital curation & trusted repositories: Seeking success. held June 11-15, 2006, <http://sils.unc.edu/events/2006jcdl/digitalcuration.html>

Lavoie, B. The Open Archival Information System Reference Model: Introductory guide. DPC Technology Watch Series Report 04-01, Dublin:

OCLC, Inc., 2004, http://www.dpconline.org/docs/lavoie_OAIS.pdf , 08/03/06

Lavoie, B. & Gartner, B. Preservation metadata. Digital Preservation Coalition, Technology Watch Series Report 05-01, September 2005, <http://www.dpconline.org/docs/reports/dpctw05-01.pdf>, 15/12/05

LeFurgy, W. PDF/A : Developing a file format for long-term preservation. In: *RLG Diginews*, Volume 7, Number 6, 15 december 2003, http://www.rlg.org/preserv/diginews/v7_n6_feature1.html, 20/06/06

Lynch, C. Institutional repositories: Essential infrastructure for scholarship in the digital age. In: *ARL Bimonthly Report* 226, February 2006, <http://www.arl.org/newsltr/226/ir.html>, 12/04/06

Metadata Advisory Group, EAD – Encoded Archival Description. MIT Libraries, 2004, <http://libraries.mit.edu/guides/subjects/metadata/standards/ead.html>, 29/06/06

Metadata Encoding & Transmission Standard (METS), <http://www.loc.gov/standards/mets/> , 07/12/05

McHugh, A. Open source for digital curation. In: *Digital Curation Manual*, Instalment, 2005, <http://www.dcc.ac.uk/resource/curation-manual/chapters/open-source/opensource.pdf>, 12/04/06

McLean, B. The ISO 15489 imperative. In: *Information Management Journal*, Nov/Dec 2002, http://findarticles.com/p/articles/mi_qa3937/is_200211/ai_n9163957, 26/05/06

McLeod, J. & Ferguson, S. ISO 15489 – Implementation and impact. In: *NorthumbriaOnline, Information Management Research Institute*, May 20, 2004, http://online.northumbria.ac.uk/faculties/art/information_studies/Imri/rarea/rm/rmproj/rmcurr/Impl/Impl.htm, 04/05/06

<http://www.msscience.com/iso.html>

National Archives of Australia. Recordkeeping metadata standard for Commonwealth agencies. Commonwealth of Australia, 1999, <http://www.naa.gov.au/recordkeeping/control/rkms/summary.htm> , 20/02/06

The National Archives, Guidance for an inventory of electronic record collections: A toolkit. Surrey: Public Record Office, 2000, http://www.nationalarchives.gov.uk/electronicrecords/advice/pdf/inventory_toolkit.pdf, 06/11/05

The National Archives, Evaluating information assets: appraising the inventory of electronic records. Surrey: Public Record Office, [2001], http://www.nationalarchives.gov.uk/electronicrecords/advice/pdf/appraisal_toolkit.pdf, 06/11/05

The National Archives, Requirements for Electronic Records Management Systems. Part 1: Functional requirements. Surrey: Public Record Office, 2002, Zie ook: <http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/requirementsfinal.pdf>

The National Archives, Requirements for Electronic Records Management Systems. Part 2: Metadata standard. Surrey: Public Record Office, 2002, Zie ook: <http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/metadafinal.pdf>

The National Archives, Requirements for Electronic Records Management Systems. Part 3: Reference document. Surrey: Public Record Office, 2002, Zie ook: <http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/referencefinal.pdf>

The National Archives, Requirements for Electronic Records Management Systems. Part 4: Implementation guidance. Surrey: Public Record Office, 2004, <http://www.nationalarchives.gov.uk/electronicrecords/reqs2002/pdf/implementation.pdf>, 06/11/05

The National Archives of Australia, The DIRKS Manual: A strategic approach to managing business information. Commonwealth of Australia, 2001 (revised 2003), <http://www.naa.gov.au/recordkeeping/dirks/dirksman/dirks.html>, 06/11/05

National Archives of Australia. Guidelines for creating, managing and preserving digital records. Exposure draft. Canberra: Commonwealth of Australia, May 2004, <http://www.naa.gov.au/recordkeeping/er/guidelines.html> , 24/03/06

National Archives of Australia. Guidelines for implementing the functional specifications for electronic records management systems software. Exposure draft. Canberra: Commonwealth of Australia, 2006, <http://www.naa.gov.au/recordkeeping/er/erms/ERMSguidelines.pdf> , 20/02/06

National Archives of Australia. Functional specifications for electronic records management systems software. Exposure draft. Canberra: Commonwealth of Australia, 2006, <http://www.naa.gov.au/recordkeeping/er/erms/ERMSspecifications.pdf> , 20/02/06

National Archives of Australia, Australian standard for Records Management – AS ISO 15489. <http://www.naa.gov.au/recordkeeping/rkpubs/advice/advice58.html>, 20/11/05

NEN-ISO 15489-1: 2001 nl. http://www.rijksarchiefinspectie.nl/publicaties/naslag_digitalisering_NEN_ISO.html, 04/05/06

NEN-ISO 15489-1 (nl), Informatie en documentatie – Informatie- en archiefmanagement – Deel 1: Algemeen (ISO 15489-1: 2001, IDT), Delft: Nederlands Normalisatie-instituut, 2001

NPR-ISO/TR 15489-2 (nl), Informatie en documentatie – Archiefbeheer – Deel 2: Richtlijnen (ISO/TR 15489-2: 2001), Delft: Nederlands Normalisatie-instituut, 2002

NISO Framework Advisory Group, A framework of guidance for building good digital collections. 2nd edition. Bethesda, MD: National Information Standards Organization, 2004, <http://www.niso.org/framework/framework2.html> , 19/04/06

Pitti, D. Encoded Archival Description. An introduction and overview. In: *D-Lib Magazine*, Volume 5, Number 11, November 1999, <http://www.dlib.org/dlib/november99/11pitti.html>, 29/06/06

http://partners.adobe.com/public/developer/pdf/index_reference.html

Photorel, Theoretisch kader. Photographic Heritage E-Learning, http://www.culturelestudies.be/student/photherel/beschrijving_3.html, 29/06/06

Portable Document Format. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Portable_Document_Format, 20/07/06

RLG/OCLR, Trusted digital repositories. Attributes and responsibilities. An RLG-OCLR Report. Mountain View: RLG, 2002; <http://www.rlg.org/en/pdfs/repositories.pdf> , 15/03/06

RLG, An audit checklist for the certification of trusted digital repositories. Draft for public comment. Mountain View: Research Libraries Group, 2005, http://www.rlg.org/en/page.php?Page_ID=20769, 12/04/06

RLG, Encoded Archival Description Activities. RLG, 2006, http://www.rlg.org/en/page.php?Page_ID=411, 29/06/06

Roberts, D. The new Australian Records Management Standard. 28 March 2000, http://www.records.nsw.gov.au/publicsector/rk/sacramento/sacra_1.htm

Rusbridge, C. Excuse me ... some digital preservation fallacies? In: *Ariadne*, Issue 46, February 2006, <http://www.ariadne.ac.uk/issue46/rusbridge/>, 12/04/06

Siepel, A. & de Vries, H. Encoded Archival Description (EAD). Het digitaliseren van toegangen op archieven. Amsterdam: Archiefschool, 2001, <http://www.archiefschool.nl/docs/siepenco.pdf>, 29/06/06

Stephens, D. The why and how of international records retention. In: *Information Management Journal*, Sept/Oct 2005, http://findarticles.com/p/articles/mi_qa3937/is_200509/ai_n15350578, 26/05/06

Stoehr, T. Un format PDF pour l'archivage a été normalisé, mais est-ce un format ouvert? 6 novembre 2005, <http://formats-ouverts.org/blog/2005/11/06/604-un-format-pdf-pour-larchivage-a-ete-normalise-mais-est-ce-un-format-ouvert->, 20/07/06

Table of commentary. AS ISO-15489-2002, Records Management. State Records New South Wales, 6 January 2004, <http://www.records.nsw.gov.au/publicsector/standards/tableofcommentary.htm> , 26/05/06

Thibodeau, K.; Moore, R. & Baru, C. Persistent object preservation: Advanced computing infrastructure for digital preservation. In: *Proceedings of the DLM-Forum 1999, European citizens and electronic information: the memory of the information society*, Luxemburg: European Communities, Brussel, 18-19 October 1999, http://europa.eu.int/ISPO/dlm/fulltext/full_thib_en.htm , 24/03/06

The Victorian Electronic Records Strategy, <http://www.prov.vic.gov.au/vers/vers/default.htm>, 10/11/05

Waldron, M. MoReq: Model Requirements for the Management of Electronic Records. In: B. Duhon (ed.), *AIIIM E-Doc Magazine*, September/ October 2002, http://www.edocmagazine.com/article_new.asp?ID=25504, 30/06/06

Weibel, S.; Godby, J.; Miller, E.; Daniel, R. OCLC/NCSA Metadata Workshop Report. <http://dublincore.org/workshops/dcl/report.shtml>, 29/06/06

Wheatley, P. Institutional repositories in the context of digital preservation. DPC Technology Watch Series Report 04-02, 2004, <http://www.dpconline.org/docs/DPCTWf4word.pdf>, 12/04/06

Onderzoek

Boudrez, F. Het digitaal archiveringssysteem: beheersinventaris, informatietoelagen en beslissingsmodel als uitgangspunt. Antwerpen: Stadsarchief, DAVID Project, 2001,

<http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Rapport4.pdf> , 17/02/06

Boudrez, F. Archief onder controle? De beheersinventaris als instrument bij de archivering van digitale archiefdocumenten. Antwerpen: Stadsarchief, 2002, zie ook: <http://www.antwerpen.be/david/website/teksten/DAVIDbijdragen/Beheersinventaris.pdf>

Boudrez, F. <XML/> en digitaal archiveren. Antwerpen: Stadsarchief, 2002, zie ook: http://www.antwerpen.be/david/website/teksten/XML_digitaalarchiveren.pdf

Boudrez, F. XML Topic Maps voor digitale archivering. Antwerpen: Stadsarchief, 2002, zie ook: http://www.antwerpen.be/david/website/teksten/DAVIDbijdragen/XML_topicmaps.pdf

Boudrez, F. Magnetische dragers voor het archief. Antwerpen: Stadsarchief, 2002, zie ook: http://www.antwerpen.be/david/website/teksten/DAVIDbijdragen/Magnetische_dragers.pdf

Boudrez, F. Standaarden voor digitale archiefdocumenten. Technisch rapport. Antwerpen: Stadsarchief, 2001 -2003, zie ook: <http://www.expertisecentrumdavid.be/davidproject/teksten/DAVIDbijdragen/Standaarden.pdf>

Boudrez, F. Preserving electronic records from database-driven information systems. In: *Digicult.Info, A newsletter on digital culture*, Issue 5, October 2003, pp. 18-20, http://www.digicult.info/downloads/digicultinfo_issue5_october_2003.pdf

Boudrez, F. & Dekeyser, H. Digitaal archiefbeheer in de praktijk. Antwerpen: Stadsarchief, 2004, zie ook: http://www.expertisecentrumdavid.be/davidproject/teksten/DAVID_vademecum.pdf

Dhérent, C. Les archives électroniques. Manuel pratique. Direction des Archives de France, 2002, http://www.archivesdefrance.culture.gouv.fr/fr/archivistique/DAF_manuel%20version%207.html , 16/02/06

Documentation Abstracts Inc. The state of digital preservation: an international perspective. Conference proceedings, Washington D.C., April 24-25 2002. Washington, D.C.: Council on Library and Information Resources (CLIR), 2002; <http://www.clir.org/pubs/reports/pub107/pub107.pdf> , 15/03/06

ERPANET, XML for digital preservation. ERPANET Workshop report, Urbino, 9-11 October, 2002, <http://www.erpanet.org/events/2002/urbino/Urbino%20Workshop%20Report.pdf>, 15/12/05

ERPANET, Metadata in digital preservation. Training Seminar, Marburg, September, 3-5, 2003, <http://www.erpanet.org/events/2003/marburg/finalMarburg%20report.pdf>, 15/12/05

ERPANET, The Long-term preservation of databases. ERPANET Workshop report, Bern, April 9-11, 2003, http://www.erpanet.org/events/2003/bern/Bern_Report_final.pdf, 15/12/05

Gilliland-Swetland, A. Advanced workshop in preserving and providing long-term access to archival electronic records. South-Carolina: Archives and History Center, June 12, 2003, <http://www.palmettohistory.org/agspresentation/index.htm>, 06/12/05

Heazlewood, J. Management of electronic records over time. In: Judith A. Ellis (ed.), *Selected essays in electronic recordkeeping in Australia*, Victoria: Australian Society of Archivists Inc., 2000, pp. 97-118

Heazlewood, J. & Quenault, H. Blurring the records boundary: Electronic records, between standards and practice. Victoria: Public Record Office, ASA & RMAA Conference, September 2001, <http://www.prov.vic.gov.au/vers/pdf/010904p.pdf> , 02/03/06

Hedstrom, M. It's about time: Research challenges in digital archiving and long-term preservation. Final report. Workshop april 12-13 2002. The Library of Congress: Digital Friedlander, A. Background. Summary of results from interviews and essays [by the Council on Library and Information Resources (CLIR)]. The Library of Congress: Digital Preservation, National digital information infrastructure and preservation program, 2001, http://www.digitalpreservation.gov/repor/interviews_summary.pdf , 16/02/06

JISC, Records Management. http://www.jisc.ac.uk/index.cfm?name=recordsman_home, 15/12/05

JISC, ITT: Electronic Records Management System implementation toolkit. http://www.jisc.ac.uk/index.cfm?name=funding_erm, 15/12/05

<http://www.nationalarchives.gov.uk/electronicrecords/advice/default.htm#toolkits>

Preservation, National digital information infrastructure and preservation program, 2003, http://www.digitalpreservation.gov/about/NSF.pdf?bcsi_scan_747D1649BA44C3E8=0&bcsi_scan_filename=NSF.pdf , 16/02/06

Rijksarchiefinspectie, De digitale erfenis. Onderzoek naar het beheer van digitale archieven. Case study 1: De Arbeidsvoorziening. 2003, http://www.rijksarchiefinspectie.nl/publicaties/onderzoek_rapp_arbeidsvoorziening.html, 15/11/05

Roberts, D. Making electronic records accessible. In: Judith A. Ellis (ed.), *Selected essays in electronic recordkeeping in Australia*, Victoria: Australian Society of Archivists Inc., 2000, pp. 119-134

Sinclair, K. The Victorian Electronic Records Strategy: A strategy developed by the state government of Victoria (Australia) to preserve electronic records for the long term. Victoria: Public Record Office, 2001, <http://www.prov.vic.gov.au/vers/pdf/acarm.pdf> , 02/03/06

Task Force on Archiving of Digital Information, Preserving digital information. Commission on Preservation and Access/ Research Libraries Group, 1996, <http://www.rlg.org/legacy/ftpd/pub/archtf/final-report.pdf>, 05/04/06

Teil, J.-P. Constance, la première migration technologique. Archives de France, 1999, <http://www.archivesdefrance.culture.gouv.fr/fr/archivistique/DAFconstance.html> , 16/02/06

Testbed Digitale Bewaring, Migratie: Context en huidige stand van zaken. Den Haag: ICTU, 2001 ; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white_paper_migratie.pdf

Testbed Digitale Bewaring, XML en digitale bewaring. Den Haag: ICTU, 2002; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white_paper_xml-nl.pdf

Testbed Digitale Bewaring, Emulatie: Context en huidige stand van zaken. Den Haag: ICTU, 2003 ; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/docs/white_paper_emulatie_Ned.pdf

Testbed Digitale Bewaring, Van digitale vluchtigheid naar digitaal houvast, Deel 1: Kosten- en beslissingsmodellen/Functionele specificaties. Bewaren van databases. Den Haag: ICTU, 2003 ; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/Bewaren_van_databases.pdf

Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 2: Bewaren van spreadsheets. Den Haag: ICTU, 2003 ; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/Bewaren_van_spreadsheet_s.pdf

Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 3: Bewaren van tekstdocumenten. Den Haag: ICTU, 2003 ; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/docs/bewaren_van_tekstdocumenten.pdf

Testbed Digitale Bewaring, digitale vluchtigheid naar digitaal houvast, Deel 4: Bewaren van e-mail. Den Haag, ICTU, 2003 ; zie ook: http://www.digitaleduurzaamheid.nl/bibliotheek/docs/bewaren_van_email.pdf

Van den Eynde, S. Wat archiveren en hoe? Op zoek naar de rol van PKI voor digitale archieven. DAVID Rapport 6. Antwerpen: Stadsarchief, 2001, zie ook: <http://www.antwerpen.be/david/website/teksten/Rapporten/Rapport6.pdf>

Woodyard, D. Significant property: digital preservation at the British Library. In: *VINE: The Journal of Information and Knowledge Management Systems*, Volume 34, Number 1, Emerald Group Publishing Limited, 2004, pp. 17-20, <http://www.emeraldinsight.com/Insight/ViewContentServlet?Filename=Published/EmeraldFullTextArticle/Articles/2870340102.html> , 16/02/06

Documenttypes

Adoption of ODF in Europe. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Adoption_of_ODF_in_Europe, 03/05/06

Boudrez, F. Hoe archiveer je digitale kantoordocumenten? In: *Lokaal voor Gemeente en OCMW*, Nummer 7, Brussel: Vereniging van Vlaamse Steden en Gemeenten vzw, 1-15 april 2003, pp. 17-19, <http://www.expertisecentrumdavid.be/davidproject/teksten/DAVIDbijdragen/Lokaal7.pdf> , 17/02/06

Deckmyn, D. Belgian government chooses OpenDocument. In: *CNET News.com*, June 23, 2006, http://news.com/Belgian+government+chooses+OpenDocument/2100-7344_3-6087275.html?tag=nefd.top, 20/07/06

Folk, M. & Barkstrom, B. R. Attributes of file formats for long-term preservation of scientific and engineering data in digital libraries. The National Archives: NCSA, 2003, http://www.ncsa.uiuc.edu/NARA/Sci_Formats_and_Archiving.doc, 15/12/05

Grath, R. E. XML and scientific file formats. Urbana-Champaign: University of Illinois, 2003, zie ook: http://www.ncsa.uiuc.edu/NARA/XML_and_Binary.pdf

Heuscher, S. SIARD under the hood. Technical aspects of SIARD. Presentation at the Urbino Erpanet Congress, 10 april 2003, www.erpanet.org/events/2003/bern/presentations/WS%20Bern%20Presentation%20Heuscher.ppt, 16/07/06

Lai, E. Google backs OpenDocument format. In: *Computerworld*, July 11, 2006, <http://www.computerworld.com/action/article.do?Command=view>

[ArticleBasic&articleId=9001713&bcsi_scan_747D1649BA44C3E8=3s8ko7306n12R7CcILyLwoAAABzvpcB&bcsi_scan_filename=article.do](http://www.clir.org/pubs/abstract/pub93abst.html), 20/07/06

Lawrence, G. W.; Kehoe, W. R.; Rieger, O.Y.; Walters, W.H. and Kenney, A.R. Risk management of digital information: A file format investigation. CLIR Report, Cornell University, 2000, <http://www.clir.org/pubs/abstract/pub93abst.html> , 17/03/06

MPEG4IP: Open source, open standards, open streaming. <http://www.mpeg4ip.net/>, 03/05/06

<http://www.oasis-open.org/home/index.php>

OASIS Open Document Format for Office Applications (OpenDocument) TC, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office, 20/07/06

OpenDocument. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Open_Document, 03/05/06

www.openoffice.org

Mission Statement, <http://about.openoffice.org/index.html>, 15/03/06

<http://www.opensource.org/>

Ross, S.; Donnely, M. & Dobрева, M. Emerging technologies for the cultural and scientific heritage sector. DigiCULT Technology Watch Report 2, February 2004, <http://www.digicult.info/downloads/html/1082026370/1082026370.html>, 16/07/06

Sustainability of digital formats, Planning for Library of Congress Collections, <http://www.digitalpreservation.gov/formats/index.shtml> , 16/02/06

Wusteman, J. Formats for the electronic library. In: *Ariadne, Issue 8*, April 1997, <http://www.ariadne.ac.uk/issue8/electronic-formats/intro.html> , 16/02/06

E-mail

Boudrez, F. Digitaal archiveren: richtlijn en advies nr. 1: Archiveren van e-mail. Antwerpen: Stadsarchief, DAVID-Project, 2005, <http://www.expertisecentrumdavid.be/davidproject/teksten/Richtlijn1.pdf>, 09/03/06

Boudrez, F.; Dekeyser, H. & Van den Eynde, S. Archiveren van e-mail. DAVID Rapport 5. 2^{de} editie. Antwerpen: Stadsarchief DAVID Project, 2003, zie ook: http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Rapport5_V_2.pdf

Boudrez, F. E-mailarchieven. E-mails: hoe bewaren en goed archiveren? Antwerpen: Stadsarchief DAVID Project, 2003; <http://www.expertisecentrumdavid.be/davidproject/teksten/tr.pdf>

Boudrez, F. Digitaal archiveren van archiefdocumenten. Casus: e-mail. Antwerpen : Stadsarchief Antwerpen, Onuitgegeven lezing, 6 maart 2006

Boudrez, F. E-mails: hoe klasseren en archiveren? Antwerpen: Stadsarchief, 2006

<http://www.novell.com/products/groupwise/>

<http://www.novell.com/news/press/archive/2001/03/pr01031.html>,
19/07/06

Novell Workgroup, A collaboration solution that supports your business goals. Business White Paper. Novell, 2006, p. 5, <http://www.novell.com/collateral/4621449/4621449.pdf>, 20/07/06

Audio

Analog-to-digital converter. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Analog_to_digital_converter , 01/03/06

Baaten, L. Rapport. Krachtlijnen conserverings- en digitaliseringsbeleid. Geluidsarchief. Antwerpen: Stadsarchief, 2003; http://www.cdavid.be/pdf/rapport_geluidsarchief.pdf , 15/03/06

Baaten, L. & Vandermaesen, M. Digitaal geluidsarchief. Krachtlijnen digitalisering: Standaarden, formaten dragers. CDAVID, Erfgoed digitaal, 2004, http://www.cdavid.be/pdf/TechnischRapport_DigitaalGeluidsarchief.pdf , 17/02/06

Brylawski, S. Preservation of digitally recorded sound. The Library of Congress: Digital Preservation, National digital information infrastructure and preservation program, [2002], <http://www.digitalpreservation.gov/index.php?nav=3&subnav=7> , 16/02/06

Digital-to-analog converter. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Digital_to_analog_converter , 01/03/06

Elsea, P. Basics of digital recording. Converting sound into numbers. University of California, Santa Cruz, Division of the Arts, 1996, http://arts.ucsc.edu/EMS/Music/tech_background/TE-16/teces_16.html , 22/02/06

IBM Corporation & Microsoft Corporation, Multimedia Programming Interface and Data Specifications 1.0. August 1991, <http://www.kk.iij4u.or.jp/~kondo/wave/mpidata.txt> of http://www.tactilemedia.com/info/MCI_Control_Info.html,

Introduction to Pulse Code Modulation (PCM). In: *ICT Technologies, Network for Capacity Building and Knowledge Exchange in ICT Policy, Regulation and Applications*, NetTel@Africa Off-line content, 2004, <http://cbdd.wsu.edu/kewlcontent/cdoutput/TR502/page13.htm> , 23/02/06

Kondratiuk, C. Technical aspects of multimedia audio. In: Jessica Keyes (ed.), *The ultimate multimedia handbook*, New York: McGraw-Hill, 1997 (2nd ed.), pp. 27.1-27.23

Morse, K. Introduction to digital video and audio. In: Jessica Keyes (ed.), *The ultimate multimedia handbook*, New York: McGraw-Hill, 1997 (2nd ed.), pp. 32.1-32.24

Roland-Mieszkowski, M. Introduction to digital recording techniques. Digital Recordings, 1989-2005, <http://www.digital-recordings.com/publ/pubrec.html> , 22/02/06

Smith, S.W. The scientist and engineer's guide to digital signal processing. California Technical Publishing, 1997 (updated 12/2001), <http://www.dspguide.com/pdfbook.htm> , 01/03/06

Teruggi, D. Can we save our audio-visual heritage? In: *Ariadne, Issue 39*, April 2004, <http://www.ariadne.ac.uk/issue39/teruggi/intro.html> , 16/02/06

Timmerer, C. Information technology – MPEG-21 multimedia framework. Department of Information Technology, University of Klagenfurt, January 2005, <http://mpeg-21.itec.uni-klu.ac.at/cocoon/mpeg21/>, 09/05/06

Seadle, M. Sound practice: A report of the Best Practices for Digital Sound Meeting, 16 January 2001 at the Library of Congress. In: *RLG Diginews, Volume 5, Number 2*, 15 april 2001, <http://www.rlg.org/legacy/preserv/diginews/diginews5-2.html#feature3> , 16/02/06

Staller, L. Understanding analog to digital converter specifications. In: *Embedded.com*, 24/02/2005, <http://www.embedded.com/showArticle.jhtml?articleID=60403334> , 01/03/06

WAV. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Wav>, 20/07/06

Waveform coding techniques. Cisco Systems Inc., 18/02/2005, http://www.cisco.com/warp/public/788/signalling/waveform_coding.pdf, 01/03/06

Waveformatex. http://msdn.microsoft.com/library/default.asp?url=/library/en-us/multimed/htm/_win32_waveformatex_str.asp, 26/07/06

Windows Media Codecs, <http://www.microsoft.com/windows/windowsmedia/forpros/codecs/audio.aspx>, 20/07/06

Video

Betacam. In: *Answers.com*, <http://www.answers.com/topic/digital-betacam>, 09/04/06

Boudrez, F. Videoarchivering: bruggen bouwen op technologisch drijfzand. Antwerpen: Stadsarchief, 2004, zie ook: http://www.antwerpen.be/david/website/teksten/Presentaties/Abstract_Technologisch_Drijfzand.pdf

Girod, C., SMPTE releases VC-1 standard. SMPTE Press Release, April 3, 2006, http://www.smpte.org/news/press%5Freleases/003_06.cfm, 09/05/06

Poynton, C. Video engineering. January 2003, <http://www.poynton.com/Poynton-video-eng.html>, 09/04/06

Stroeykens, S. MPEG-21 scheidt orde in de multimedia-chaos. In: *De Standaard*, 21 oktober 2002, p. 17

TASI, Advice paper. Digital cameras. Technical Advisory Service for Images, 2003, <http://www.tasi.ac.uk/advice/creating/camera.html>, 15/03/06

Vandermaessen, M. Rapport. Digitaal beeldarchief. Aanbevelingen voor het realiseren van beeldbanken, aan de hand van 'lessons learnt' en 'best practices'. Antwerpen: Stadsarchief, cDAVID, 2005; http://www.cdavid.be/pdf/Rapport_beeldbank_LR.pdf, 02/03/06

Van Horik, R. Permanent pixels. Building blocks for the longevity of digital surrogates of historical photographs. Proefschrift (doctoraatsverhandeling), Delft: Technische Universiteit Delft, 2005, http://www.dans.knaw.nl/over_dans/organisatie/rene_van_horik/publicaties/permanentpixels/Permanent_pixels_proefschrift.pdf, 16/03/06

Wactlar, H. D. & Christel, M.G. Digital video archives: managing through metadata. Carnegie: Mellon University, [2002], <http://www.digitalpreservation.gov/index.php?nav=3&subnav=10> , 16/02/06

Windows Media Video. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Windows_Media_Video, 09/04/06

WMV. In: *Audio Video Affair*, <http://www.audio-video-affair.com/wmv.html>, 09/05/06

Young, R. Digital video for multimedia basics. San Francisco State University's College of Extended Learning, Multimedia Studies Program, juli 1998 <http://msp.sfsu.edu/Instructors/rei/VIDEOPGE.HTM> , 22/02/06

Websites

Arms, W. ; Adkins, R. ; Ammen, C. & Hayes, A. Collecting and preserving the web: The Minerva Prototype. In: *RLG Diginews, Volume 5, Number 2*, 15 april 2001, <http://www.rlg.org/legacy/preserv/diginews/diginews52.html#feature1>, 16/02/06

Aschenbrenner, A. Long-term preservation of digital material – Building an archive to preserve digital cultural heritage from the internet. [2002?], http://www.ifs.tuwien.ac.at/~aola/publications/thesis-ando/Long_Term_Preservation.html , 08/03/06

Boudrez, F. & Van den Eynde, S. Archiveren van websites. Antwerpen: Stadsarchief DAVID Project, 2002, zie ook: <http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Rapport7.pdf>, 17/02/06

Boudrez, F. Digitaal Archiveren: richtlijnen & advies nr5. Websitebeheer en content management voor digitale archivering. Antwerpen: Stadsarchief: DAVID, 2005, <http://www.expertisecentrumdavid.be/davidproject/teksten/Richtlijn5.pdf> , 15/03/06

Boudrez, F. Archiveren van websites: een kwestie van waardering en 'capture'. Antwerpen: Expertisecentrum DAVID vzw, 2005, http://www.expertisecentrumdavid.be/docs/archiveren_websites.pdf , 17/02/06

Dollar, C. Archival preservation of Smithsonian web resources: Strategies, principles, and best practices. July 2001, http://siarchives.si.edu/pdf/dollar_report.pdf , 16/03/06

Hakala, J. Collecting and preserving the web: Developing and testing the NEDLIB Harvester. In: *RLG Diginews, Volume 5, Number 2*, 15 april 2001, <http://www.rlg.org/legacy/preserv/diginews/diginews5-2.html#feature2> , 16/02/06

Hokke, H.A. Naar archivering van websites. Implementatieadvies bij het onderzoeksrapport "Blijvend beschikbaar". Amsterdam: Archiefschool, Ministerie van Verkeer en Waterstaat, 2003, http://www.elo.nl/elo/Images/implarchwebs_tcm70-20515.doc , 16/03/06

Hokke, E. Eerste hulp bij website archivering. Bibliografie websitearchivering. Amsterdam: Archiefschool, februari 2006, <http://www.archiefschool.nl/docs/ehbw.pdf> , 02/03/06

Lyman, P. Archiving the World Wide Web. Berkeley: University of California, [2001], <http://www.digitalpreservation.gov/index.php?nav=3&subnav=5> , 16/02/06

NARA, Guidance on managing web records. January 2005, <http://www.archives.gov/records-mgmt/pdf/managing-web-records-index.pdf> , 16/03/06

National Archives of Australia. Guidelines for keeping records of web-based activity in the Commonwealth government. Commonwealth of Australia, 2001, http://www.naa.gov.au/recordkeeping/er/web_records/guide_contents.html , 22/02/06

Voorburg, R.J.J. Webontwerp: richtlijnen voor archivering. Uselab.com – Ministerie van Verkeer en Waterstaat, 2004, http://digidiv.amsterdam.nl/handleidingen_en_instrumenten/documentvormen/websites/webont.pdf , 16/03/06

ICT

Biemond, H. Woordenboek ICT. Arnhem: Elsevier Bedrijfsinformatie BV, 2002

Boudrez, F. Digitale handtekeningen en archiefdocumenten. Antwerpen: Expertisecentrum DAVID vzw, 2005, zie ook: http://www.expertisecentrumdavid.be/docs/digitalehandtekeningen_archiefdocumenten.pdf, 29/03/06

Codd, E.F. A relational model for large shared data banks. In: *Communications of the ACM*, Vol. 13, No 6, Association for Computing Machinery Inc., June 1970, pp. 377-387, zie ook: <http://www.acm.org/classics/nov95/toc.html>, 19/03/06

Company profile Arbor, <http://www.arbor-audio.com/index.php?Action=company>, 20/05/06

Committee on Innovations in Computing and Communications: Lessons from History, Funding a revolution. Government support for computing research. Washington: National Academy Press, 1999, <http://www.nap.edu/readingroom/books/far/ch6.html>, 20/06/07

Computer cluster. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Computer_cluster, 19/07/06

Cumming, A. A gentle introduction to SQL. Edinburgh: School of Computing, Napier's University, 2005, <http://sqlzoo.net/w.htm>, 16/11/05

Data center. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/Datacenter>, 19/07/06

Digital signature. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Digital_signature, 29/03/06

Electronic signature. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Electronic_signature, 29/03/06

Ellison, C. & Schneier, B. Ten risks of PKI: What you're not being told about Public Key Infrastructure. In: *Computer Security Journal*, Volume XVI, Number 1, 2000, <http://www.schneier.com/paper-pki.pdf>, 29/03/06

End-of-life. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/End-of-life>, 19/07/06

<http://www.fastpcnet.net/datacenter.html>, 16/07/06

Gagné, M. New life for old Microsoft Access data. In: *Oracle Technology Network*, 2004, http://www.oracle.com/technology/pub/articles/gagne_access.html, 15/04/06

Greenwald, R.; Stackowiack, R. & Stern, R. Oracle essentials: Oracle9i, Oracle8i & Oracle8 (2nd edition). US: O'Reilly & Associates, 2001, <http://www.developer.com/db/article.php/1582621>, 07/04/07

Hash function. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Hash_function, 29/03/06

Hillesund, T. Digital text cycles: From medieval manuscripts to modern markup. In: *Journal of Digital Information (JoDI)*, Volume 6, Issue 1, 09/03/2005, <http://jodi.tamu.edu/Articles/v06/i01/Hillesund/>, 02/03/06

Information Security Committee, Digital signature guidelines. Tutorial. American Bar Association, Section of Science and technology, <http://www.abanet.org/scitech/ec/isc/dsg-tutorial.html>, 29/03/06

<http://java.sun.com/javacc/>

Java 2 Enterprise Edition. In: *Wikipedia, De Vrije Encyclopedie*, <http://nl.wikipedia.org/wiki/J2EE>, 15/02/06

Kroenke, D. Database processing. Fundamentals, design, and implementation. International edition. New Jersey: Pearson Education Inc., 2006

Litchfield, D. Hackproofing Oracle Application Server (A guide to securing Oracle 9). NGSSoftware Insight Security Research, 2002, <http://www.ngsssoftware.com/papers/hpoas.pdf>, 22/03/06

Lawrence Livermore National Laboratory From kilobytes to petabytes in 50 years. University of California, 2003, <http://www.llnl.gov/50science/computation.html>, 26/07/06

Load balancing. In: *Webopedia*, http://www.webopedia.com/TERM/L/load_balancing.html, 19/07/06

LogDepot Audio Specification, <http://www.arbor-audio.com/index.php?action=product&productid=2&view=specification>

<http://www.pki-page.org/> , 29/03/06

Menezes, A.; van Oorschot, P. & Vanstone, S. Handbook of applied cryptography. Boca Raton: CRC Press, 1996

Microsoft Access. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Microsoft_Access, 20/07/06

Naudé, F. Oracle Internet/Web Application Server (iAS) FAQ. Version 2.82, 2001, <http://www.orafaq.com/faqweb.htm>, 22/03/06

Wat? In: *DigitaleHandtekening.be*, <http://www.digitalehandtekening.be/what.cfm>, 27/07/06

Oracle Application Server Portal. Graz University of Technology, 2004, http://fstgss16.tu-graz.ac.at/portalHelp/ohw/state?navSetId=_&navId=0&destination=, 22/03/06

Panda, D. Oracle Containers for J2EE (OC4J). In: *OnJava.com, The independent source for enterprise Java*, 2002, <http://www.onjava.com/pub/a/onjava/2002/01/16/oracle.html>, 15/02/06

Public-key cryptography. In: *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Public-key_cryptography , 29/03/06

RSA Laboratories, RSA Laboratories' Frequently Asked Questions about today's cryptography, Version 4.1. RSA Security Inc., 2000, <http://www.rsasecurity.com/rsalabs/node.asp?id=2152> , 29/03/06

SAN. In: *Webopedia*, 2006, <http://www.webopedia.com/TERM/S/SAN.html>, 20/07/06

Schneier, B. Why cryptography is harder than it looks. <http://www.schneier.com/essay-037.html> , 29/03/06

SAP history: From start-up software vendor to global market leader. <http://www.sap.com/company/history.epx>, 20/07/06

<http://www.sap.com/solutions/index.epx>

SAP AG. In: *Wikipedia, Die freie Enzyklopädie*, http://de.wikipedia.org/wiki/SAP_AG, 20/07/06

SAP: Profile for SAP Aktiengesell ADS. In: *Yahoo! Fincance*, <http://finance.yahoo.com/q/pr?s=SAP>, 20/07/06

SQL. In: *Wikipedia, The Free Encyclopedia*, <http://en.wikipedia.org/wiki/SQL>, 16/11/05

Tate, J.; Lucchese, F.; Moore, R. Introduction to Storage Area Networks. IBM Redbooks, 2004, <http://www.redbooks.ibm.com/abstracts/sg245470.html>, 20/07/06

Van den Eynde, S. Wat archiveren en hoe? Op zoek naar de rol van PKI voor digitale archieven. Antwerpen: Stadsarchief, DAVID Project, 2001, <http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Rapport6.pdf> , 29/03/06

Van Rompay, B.; Preneel, B.; Vandewalle, J. The digital timestamping problem. In: A. Barbé, E.C. van der Meulen, P. Vanroose (eds.), *Proceedings of the 20th symposium on information theory in the Benelux*, Werkgemeenschap voor Informatie- en Communicatietheorie, 1999, pp. 71-78, <http://www.cosic.esat.kuleuven.ac.be/publications/article-122.pdf> , 29/03/06

Wat? In: *DigitaleHandtekening.be*, <http://www.digitalehandtekening.be/what.cfm>, 27/07/06

Wat is PKI? <https://www.pkioverheid.nl/asp/get.asp?xdl=../views/pki/xdl/Page&VarIdt=00000001&SitIdt=00000002&ItmIdt=00000444> , 29/03/06

Het wat en waarom van PKI. https://www.pkioverheid.nl/contents/pages/00000447/het_wat_en_waarom_van_pki.pdf , 29/03/06

Williams, S. The associative model of data. Lazy Software Ltd., 2002, 2nd edition; zie ook: http://www.lazysoft.com/docs/other_docs/AMD.pdf , 22/02/06

Zaïane, O. CMPT 354: Database systems and structures. Simon Fraser University, School of Computing Science, 1998, <http://www.cs.sfu.ca/CC/354/zaiane/material/notes/Chapter5/node2.html> , 13/03/06

Wetgeving

Coppein, B. Een onneembare burcht? Een toelichting bij de Belgische archiefwet en de pogingen om haar te vervangen. Lezing gehouden op de FAAD-studiedag “50 jaar Belgische archiefwet” van 23 april 2005, <http://www.vub.ac.be/archivistiek/FAAD/files/bc230405.doc> , 05/04/06

Dekeyser, H. Digitale archivering: een juridische stand van zaken vanuit Belgisch perspectief. Deel 2. Auteursrecht, technische beschermingsmaatregelen en wettelijk depot. Antwerpen, Stadsarchief, DAVID Project, 2004, <http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Rapport8.pdf> , 29/03/06

Directive 1999/93/EC of the European Parliament and of the council of 13 December 1999 on a Community framework for electronic signatures. In: *Official Journal of the European Communities*, 19 January 2001, http://europa.eu.int/eur-lex/pri/en/oj/dat/2000/l_013/l_01320000119en00120020.pdf , 05/04/06

<http://eid.belgium.be/>

<http://www.microsoft.com/belux/nl/office/eid/word.aspx>

Fedict, Elektronische identiteitskaart: eID. Februari 2006, <http://www.belgium.be/eportal/application?languageParameter=nl&pageid=contentPage&docId=6418>, 19/07/06

Fedict, eID gebruiken. <http://eid.belgium.be/nl/navigation/documents/39763.html>, 19/07/06

Public Law 106-229 – June 30, 2000, Electronic signatures in global and national commerce Act, http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=106_cong_public_laws&docid=f:publ229.106.pdf, 05/04/06

Regeling geordende en toegankelijke staat archiefbescheiden. 23 februari 2002, nr. WJZ/2002/6839(8080), verschenen in: *Staatscourant*, nr. 43, 1 maart 2002, p. 8; zie ook https://secure-web.eur.nl/paz/Regelingen/index/edit/Bijlage_2.pdf , 22/03/06

Van den Eynde, S. Digitale archivering: een juridische stand van zaken vanuit Belgisch perspectief. Deel 1. Antwerpen, Stadsarchief, DAVID Project, 2001, <http://www.expertisecentrumdavid.be/davidproject/teksten/Rapporten/Rapport3.pdf> , 29/03/06

Wet tot invoering van het gebruik van telecommunicatiemiddelen en van de elektronische handtekening in de gerechtelijke en buitengerechtelijke procedure, 20 oktober 2000 (BS 22 december 2000), <http://www.ejustice.just.fgov.be/cgi/article.pl>, 05/04/06