

Richtlijnen voor het gebruik van elektronische informatie

Hoe om te gaan met machineleesbare gegevens en elektronische documenten

DLM-Forum on Electronic Records

First published in English as
Guidelines on best practices for using electronic information
How to deal with machine-readable data and electronic documents
By the Office for Official Publications of the European Communities,
2 rue Mercier, L-2985 Luxembourg (1997)
for the Commission of the European Communities.
©European Communities, 1997
This translation © Van Heijst Information Consulting, 1998.

Verantwoordelijkheid voor de vertaling ligt geheel bij
Van Heijst Information Consulting

EUROPEAN
ARCHIVES NEWS

INSAR

. Secretariaat-Generaal van de Europese Commissie

Comité van voorbereiding:

Lino FACCO

Nicola BELLIENI

Hans HOFMANN

Christina BECKERS

Peter BERNINGER

Bas VAN DEN BARG

Emilio ROSSETTI

Johanna MAAS

Voor meer informatie:

Europese Commissie,

Tel. (32-2) 296 08 27

295 05 57

Fax (32 2) 296 10 95

E-mail: archis@sg.cec.be

Adres: Wetstraat 200

Office: SDME 5/64

B-1049 Brussels

Deze multidisciplinaire richtlijnen zijn oorspronkelijk samengesteld door Jean-Michel Cornu, adviseur, in nauwe samenwerking met de Historische archieven van de Europese Commissie en experts van de lidstaten. Een ontwerpversie werd opgesteld en gedistribueerd aan de deelnemers van het DLM-Forum on electronic records (gehouden in Brussel, 18 tot 20 december 1996): ISBN 92-827-9129-7 (EN,FR,DE). Het ontwerp werd geactualiseerd door Syllis (Lille, Frankrijk) en bevat commentaren en annotaties die zijn ontvangen gedurende het Forum en die nadien werden toegezonden aan de Historische Archieven van de Europese Commissie. De vertaling is van A.L.M. van Heijst, adviseur documentaire informatievoorziening.

Deze geactualiseerde en vermeerderde editie van de Richtlijnen is één van de concrete resultaten van het DLM-Forum. Het illustreert de voortgaande samenwerking van diverse disciplines (overheidsadministratie, archieven, industrie en onderzoek) op het gebied van de elektronische informatie en het informatiemanagement in de zich snel ontwikkelende informatiemaatschappij.

Aanvullende informatie over de Europese Unie is te vinden op Internet. Het kan worden bereikt via de Europa Server (<http://europa.eu.int>).

INHOUDSOPGAVE

VOORWOORD

1. INTRODUCTIE

- 1.1 Over de multidisciplinaire richtlijnen
- 1.2 Het gebruik van de multidisciplinaire richtlijnen
- 1.3 Van individuele productie naar algemeen gebruik
- 1.4 Drie stadia in de informatie levenscyclus

2. VAN GEGEVENS NAAR GESTRUCTUREERDE INFORMATIE

- 2.1 Wat is informatie?
- 2.2 Wat zijn gegevens?
- 2.3 Wat maakt elektronische informatie zo anders?
- 2.4 Het aanmaken van elektronische records
- 2.5 Twee manieren om gegevens te structureren
 - 2.5.1 Documenten
 - 2.5.2 Databases
- 2.5.3 Aanschaf van hard- en software voor elektronische informatie

3. DE INFORMATIELEVENSCYCLUS EN HET TOEKENNEN VAN VERANTWOORDELIJKHEDEN

- 3.1 Overzicht van de informatie-levenscyclus
- 3.2 Het definiëren van verantwoordelijkheden voor elk stadium

4. ONTWERP, CREATIE EN ONDERHOUD VAN ELEKTRONISCHE INFORMATIE

- 4.1 Het ontwerpstadium
- 4.2 De creatie van elektronische informatie
- 4.3 Integratie, conversie en waardering van elektronische informatie
 - 4.3.1 Integratie van informatie
 - 4.3.2 Conversie van informatie
 - 4.3.3 Van papier en microfilm naar een gescand image
 - 4.3.4 Van een gescand image naar een gecodeerd formaat (via OCR)
 - 4.3.5 Van het ene digitale formaat naar het andere
 - 4.3.6 De waardering van elektronische informatie
- 4.4 Management en classificatie van elektronische gegevens
- 4.5 Overdracht

5. BEWARING VAN GEGEVENS VOOR KORTE EN LANGERE TERMIJN

5.1 Media voor de opslag van gegevens

5.2 Bestandsformaten

5.2.1 Compressie van bitmaps

5.2.2 Vectorbestanden

5.2.3 Tekstbestanden

5.2.4 Gegevens en programma's

6. HET VERKRIJGEN VAN TOEGANG TOT- EN DE VERSPREIDING VAN INFORMATIE

6.1 Naar een informatiemaatschappij

6.2 Standaards voor de toegang van informatie

6.3 Beveiliging

6.3.1 Toegangsrechten

6.3.2 Encryptie en het bewijs van de authenticiteit

6.4 Toegang tot gegevens

7. CONCLUSIES

8. ANNEXEN

8.1 Terminologie

8.2 Lijst van nog op te lossen problemen

8.3 Het waarom van standaardisatie

8.4 Checklist voor een strategie op het gebied van elektronische informatie

8.5 Prototype: welke metadata moet worden ontwikkeld?

8.6 Prototype: hoe de juiste standaarden te selecteren

8.7 De DLM-vervolgactiviteiten

8.8 Bibliografie

VOORWOORD

Het DLM-Forum, georganiseerd in samenwerking tussen de lidstaten van de Europese Unie en de Europese Commissie in Brussel in December 1996, bracht een groep experts samen uit de industrie, de onderzoekswereld, de administratie en archieven om een onderwerp te bespreken dat steeds belangrijker wordt: het geheugen van de informatiemaatschappij. Enkele jaren geleden had nog niemand zich kunnen voorstellen dat archiefwerkzaamheden zo revolutionair zouden veranderen door technologische innovaties en dat, door het ontbreken van tijdige en doeltreffende maatregelen, de authenticiteit en lange-termijn conservering van vitale informatie ernstig zou worden bedreigd in de nabije toekomst.

Archiefdiensten vormen een essentieel onderdeel van de informatiemaatschappij. Sinds de publicatie van het rapport "Europe and the global information society (het Bangemann Rapport) in 1994 is het belang van archieven in de actie- en ondersteuningsprogramma's van de Commissie toegenomen. Archiefdiensten zullen bovendien een belangrijke rol spelen bij het moderne informatiebeheer. Tegen deze achtergrond heeft het DLM Forum een helder signaal gegeven en is de bal nu aan het rollen geraakt.

De Richtlijnen voor het gebruik van elektronische informatie vormen een van de belangrijkste resultaten van het DLM-Forum. Een voorlopig ontwerp van de Richtlijnen werd aan de deelnemers gedistribueerd als basis voor discussie. Hierop werden door landelijke deskundigen tijdens en na het Forum meer dan 300 voorstellen tot wijziging en verbetering ingediend bij de Europese Commissie. Als gevolg daarvan werd deze herziene en uitgebreide editie van de Richtlijnen geproduceerd.

Ik ben ervan overtuigd dat deze multidisciplinaire, in samenwerking ontwikkelde Richtlijnen zullen helpen bij het definiëren van een strategie voor de korte- en middellange termijn, die nodig is om de acute problemen in het beheer en de conservering van elektronisch opgeslagen data op te lossen. Beslissers bij de overheid en in de industrie worden opgeroepen om samen met deskundigen op het gebied van elektronische archivering de noodzakelijke maatregelen te nemen en praktische oplossingen te ontwikkelen, waarbij op Europese schaal wordt samengewerkt. Hierdoor zou het vertrouwen van bedrijven en private personen in de betrouwbaarheid van elektronische informatie, opgeslagen op nieuwe media, worden versterkt.

De structuur van de Richtlijnen maakt ze tot een rijke informatiebron voor deskundigen en andere geïnteresseerden. Zij vormen een belangrijke stap voorwaarts in zowel het zeker stellen van het geheugen van de informatiemaatschappij door conservering als in het verhelfen van het beslissingsproces in administraties van de lidstaten en de instituties van de Europese Unie. Dit voor alle burgers.

Dr. Martin Bangemann

Lid van de Europese Commissie, verantwoordelijk voor
Industriële aangelegenheden, informatie- en telecommunicatietechnologie.

1. INTRODUCTIE

1.1 Over de multidisciplinaire richtlijnen

Documenten en gegevensbestanden: toenemend gebruik van elektronische media

De overgrote meerderheid van documenten komen nog steeds in papieren vorm voor. Maar vele hiervan -elektronische berichten, memo's of verslagen van de vergaderingen die ons werkzame leven structuur geven- worden geproduceerd met behulp van elektronische hulpmiddelen. Door de snelheid van elektronische post is het aantal elektronische documenten binnen overheid en industrie drastisch toegenomen. Als resultaat hiervan begint het digitale formulier zelfs het papier te vervangen. Om die reden bestaat er een toenemende behoefte om de invloed van dit fenomeen te beschouwen op de dagelijkse praktijk en de regels voor het gebruik van documenten en de archivering.

Veel organisaties hebben ook hun eigen gegevensbestanden ontwikkeld. Deze bevatten vaak informatie die, nadat deze een bepaald doel heeft gediend- voor een bepaalde termijn bewaard dient te worden uit juridische overwegingen of voor later onderzoek. Er dient daarom aandacht te worden geschonken aan de opslag op lange termijn, de toegankelijkheid van de informatie en de mogelijkheden om uit het gegevensbestand nieuwe informatie te genereren.

Voorbeelden van "de beste aanpak" die helpen in het definiëren van een strategie voor elektronische informatie

Deze multidisciplinaire richtlijnen hebben niet de pretentie om een volledig en definitieve antwoord te geven op elke vraag. Er zal nog veel discussie nodig zijn tussen verschillende betrokken disciplines, zoals de overheidsadministratie, de archiefdiensten, de industrie en de onderzoekswereld. De richtlijnen geven echter al een aantal voorbeelden hoe in bepaalde gevallen bij voorkeur gehandeld dient te worden en kunnen organisaties tot hulp zijn bij het definiëren van hun eigen strategie op het gebied van elektronische informatie. De doelstelling is niet om een éénduidige, Europa-brede projectaanpak te presenteren, maar meer om de deskundigheden uit verschillende landen en verschillende disciplines te bundelen, zodat ieder er voordeel mee kan doen.

Een ontwerp van deze multidisciplinaire richtlijnen werd door de Europese Commissie gedistribueerd aan de deelnemers van het DLM-Forum over elektronische records, dat plaatsvond in Brussel van 18 tot 20 december 1996. Als gevolg hiervan werden commentaar en aanvullingen ontvangen van nationale deskundigen en werden verdere discussies gehouden over de voorbereiding van deze geactualiseerde en uitgebreide editie.

De richtlijnen kunnen ook worden gebruikt samen met de "*Guide on electronic records*" van de International Council on Archives, die het elektronisch archiefbeheer behandelen vanuit het gezichtspunt van de archivaris.

De DLM-keten houdt verschillende typen acties in:

- *specifieke indicaties;*
- *studies*
- *ondersteuning van gebruikers*

De voortgezette activiteiten van het DLM-Forum werden samengevat in de zogenaamde "Tien Punten"⁽¹⁾ (zie de cirkels in het volgende figuur) Zij zijn gerelateerd aan de gehele keten van de

¹ 'Proceedings of the DLM-Forum on electronic records, Brussel 18-20 december 1996', *INSAR - European Archives News*, supplement II, EUR-OP, Luxembourg, 1997, p. 353 en op internet: <http://www.echo.lu/dlm/en/home.html>

productie en het onderhoud van elektronische records. Zij staan onder toezicht van het DLM-Monitoring Committee dat als één van de vervolgmaatregelen werd ingesteld. Deze acties houden in:

- Ondersteuning van de gebruikers met training en multidisciplinaire richtlijnen;
- Een uitgebreide studie over de relatie tussen overheidsadministraties en archieven;
- Specifieke aanwijzingen voor functionele vereisten van DLM-specificaties, voor IT-standaards ten behoeve van leveranciers van apparatuur en programmatuur en standaardisatie-instituten, voor juridische aspecten ten behoeve van DLM-managers en voor de toegankelijkheid van informatie voor de burger en de onderzoeksgemeenschap.

Een multidisciplinaire aanpak, waarbij rekening wordt gehouden met de behoeften van alle betrokken partijen

Deze multidisciplinaire richtlijnen zijn zowel bedoeld voor personen met basiskennis van elektronisch records management als voor personen met een uitgebreider kennis uit verschillende werkvelden, in het bijzonder op het gebied van

- Overheidsadministratie
- Archiefdiensten
- Industrie (software en hardwareleveranciers)
- De onderzoeksgemeenschap

In de lidstaten en instituten van de Europese Unie.

1.2 Over het gebruik van de multidisciplinaire richtlijnen

Om het de lezer gemakkelijker te maken om de juiste informatie te vinden die aansluit bij zijn of haar kennisniveau, is gebruik gemaakt van kaders.

De basisbegrippen: dit kader geeft een korte beschrijving van de basisconcepten die op verschillende gebieden worden gehanteerd

CD-ROM Basisbegrippen (volgt een beschrijving van de basisconcepten)

Voor gevorderden
Dit kader geeft meer technische uitleg voor hen die detailinformatie wensen

Grafische bestanden Voor gevorderden (volgt een beschrijving van een detail-onderwerp)

Voorbeeld
Een voorbeeld geeft een toegepaste oplossing

Duitsland Voorbeeld (wordt voorbeeld gegeven)
--

Optie
Dit kader geeft advies of een beslissingsboom waardoor het mogelijk wordt om de juiste keuze te maken

Scanning van een document Opties (in het kader wordt een advies omschreven dat helpt een keuze te maken)

Standards. Niet alle standards hebben een zelfde gewicht

Niet alle standards die in deze richtlijnen zijn geciteerd, zijn van een gelijk niveau. Wij hebben steeds aangegeven in hoeverre zij betrouwbaar zijn.

1.3 Van individuele productie tot algemeen gebruik

Een document of een gegevensbestand, vervaardigd door een individueel persoon, maakt onderdeel uit van het informatiesysteem van de organisatie

Waarom is het zo belangrijk om een strategie voor de langere termijn te hebben bij de creatie, het actualiseren of het leveren van informatie?

Zelfs wanneer documenten of gegevensbestanden worden vervaardigd door één of meerdere individuen, zullen zij vaak voor meer mensen van belang zijn dan aanvankelijk werd verondersteld.

- Ze kunnen gebruikt of geactualiseerd worden door iemand die geen deel had aan de creatie.
- Ze zullen vaak worden gebruikt in samenhang met andere records, waarvan enkele niet eens bekend zijn bij de originele auteur.
- De informatie die zij bevatten kan nog lang na de dynamische fase worden gebruikt, bij voorbeeld voor juridische of historische doeleinden.

Een document of gegevensbestand dient daarom te worden gezien als een onderdeel van een breder informatiesysteem. Het is van vitaal belang dat het behoud, de toegang tot de informatie en de bescherming van de privacy vanaf de aanvang worden meegewogen.

1.4 Drie stadia in de informatie levenscyclus

Drie stadia : ontwerp, creatie, onderhoud

Er zijn drie hoofdstadia in de levensperiode van elektronische informatie.

Ontwerp

In dit stadium wordt de totaalstrategie neergelegd.

Creatie

In dit stadium worden de gegevens vervaardigd, gewoonlijk door een beperkt aantal mensen

Onderhoud

Dit houdt in het gebruik en het behoud van de gegevens.

Deze drie stadia kunnen niet alle op dezelfde wijze worden benaderd. Er zijn verschillen aangaande de wijze waarop de gegevens worden geactualiseerd, het aantal maal dat zij worden benaderd, de verantwoordelijkheden enz.

2. VAN GEGEVENS NAAR GESTRUCTUREERDE INFORMATIE

2.1 Wat is informatie?

***Informatie** is een mededeling of een gebeurtenis die ter kennis wordt gebracht van een persoon of een groep personen*

Een informatie-eenheid is een mededeling of een gebeurtenis die ter kennis wordt gebracht aan een persoon of groep personen. Informatie kan worden aangemaakt, onderhouden, bewaard en verspreid.

Informatie is de basis waarop bedrijfsprocessen worden ontwikkeld. Informatie is inmiddels zo belangrijk geworden dat we al spreken over de informatiemaatschappij, en over de informatierevolutie als een nieuwe stap in de industriële maatschappij.

2.2 Wat zijn gegevens?

***Gegevens** kunnen worden samengevoegd tot een document of een lijst*

Gegevens zijn de basiseenheden van informatie.

In een document bij voorbeeld worden veel gegevens-elementen samengebracht om een argument te presenteren of een activiteit te beschrijven. Tot voor kort werden de meeste gegevens op papier vastgelegd, bewaard en uitgewisseld (vroeger op nog andere media, zoals steen).

Soms treffen we gegevens aan in de vorm van lijsten, zoals een telefoonboek, waarbij de bedoeling niet is om een mening te geven of een activiteit te beschrijven, maar om materiaal te leveren voor een verdere actie (bij voorbeeld het vinden van een telefoonnummer). Dan is het belangrijk dat gegevens dusdanig worden ingericht dat het zoeken zo snel mogelijk gaat en trefzeker is. Om die reden zijn de namen in alfabetische volgorde geplaatst, gerangschikt naar stad of gebied.

Het is mogelijk om gegevens op andere media dan papier op te slaan. Informatie kan bijvoorbeeld elektronisch worden opgeslagen om de bewerking te vereenvoudigen.

***Gegevensbeheer** is van een toenemende complexiteit*

Doordat de technologische ontwikkelingen zo snel gaan wordt het steeds moeilijker om een duurzame oplossing te vinden voor het bewaren en conserveren van de gegevens. Elektronische media ontwikkelen zich sneller en hebben een veel kortere levensverwachting dan papier.

Maar ook het volume van de geproduceerde informatie is enorm gegroeid, vooral bij de overheid. Als gevolg hiervan wordt het classificeren en structureren van gegevens steeds complexer. Er moeten immers waarborgen worden geschapen dat de gegevens lang na hun productie nog volledig, authentiek, betrouwbaar, toegankelijk, beschikbaar en leesbaar zijn.

2.3 Wat maakt elektronische informatie zo anders?

***Het medium is niet de boodschap.** Er is een hulpmiddel nodig om elektronische informatie te kunnen lezen*

Als gegevens zijn opgeslagen op een elektronisch medium kunnen ze niet worden gelezen zonder een specifiek hulpmiddel te gebruiken. Gewoonlijk is dat een computer, maar ook software is nodig om de gegevens te kunnen lezen.

Het medium maakt echter doorgaans geen deel uit van de boodschap, zoals dit wel het geval is bij een papieren document. In deze multidisciplinaire richtlijnen wordt met elektronische informatie bedoeld: de gegevens die zijn opgeslagen in een dussdanige vorm, dat zij automatisch kunnen worden bewerkt. In het algemeen gesproken zal dit zijn op een elektronisch medium.

Elektronische informatie brengt voordelen en beperkingen voor het gegevensbeheer

Er zijn onmiskenbare voordelen om gegevens op een medium op te slaan dat direct door een machine gelezen kan worden:

- Het is gemakkelijker om gegevens machinaal te bewerken. Is verandering van een detail noodzakelijk, dan is het niet noodzakelijk om alle gegevens opnieuw in te voeren;
- Elektronische opslag vraagt minder ruimte;
- Het kopiëren van documenten is eenvoudig;
- informatie is snel te verplaatsen

Het gebruik van elektronische media houdt echter ook belangrijke beperkingen in.

- De mens heeft steeds een hulpmiddel nodig om de gegevens te lezen;
- Elektronische media hebben een korter leven dan papier of microfilm;
- Doordat een origineel eenvoudiger te dupliceren is, ontstaan problemen met de bewijskracht van documenten en met de vraag naar authenticiteit;
- De snelle verandering in technologie en op de informatiemarkt maakt het moeilijk om stabiele softwareformats te gebruiken die langdurig bruikbaar zijn.

2.4 Het aanmaken van elektronische records

Gegevens zijn opgeslagen op een medium. Bij elektronische informatie kunnen de gegevens worden bewerkt, gecommuniceerd en geïnterpreteerd door een computer. Een hoeveelheid gegevens kan een document vormen.

Wat is een record? Definitie van de International Council on Archives

Het Committee on Electronic Records van de International Council on Archives (ICA) definieert een record als: "een specifieke hoeveelheid vastgelegde informatie, verzameld of ontvangen bij het begin, de uitvoering of afwikkeling van een activiteit en die voldoende inhoud, omgevingscontext en structuur heeft om bewijs te leveren van een activiteit."

Bij de overheidsadministratie heeft een document juridische waarde, omdat het dient ter informatie of bewijs. De term "record" wordt hier begrepen in administratieve en archivalische betekenis, *niet in de technische zin zoals dit gebruikt wordt in de computerterminologie (als een record van een databank)*.

Dat het record los staat van het medium waarop het is vastgelegd, heeft verschillende consequenties. Het is bij voorbeeld gemakkelijk om een elektronisch record te kopiëren van het ene medium naar het andere. Dit maakt het eenvoudiger om informatie te kopiëren of te verspreiden, maar het wordt ook moeilijker -om niet te zeggen cruciaal- om aan te geven wat nu het origineel is en wat de kopie (zie hiervoor annex 8.2- de lijst van nog op te lossen problemen).

De gegevens die deel uitmaken van het elektronisch record moeten een samenhangende en consistente hoeveelheid informatie leveren. Het opstellen van een classificatieschema is één van de belangrijkste taken om records goed te kunnen definiëren. Er zijn vaak verschillende mogelijkheden om de gegevens in een record te groeperen, afhankelijk van het gewenste detailniveau. Een record kan bij voorbeeld bestaan uit een database, maar ook uit een samenhangende set gegevens uit die database.

Elektronische records bestaan uit vier elementen:

- inhoud
- structuur
- context
- presentatie. Deze laatste kan niet worden geconserveerd, omdat dit vooral afhankelijk is van het medium waarop de gegevens zijn opgeslagen.

Elektronische records

Voor gevorderden

1. Een elektronisch record bestaat uit vier hoofdelementen. De eerste drie moeten worden behouden.
2. De inhoud van een record kan verschillende typen gegevens bevatten:
 - Tekst (bladzijden, paragrafen, zinnen, woorden);
 - Cijfers (absolute getallen zowel als variabelen);
 - Tabellen (complete tabellen of cellen);
 - Tekeningen, grafische beelden, geluid en video;
 - Hypertext-verbindingen met andere documenten.
3. De logische structuur van een record kan onderdeel uitmaken van het document of de database zelf, maar kan daarvan ook los staan, waarbij dezelfde structuur gebruikt kan worden voor meerdere records. De logische structuur kan sterk afwijken van de fysieke structuur van het record.
4. De context van een record wordt beschreven in een gerelateerd record dat de meta-data levert aan de hand waarvan het document kan worden teruggezocht. De metadata kunnen inhouden:
 - De technische metadata (hard- en software-omgeving, versienummers, bestandsstructuur, een beschrijving van de gegevens en de geschiedenis van de verbindingen of links met andere records, logboekgegevens over het gebruik en het wijzigen van de gegevens);
 - Een beschrijving van de administratieve context. Is een record geïntegreerd in een netwerk-architectuur, dan zal de context, zoals deze beschreven is in de documentatie, zeer complex zijn.
5. De presentatie (in het bijzonder van documenten) wordt steeds vaker apart behandeld van het record zelf, waardoor de informatie onafhankelijk is van de presentatiewijze. Verspreiding van informatie op verschillende media noemt men wel 'cross-media management'. Er zijn nog geen hulpmiddelen op de markt die het mogelijk maken om opgeslagen gegevens na een langere termijn, bij voorbeeld enkele jaren, terug te halen op de wijze waarop zij nu worden gerepresenteerd. Dit moet nog worden uitgevonden.

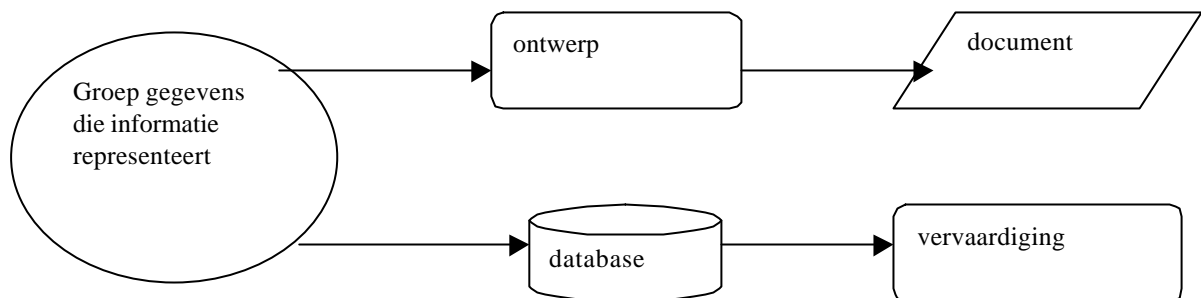
2.5 Twee manieren om gegevens te structureren

***Gegevensbestanden en documenten:** twee typen gegevens die steeds vaker worden samengevoegd tot een 'compound (samengesteld) document'.*

Om een specifiek informatie-item te kunnen vinden, dient het gestructureerd te zijn. Afhankelijk van de doelstelling van de informatie zijn er twee manieren om gegevens te structureren.

- **Gegevensbestanden** of databases: gegevens zijn geplaatst in een omgeving, van waaruit zij kunnen worden teruggehaald en geactualiseerd.

- **Documenten:** deze structuur wordt gebruikt wanneer gegevens zijn samengesteld in een bepaalde samenhang om een argument weer te geven of een activiteit te beschrijven. Een document kan dienen als bewijs voor een bepaalde activiteit, bij voorbeeld als juridisch instrument. In dit geval dient het als record te worden vastgelegd.



Figuur 1: document of gegevensbank?

Natuurlijk zijn er veel verschillende manieren waarop gegevens kunnen worden gestructureerd. Nu informatie een steeds belangrijker rol gaat spelen in organisaties, wordt de structuur van zowel documenten als gegevensbestanden steeds complexer. (Denk bij voorbeeld aan een on-line pagina of op CD-ROM die automatisch wordt gevuld met gegevens die uit een database worden opgehaald).

Bovendien is er vaak een hechte band tussen een gegevensbestand en een hoeveelheid documenten. Wanneer bij voorbeeld de gegevens in een databank worden geactualiseerd dan kan dit gevolgen hebben voor een document dat automatisch volgens een administratieve procedure wordt geproduceerd uit de gegevens van die databank. Deze benadering wordt steeds meer gebruikt bij workflow-hulpmiddelen, die de werkprocessen binnen een bedrijf regelen.

Er zijn een aantal verschillende benaderingen mogelijk, afhankelijk van de vraag of we te maken hebben met een document of een gegevensbestand. Bij voorbeeld:

- Het behoud van elk type (gegevensbestanden en documenten) leidt tot verschillende eisen;
- Een document moet consistent zijn en steeds hetzelfde blijven (actualisering moet dus leiden tot een geheel nieuw document dat opgeslagen wordt als een nieuwe versie);
- Een database daarentegen kan echter regelmatig worden geactualiseerd;
- Uitgaande van deze twee tegengestelde benaderingen, leidt informatie die is samengesteld uit een gegevensbestand en een document tot geheel nieuwe vragen.

2.5.1 Documenten

Er zijn verschillende stadia in de omschakeling van een kantoor waarin de informatie nog is gebaseerd op papieren documenten naar een 'less paper office', waar informatievoorziening is gebaseerd op elektronische documenten.

1. Het traditionele kantoor, gebaseerd op papieren informatiestromen en -documenten;
2. Een kantoor waar een mix voorkomt van papier en elektronische informatie;
3. De conversie van papieren documenten, door bij voorbeeld scanning;
4. Een volledig elektronisch kantoor waar alle documenten worden geproduceerd, ontvangen en verspreid op elektronische wijze.

Deze richtlijnen richten zich hoofdzakelijk op de stadia 2 en 3, omdat deze de huidige situatie in kantoren weergeven. Zij geven de relatie aan tussen elektronische documenten en traditionele documenten (hoofdzakelijk op papier en microfiche). Stadium 4 geeft eerder een richting aan waarin organisaties zich in de toekomst zouden kunnen ontwikkelen dan een echte fase. Papieren en

elektronische documenten zullen waarschijnlijk nog gedurende een lange periode naast elkaar gebruikt blijven, alhoewel door nieuwe kantoorconcepten in combinatie met telewerken wel steeds meer informatie elektronisch zal worden opgeslagen en beschikbaar gesteld.

Documenten komen in verschillende vormen voor: brieven, notities, memo's, formulieren, rapporten enz.-, elke vorm heeft een eigen behandeling nodig.

Documenten kunnen gegroepeerd worden tot dossiers of 'mappen', waardoor een samengestelde informatie-eenheid ontstaat. Het toegankelijk maken van documenten is een belangrijke factor in het eenvoudig en snel terugvinden van de juiste informatie. Worden de stukken zaaksgewijze gegroepeerd, dus alle correspondentie betreffende een specifieke casus bij elkaar, dan is "alle informatie betreffende een samengesteld, eindigend complex van handelingen, gericht op een bepaald doel" samengebracht. Deze groepering dient dan nog wel een specifieke aanduiding te krijgen, zodat de stukken kunnen worden geïdentificeerd als bij elkaar horende verzameling.

Classificatie en trefwoorden: twee aanvullende hulpmiddelen bij het zoeken naar informatie

Basisbegrippen

Ontsluiting van documenten

Er zijn verschillende manieren waarop documenten toegankelijk kunnen worden gemaakt. De belangrijkste zijn:

- chronologisch: elk document krijgt een serienummer zodra het is vastgelegd;
- onderwerpsgewijze: elk document krijgt een nummer volgens een classificatieschema oftewel een dossier-ordeningsplan.

De tweede hiervan is doorgaans het meest effectief om een record snel en eenvoudig te vinden. Er dient dan wel eerst een efficiënte archiefcode (classificatieplan) te worden ontworpen.

Een andere oplossing om informatie te vinden is met behulp van trefwoorden, of om in de tekst zelf te zoeken via full-text search.

De relatie tussen het classificatieplan en het trefwoordensysteem lijkt op die tussen een inhoudstabel en een alfabetische index van een boek. Beide zijn efficiënte zoekinstrumenten, maar ze kunnen niet los van elkaar worden gebruikt.

Moderne informatietechnieken kunnen nieuwe manieren om informatie toegankelijk te maken toevoegen aan de bestaande. Vooral op het gebied van de full-text search wordt veel onderzoek gedaan naar geautomatiseerde informatie-ontsluiting.

2.5.2 Gegevensbestanden

Het behoud van een gegevensbestand is gecompliceerd; doordat standaard-formaten onbreken kan de toegankelijkheid op langere termijn niet worden gegarandeerd.

Gegevensbestanden kunnen een probleem vormen wanneer gegevens moeten worden teruggeroepen lang nadat de gebruikscyclus van de database is geëindigd. Er zijn momenteel weinig standaards voor de vorm van gegevensbestanden. Een dergelijke vorm wordt 'format' genoemd. Vaak zijn er maar twee opties:

- Het gegevensbestand kopiëren naar een structuur van een lagere orde: de database wordt als platte tekst of in ISAM-format opgenomen (indexed sequential access method, de gegevens zijn in

opeenvolgende volgorde opgeslagen, zonder de rekenformules waarmee onderlinge verbanden tussen cellen waren gelegd maar wel met scheidingstekens tussen de gegevensgroepen);

- De programmatuur bewaren, waarmee de database werd ontwikkeld en beheerd, inclusief de documentatie ervan (het database management systeem, de rekenformules etc.).

De eerste oplossing kan betekenen dat structurele elementen van het bestand verloren gaan. De tweede betekent dat niet alleen de applicatie dient te worden bewaard, maar ook het computersysteem waarop het heeft gedraaid, evenals de kennis hoe met de apparatuur en het pakket om te gaan. Niet gemakkelijk na verloop van jaren!

In veel gevallen is de database volledig geïntegreerd in een merk- of leveranciergebonden applicatie. Dit is vaak het geval met beheersprogramma's.

2.5.3 Aanschaf van hard- en software voor elektronische informatie

De aanschaf van apparatuur, programmatuur en diensten is een belangrijk aspect bij het gebruik en de archivering van elektronische gegevens. Er bestaat vandaag de dag een grote verscheidenheid aan hard- en softwareproducten en -diensten, gebaseerd op een variëteit van technologieën, zowel standaard- als specifieke toepassingen. Met de snelle vooruitgang van ICT, de informatie- en communicatietechnologieën is de zorg gerechtvaardigd dat de duurzame overleving op lange termijn in hoge mate afhankelijk is van de vorm, de structuur, waarin de informatie is bewaard. Het is mogelijk dat deze vorm door komende softwareproducten niet wordt ondersteund en dat de gegevens dus niet meer te lezen zijn.

EPHOS is het programma van de Europese Unie dat als doelstelling heeft om de inkopers binnen de publieke sector van dienst te zijn.

In de Europese Unie wordt het aanschafbeleid van de publieke sector geregeld door het Besluit van de Raad van Ministers nummer 87/95/EEC. Dit besluit verplicht alle inkoopende instanties om te refereren aan *de jure* standaards (dit zijn standaards die worden goedgekeurd door formele normalisatie-organisaties) in hun aanbestedingsprocedures. EPHOS (European Procurement Handbook for Open Systems) is een programma van de Europese Unie dat als doel heeft om aanbestedende instanties binnen de publieke sector te helpen om het Raadsbesluit 87/95/EEC op een juiste manier toe te passen. Dit gebeurt door te voorzien in een serie handboeken die een strategisch aankoopadvies voor open systeem-technologie bevatten. EPHOS refereert waar mogelijk aan internationale standaards. De huidige EPHOS modules refereren uitsluitend aan *de jure* standaards, alhoewel sommige modules die voorbereid worden, ook ingaan op het gebruik van algemeen beschikbare specificaties.

Aanschaf volgens de jure standaards levert praktische problemen op doordat bepaalde producten niet beschikbaar zijn, door de kosten van de producten en doordat er in de praktijk de facto standaards worden gehanteerd die afwijken van de de jure standaards.

De ervaring heeft geleerd dat in sommige gebieden van informatie en communicatie, in het bijzonder binnen de sector van de informatietechnologie, de voorwaarden van Raadsbesluit 87/95/EEC op praktische problemen stuiten. Dit komt doordat:

- producten die *de jure* standaards onderschrijven, nog niet leverbaar zijn;
- de producten die de standaards onderschrijven, erg duur zijn;
- er gevaar bestaat dat de standaards in de toekomst buiten gebruik raken en dus niet ondersteund worden door toekomstige producten;
- er elkaar beconcurrerende *de facto* standaards ontstaan die door de markt worden gedictieerd, welke worden geïmplementeerd in producten, terwijl de producten zelf breed worden gebruikt (bij voorbeeld Internet specificaties).

Welke standaard dan ook: eerlijkheidshalve moet worden gezegd dat het probleem van het eventueel gebruik van elektronische gegevens in de toekomst zich zowel kan voordoen bij producten die gebaseerd zijn op de facto standaards zowel als op de jure standaards. Hierover zijn de experts het echter nog niet eens.

Aanschaf van producten: *stabiele en open standaards als basis zijn een belangrijk criterium voor het evalueren van producten.*

Het bovenstaande suggereert dat het openbare aankoopbeleid eerder op producten dan op standaards gebaseerd zou moeten zijn. Gebruikers die machineleesbare gegevens produceren of onderhouden moeten zich ervan verzekeren dat ze alle hardware, programmatuur en programma-documentatie hebben die in de toekomst nodig zijn om de gegevens en documenten die door de applicatie zijn geproduceerd, terug te halen. Het is echter niet altijd praktisch of realistisch om over een langere periode alle apparatuur, software en documentatie te bewaren. Het zal duidelijk zijn dat het voldoen aan stabiele en open standaards een belangrijk criterium is voor de evaluatie van producten.

Gebruikers dienen duidelijk te zien welke standaards door IT-producten worden ondersteund, zoals een vaststelling van de stabiliteit en het open karakter van de standaards. Dit leidt tot de noodzaak van een lange-termijn aanschafpolitiek. Hiervoor horen algemene regels te worden ontwikkeld.

Documentatie van de vorm: *Om te zorgen dat het beheer en gebruik van gegevens over langere termijn gewaarborgd blijft is het aan te bevelen om de leverancier te verplichten alles te doen wat nodig voor het in stand houden van gegevens, die worden gegenereerd door een bepaalde, aan te schaffen, applicatie.*

Optie

Clausule in aankoopcontracten

Zolang het nog niet zover is, adviseert de Commissie om de volgende standaard-clausule op te nemen in calls for tenders:

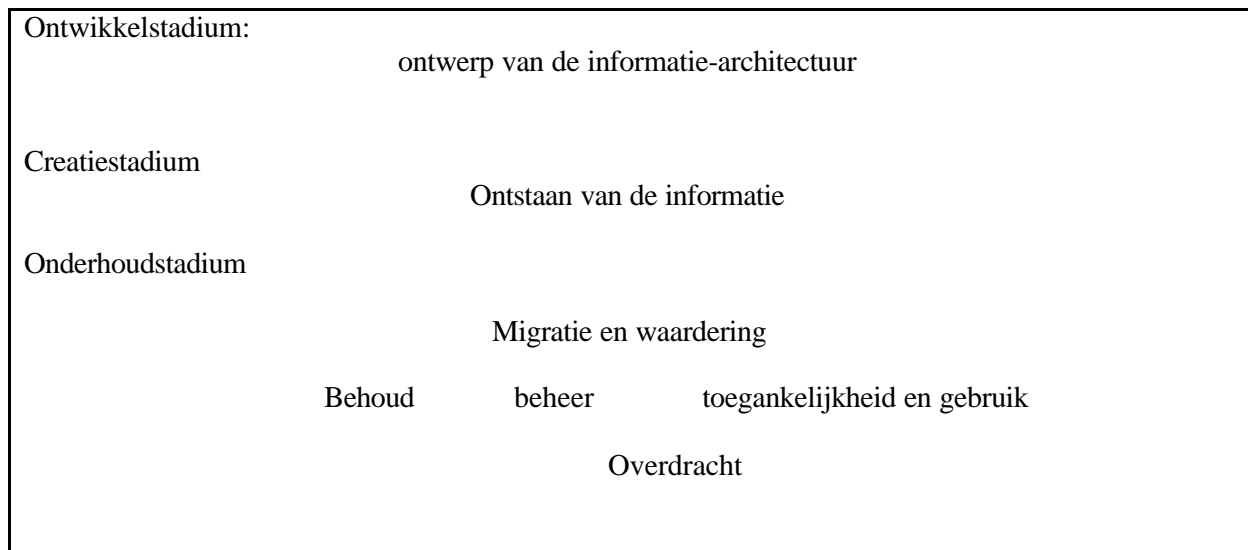
"Om het behoud van en de toegang tot de huidige gegevens voor de lange termijn te garanderen, zal de leverancier alle hardware, software en documentatie leveren die noodzakelijk is om de gegevens, die via deze applicatie worden aangemaakt, te kunnen herstellen, converteren en om deze te exporteren naar andere omgevingen en formats.

3. DE INFORMATIELEVENSCYCLUS EN HET TOEKENNEN VAN VERANTWOORDELIJKHEDEN

3.1 Overzicht van de levenscyclus

Elektronische informatie gaat door een aantal stadia gedurende de levenscyclus. Voor de verschillende fasen kunnen verschillende taken worden vastgesteld. Het is belangrijk om regels en standaarden vast te leggen, waardoor de coördinatie tussen iedereen die met bepaalde taken is belast wordt geborgd.

Levenscyclus: een aantal stadia die zo veel mogelijk moeten worden gecoördineerd.



Figuur 2: de levenscyclus van elektronische informatie

Er zijn verschillende onderhoudsstadia waarin de verantwoordelijkheid voor elektronische informatie wijzigt, soms overgaat van de ene organisatie op de andere.

Hoofdstuk 4 beschrijft de taken die voortvloeien uit de creatie en het beheer van elektronische informatie en geeft enkele suggesties om te komen tot een samenhangende strategie.

Hoofdstuk 5 bevat aanbevelingen voor het beheer van elektronische informatie (vooral bij gebruik van verschillende typen media en bestandsformaten).

Hoofdstuk 6 behandelt enkele manieren om de toegang en het gebruik van elektronische informatie te organiseren, waarbij rekening wordt gehouden met de organisatietechnische en veiligheidsaspecten en de standaards voor het onderling uitwisselen van gegevens.

3.2 Het definiëren van verantwoordelijkheden voor elk stadium

Definiëren van verantwoordelijkheden. Welke keuze een organisatie ook maakt, de verantwoordelijkheden voor elk stadium dienen helder te worden gedefinieerd.

Documenten en gegevensbestanden zijn van vitaal belang voor het openbaar bestuur, zoals voor elke andere organisatie. Informatie is de brandstof waarop de organisatie werkt. Een van de belangrijkste succesfactoren is de mogelijkheid om korte-, middenlange- en lange termijn strategieën voor de bewerking, de opslag en het toegankelijk maken van informatie te definiëren. Dat is beter dan te vertrouwen op traditioneel gegroeide verhoudingen tussen afdelingen en organisaties.

De manier waarop verantwoordelijkheden worden toegekend zijn van land tot land -en van organisatie tot organisatie- verschillend. Welke opties ook worden gekozen, het is van vitaal belang om verantwoordelijkheden helder vast te stellen als onderdeel van een totaalstrategie.

Voorbeeld

Toewijzen van verantwoordelijkheden

Organisaties op nationaal en Europees niveau hebben verschillende keuzes gemaakt aangaande de manier waarop verantwoordelijkheden worden toegekend.

- Unesco heeft aanbevolen om 'de expertise van de archivaris te gebruiken bij het waarderen en selecteren van informatie, maar dat het fysieke beheer van het record aan de bron moet plaatsvinden, bij de herkomst van het record' (Unesco, RAMP studie).
- In de Verenigde Staten worden records beheerd door de National Archives en Records Administration (NARA) in meer dan 20 locaties, verspreid over de VS. Tevens zijn er verwante archieven die door de NARA worden gecontroleerd. De records vallen onder verantwoordelijkheid van de NARA, hoewel de aangesloten archieven verantwoordelijk zijn voor hun behoud, beheer en toegankelijkheid.
- In de Europese Commissie zijn de archiefdiensten/registratieraafdelingen in sommige Directoraten-Generaal verantwoordelijk voor de distributiestroken voor de post, terwijl andere Directoraten de post op een andere manier afhandelen.

Het op een efficiënte wijze toekennen van verantwoordelijkheden betekent dat rekening wordt gehouden met zowel de aard van elke taak als de cultuur en het kennisniveau van de organisatie. De betrokken afdelingen en de archivarissen zouden vanaf de start betrokken moeten zijn bij de opzet van het records management.

Als elektronische informatie door verschillende organisaties wordt geproduceerd en onderhouden, kunnen verantwoordelijkheden worden vastgesteld in dialoog tussen die organisaties en hun archiefdiensten.

Wat is de rol van de archivarissen? Zij hebben waardevolle vaardigheden in het waarderen en selecteren van records en kunnen een bruikbare inbreng leveren vanaf het prille begin van de levenscyclus van het record.

De archivaris is verantwoordelijk voor het beheer van de records. Hij heeft ook bekwaamheden om de waarde van een record te bepalen. Zijn rol verandert van de passieve ontvangst van de records die het einde van hun actieve levensstadium hebben bereikt, naar een actieve betrokkenheid vanaf de start van de levenscyclus.

Het is noodzakelijk dat archivaris en de overige functionarissen die betrokken zijn bij de elektronische informatie bij de overheid en de private sector nauwer samenwerken.

Verskillende soorten informatie vragen om verschillende deskundigheid

- *De actuele informatie wordt vaak beheerd door de auteur van het document of de beheerder van het gegevensbestand*
- *De archivaris heeft expertise in het beheer van records*
- *De bibliothecaris heeft expertise in het beheer van publicaties*

Drie typen informatie Voor gevorderden

Er zijn drie typen informatie te onderscheiden. De verschillende typen informatie vragen om een verschillende expertise.

- a. Dynamische informatie. Deze wordt door de organisatie geproduceerd en niet geregistreerd. Dit kan een kortstondig karakter hebben, waarbij registratie niet nodig is. Ook kan het document geregistreerd worden, waardoor het tot een record wordt.
- b. De records: archivari² zijn de experts om informatie gedurende de gehele levenscyclus te beheren. Informatie kan direct worden geproduceerd als een record, maar ook kan dynamische informatie worden geregistreerd en zodoende een record worden. De archivaris noemt dit het 'vastleggen' van een record. Op dat moment wordt de informatie 'bevroren' en worden de contextuele gegevens vastgelegd om de toegankelijkheid en het gebruik van records mogelijk te maken.
- c. De publicaties: bibliothecarissen zijn de experts om deze informatie te behandelen.

De multidisciplinaire richtlijnen presenteren elektronische informatie en records vanuit het gezichtspunt van het bedrijfsproces waarin verschillende disciplines een rol spelen. De ICA-gids over elektronische records geeft meer het gezichtspunt van de archivaris weer. Zie hiervoor annex 8.9

² Waar hier steeds uitsluitend archivariissen worden genoemd, wagen wij ook de documentaire informatieverzorgers als experts op het gebied van het dynamische archiefbeheer te benoemen.

4.ONTWERP, CREATIE EN ONDERHOUD VAN ELEKTRONISCHE INFORMATIE

Elk van de fasen van de informatie-levenscyclus heeft zijn eigen karakteristieke eigenschappen. Dit hoofdstuk behandelt het ontwerp, de creatie en het onderhoud van elektronische informatie.

4.1 Het ontwerpstadium

Strategisch team. *De beste resultaten worden behaald door bij aanvang een multidisciplinair team samen te stellen dat een strategie voor elektronische informatie definieert en op de uitvoering hiervan toezicht houdt.*

De beste manier om het beheer van elektronische informatie te optimaliseren is om al aan het begin een samenhangende, volledige strategie te definiëren. Als oplossing kan een multidisciplinair team worden opgezet om de strategie vast te stellen en te bewaken. Alhoewel deze methode nog niet wijd verbreid is, kan deze benadering worden aangewezen als 'beste werkwijze'.

Opties

Hoe een strategie vast te stellen

De verantwoordelijkheden van het multidisciplinaire team dat zich bezighoudt met elektronische informatie zouden het volgende kunnen inhouden:

- Vaststellen van de gebruikseisen (inclusief de juridische aspecten)
- Vaststellen van de belangrijke records
- Regels voor een efficiënt classificatieplan vaststellen;
- Standaarden en specificaties vaststellen om er zeker van te zijn dat de gegevens onafhankelijk zijn van de media waarop ze zijn opgeslagen en om hun duurzaamheid te garanderen;
- Een selectielijst vervaardigen;
- Per taakgebied in de informatie-levenscyclus aangeven wie verantwoordelijk is;
- Een trainings- en bewustwordingsproces starten in de organisatie ³
- De invoering van nieuwe systemen bewaken en volgen ⁴

De juridische aspecten zijn één van de belangrijkste onderwerpen die in gedachte moeten worden gehouden bij elke elektronische informatiestrategie.

Voor het openbaar bestuur behoren de juridische aspecten tot de belangrijkste onderwerpen die in gedachten moeten worden gehouden. Sommige vragen aangaande elektronische informatie zijn nog onvoldoende onderzocht. Daarom is grote zorgvuldigheid geboden (zie hiervoor de uitstaande onderwerpen in annex 8.2).

4.2 De creatie van elektronische informatie

Het is belangrijk om elk van de taakgebieden binnen de informatielevenscyclus in een zo vroeg mogelijk stadium te overwegen. Daarom zouden bij de creatie van elektronische informatie algemene richtlijnen moeten worden toegepast. Dit zal de communicatie tussen de mensen die de informatie creëren en hen die de later de informatie moeten bewerken en beheren vergemakkelijken.

³ In deopleiding voor archivariissen en ambtenaren in de openbare sector dient elektronisch informatiebeheer te zijn opgenomen.

⁴ Zie hiervoor ook annex 8.4 van de checklist voor een strategie voor elektronische informatie

Regels voor de creatie van elektronische informatie. *Het in acht nemen van bepaalde regels zal het beheer in latere stadia van de informatielevenscyclus vergemakkelijken.*

De creatie van elektronische informatie

Opties

De volgende regels gelden bij de creatie van elektronische informatie:

- Elk document of elk gegevensbestand dient helder te worden geïdentificeerd door de organisatie die verantwoordelijk is voor het elektronische informatiebeheer. Dit betekent dat de relevante gegevens over het document of de gegevensbank op een unieke wijze worden vastgelegd, zodat het document zich onderscheidt van andere.
- Van alle elektronische informatie dient achtergronddocumentatie te worden geleverd. Deze wordt gedurende alle stadia van de informatielevenscyclus bewaard door de persoon die verantwoordelijk is voor het beheer van de elektronische informatie
- Er moeten procedures worden vastgesteld voor het bewerken van elektronische informatie van onbekende oorsprong.
- Gegevens mogen niet worden vernietigd of gewijzigd zonder dat een voorafgaande goedkeuringsprocedure is doorlopen. Dit verzekert dat de context behouden blijft. Vernietiging of actualisering houdt alle activiteiten in waarbij de mogelijkheid om gegevens te combineren, herkennen, terug te zoeken of te identificeren verdwijnt.

Informatie kan oorspronkelijk op papier zijn geproduceerd en daarna gedigitaliseerd, maar kan ook direct worden gecreëerd in digitale vorm (via tekstverwerking, een database programma, elektronische post etc.). Tijdens al deze momenten zou het in een gestandaardiseerd formaat moeten worden geproduceerd of ernaar moeten worden overgezet (tenzij hierdoor de authenticiteit in gevaar wordt gebracht). Hoofdstuk 5.2 geeft details aan van de meest geschikte formats.

Elektronische informatie of papieren documenten?

Opties

- Doorgaans wordt informatie op papier opgeslagen. Elektronische zoekinstrumenten kunnen helpen bij het beheer en terugzoeken van papieren bestanden.
- Elektronische informatie wordt op elektronische media vastgelegd en bewaard.

Papieren documenten en de contextuele informatie die zij bevatten kunnen worden gescand. De hierdoor verkregen elektronische records vergemakkelijken het zoeken en raadplegen. De contextuele informatie van een document wordt gewoonlijk aangeduid met de term metadata.

4.3 Integratie, conversie en waardering⁵ van elektronische informatie

4.3.1 Integratie van informatie

Het reorganiseren, selecteren, samenvoegen of afleiden van gegevens uit een bestand om hierdoor een compactere gegevensverzameling te ontwikkelen is een delicate taak. Het recht op privacy dient te worden gewaarborgd (zie sectie 6.3 over anonimiteit).

Soms kan het noodzakelijk zijn om gegevens van verschillende organisaties samen te brengen. Toekomstige onderzoekers die gebruik maken van openbare informatie zijn vooral geïnteresseerd in het zoeken per onderwerp, per organisatie of per bron of afkomst.

Documentatie, structuur en een samenhangend geheel van tekst en gegevens

Er zijn altijd meerdere mogelijkheden om informatie te groeperen. Dient een rapport te worden samengebracht met de verschillende geamendeerde versies? Welke documenten uit een bepaald gegevensbestand dienen te worden bewaard? Welke oplossing ook wordt gekozen: de aanvullende informatie van het document of de database dient de eigen documentatie te bevatten (inclusief de metadata) en structuur (als afzonderlijk bestand of opgenomen in het hoofdbestand).

Documenten gecreëerd vanuit gegevensbestanden

Voorbeeld

Als een document dynamisch vanuit een gegevensbestand wordt vervaardigd, dan zijn er verschillende oplossingen voor de opslag over een lange termijn:

- Het gegevensbestand 'bevrozen' met een specifieke vraagstelling en zo een traditioneel document produceren, waarin mogelijk enkele combinaties verloren zijn;
- Als de applicatie een volledig audit trail heeft kan de volledige database worden geëxporteerd als het systeem niet meer gebruikt wordt;
- Het vastleggen van de database en de applicatie die het document genereert, de apparatuur waarop de applicatie heeft gedraaid en waarmee het document is vervaardigd.

Wanneer we te maken hebben met samengestelde documenten zijn de oplossingen niet altijd vanzelfsprekend. Samengestelde documenten zijn complexer dan enkelvoudige documenten (denk bij voorbeeld bij een samengesteld document aan informatie uit een tekstverwerkings- een rekenschema- een presentatie- en een databasepakket).

De omvang van de informatie-eenheid kan variëren, afhankelijk van de keuzes die worden gemaakt. Het hoge aantal referenties en hyperlinks tussen documenten zou tegenwoordig betekenen dat alle gegevens op deze aardbol zouden moeten worden bewaard als één groot record! Anderzijds moeten informatie-eenheden een redelijke omvang hebben om bruikbaar te kunnen zijn.

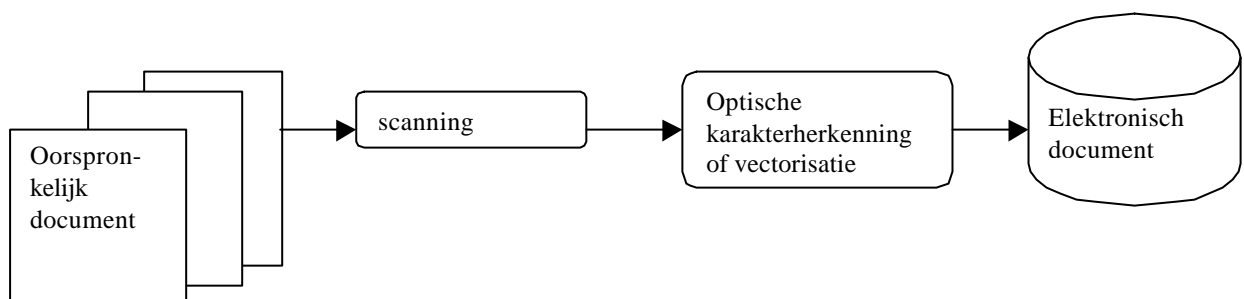
⁵ De term "appraisal" wordt hier vertaald met "waardering". Hiermee wordt het proces bedoeld waarin de informatie naar waarde wordt beoordeeld en waarbij een beslissing wordt genomen over de bewaring of de vernietiging van de informatie. Deze vernietiging is aan bepaalde termijnen gebonden. Deze termijnen kunnen voorgeschreven zijn in de wetgeving, maar kunnen ook worden bepaald door veronderstellingen omtrent het toekomstige gebruik door bijvoorbeeld historici.

4.3.2 Conversie van informatie

- *Conversie van papier naar een elektronische vorm:*
- *Eenvoudige scanning*
- *Scanning en codering*

Er zijn twee oplossingen om een papieren document naar een digitale vorm te converteren:

- Het document scannen om hierdoor een elektronische afbeelding te verkrijgen (het 'image', dit wordt behandeld als een digitale fotokopie en kan niet verder worden bewerkt);
- het document scannen en dan coderen naar een elektronische vorm (door bij voorbeeld OCR-technieken te gebruiken, Optical Character Recognition of grafische vectorisatie, zoals beschreven in sectie 4.3.4).



OCR-toepassen of niet?

Opties

Het simpelweg scannen van een tekst is een ongecompliceerde, eenvoudige oplossing en vraagt geen verdere bewerking. Optical character recognition wel.

Bij toepassing van OCR neemt het digitale document slechts enkele kilobytes aan ruimte in beslag, doordat het mogelijk is om compressie toe te passen. Bij het scannen zonder toepassing van OCR zal een document meer dan 50 kilobyte per pagina omvatten, omdat het wordt opgenomen als foto-afbeelding. In aanvulling hierop is het wel veel eenvoudiger om een OCR-tekst -die uit karakters bestaat- te wijzigen dan een image van die tekst. Ook het indexerend is veel eenvoudiger.

Een andere optie bestaat uit een combinatie van deze twee technieken, waarbij het 'ruwe' image wordt opgeslagen en daarna OCR-techniek wordt gebruikt voor de tekstgedeelten. Details over formats die bruikbaar zijn voor bewaring op lange termijn worden gegeven in hoofdstuk 5 (korte en lange-termijn bewaring van elektronische informatie), terwijl het vraagstuk van het benaderen van de gegevens wordt behandeld in hoofdstuk 6 (het toegang verkrijgen en verspreiden van informatie).

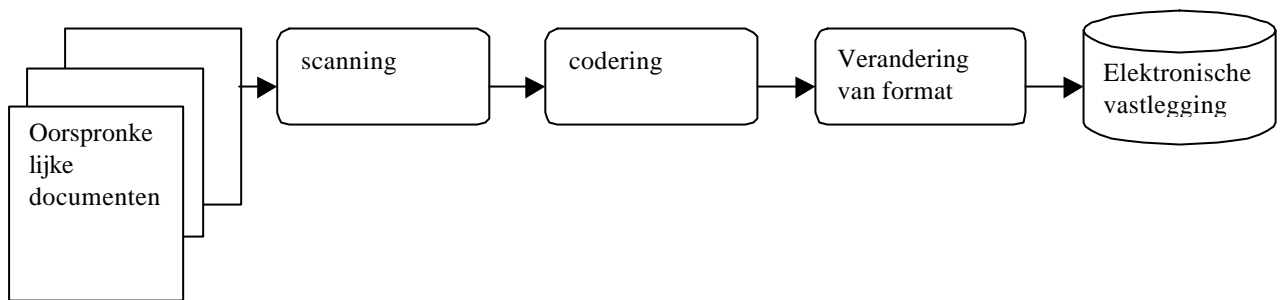
Digitale conversie: *conversie van de ene naar een ander digitale vorm is mogelijk met elk type record.*

Een derde type conversie houdt in het veranderen van het ene naar het andere digitale format. Dit is alleen raadzaam wanneer het nieuwe formaat beter gestandaardiseerd is en hierdoor langduriger

houdbaarheid garandeert (wanneer het bij voorbeeld structuur geeft aan een platte tekst). Het kan ook worden toegepast als derde stadium bij de conversie van een papieren document om een elektronisch record te verkrijgen met een "hogere" structuur (bij voorbeeld een document met een expliciete structuur of een gegevensbestand).

Alhoewel het omzetten van papieren records naar een digitaal formaat hoofdzakelijk documenten betreft, kan in dit derde stadium aan het elektronisch document een database-structuur worden gegeven. Zie hiervoor sectie 4.3.3

Figuur 4: conversie van een digitaal formaat



4.3.3 Van papier of microfilm naar een gescand image.

Scanner en fax

Basisbegrippen

Het scannen van een document maakt het mogelijk om de inhoud van papieren pagina's te bewaren als een computerbestand. Dit bestand bevat dan de digitale kopie van het originele document, gekopieerd met een bepaalde resolutie.

De resolutie van een scanner wordt gemeten in 'dots per inch (dpi)'. Scanners kunnen gemakkelijk een resolutie aan van 300 of 600 dpi kleurendruk. Deze dpi's zijn te beschouwen als het raster, waarmee het beeld wordt opgebouwd, zoals bij voorbeeld gebeurt met de puntjes van een televisiescherm.

Een faxapparaat bestaat uit een scanner, een systeem om gegevens te verzenden via een telefoonlijn en een printer op de bestemming. De scanresolutie van een fax is veel lager (tussen de 100 en 200 dpi).

Scanning Om een bruikbaar resultaat te verkrijgen, dienen enkele regels in acht te worden genomen.

Bij het scannen van documenten dienen een aantal basisregels te worden gevolgd. De kwaliteit van het resultaat is hier hoofdzakelijk van afhankelijk.

Tekst zou niet opgeslagen moeten worden als een image, behalve wanneer het gaat om een kort bericht of wanneer optische karakterherkenning onmogelijk is, bij voorbeeld bij handtekeningen. In andere gevallen is het aan te bevelen om de tekst via OCR te scannen.

Het scannen van een document

Opties

- Alle onderdelen die nodig zijn voor het uiteindelijke record (denk bij voorbeeld aan een dossier) moeten worden overgezet naar hetzelfde medium (bij voorbeeld: dezelfde schijf).
- Verbindingen (links) die bestaan tussen het record en het archiefsysteem waar het deel van uitmaakt moeten worden bewaard (bij voorbeeld andere referentierecords).
- De kwaliteit van het origineel (contrast, omvang van de karakters) moet voldoende zijn om na conversie naar een digitaal medium de best mogelijke afdruk te maken
- Voordat documenten worden gescand, dient een test te worden uitgevoerd
- Zelfs wanneer scanning is uitbesteed, dient de uitvoerende organisatie de digitale documenten te checken op kwaliteit en compleetheid

Bij het ontwerp van formulieren dient al rekening te worden gehouden met een mogelijke conversie naar elektronisch formaat (lettergrootte, positie van velden, lijnenpatronen).

4.3.4 Van gescand image naar gecodeerd formaat

Optische karakterherkenning (OCR, Optical Character Recognition) maakt het mogelijk om van tekstimages tekstbestanden te maken.

Optische karakterherkenning

Basisbegrippen

OCR zorgt ervoor dat de computer een tekst kan lezen.

OCR-programmatuur bewerkt een bestand dat een afbeelding bevat van de tekst die gelezen dient te worden, bij voorbeeld een bestand dat door een scanner is vervaardigd. De tekstkarakters worden herkend en er wordt een tekstbestand aangemaakt, dat door elk tekstverwerkingspakket kan worden bewerkt.

De tekstherkenning is niet perfect. Als de computer een karakter niet herkent, wordt het gemarkeerd zodat een menselijke operator dit kan herkennen. Het gebeurt soms dat de computer een karakter verkeerd inleest. Automatische correctiehulpmiddelen kunnen de correctie verbeteren, maar een mens zal altijd een controle moeten uitvoeren nadat optische karakterherkenning heeft plaatsgevonden. Deze controle kan ook met behulp van software worden uitgevoerd.

De herkenningratio voor een schoon image van een getypte tekst is ongeveer 95% (dit houdt in: een of twee fouten per tekstregel). Afhankelijk van de taal of het script van de bron kan deze ratio verschillen.

Nadat een document dat tekst en beeld bevat is gescand, kan de tekst met OCR-techniek worden bewerkt. Hierdoor is het echter mogelijk om de tekst te manipuleren, om gedeelten ervan opnieuw te gebruiken of om de tekst te indexeren, waardoor het terugvinden en raadplegen wordt vereenvoudigd.

Faxberichten

Voor gevorderden

Vaak is het moeilijk om bij faxen optische karakterherkenning te gebruiken. Ze hebben een zwakke resolutie en de afdruk op papier is van slechte kwaliteit.

Een oplossing kan zijn om de faxen in hun oorspronkelijke digitale vorm op te slaan (ITU-T groep III of IV). Alhoewel het hierdoor niet mogelijk is de tekst opnieuw te gebruiken, geeft het wel een gecompriëerde versie van beelden en tekst, waardoor minder geheugenruimte wordt gebruikt.

Vectorisering werkt alleen bij kaarten, grafieken en andere afbeeldingen die zijn samengesteld uit lijn- en figuurcomponenten.

Kaarten en grafieken in het originele document kunnen worden gevectoriseerd om ruimte te besparen. In sommige gevallen -bij voorbeeld bij technische tekeningen- heeft vectorisering aanvullende voordelen.

Vectorgrafieken

Basisbegrippen

Vectorgrafieken zijn gemaakt van eenvoudige elementen (rechte en gebogen lijnen, rechthoeken, ovaal en cirkels etc.). In plaats van een afbeelding te bewaren die is samengesteld uit puntjes, zoals bij een foto, herkent vectorisering de basiselementen waaruit het grafische beeld is opgebouwd.

Zoals bij OCR heeft vectorisering twee voordelen:

- Een vectorgrafiek gebruikt minder geheugen dan een fotografische afbeelding (een 'bitmap') van een tekening. Alleen het soort element en de plaatsingscoördinaten van begin en eind worden opgeslagen;
- Het is eenvoudiger om de grafische afbeelding te wijzigen of een gedeelte eruit te hergebruiken.

Coderen: enkele eenvoudige regels die zorgen voor een kwalitatief goed resultaat

Wanneer 'ruwe' gegevens worden gecodeerd met behulp van OCR-techniek of grafische vectorisering, dan moeten bepaalde basisregels worden gevolgd. Hierdoor wordt het gemakkelijker om de elektronische informatie in alle latere stadia van de informatielevenscyclus te bewerken.

Het coderen van tekst en grafieken

Optie

Delen van een document die niet gecodeerd zijn (zoals foto's of afbeeldingen) zouden afzonderlijk van de tekst moeten worden opgeslagen in een ander bruikbare vorm.

4.3.5 Van het ene digitale formaat naar het andere

Er zijn twee momenten waarop het nuttig is om informatie van het ene digitale formaat naar het andere over te zetten:

- Bij het overgaan van een formaat naar een meer duurzaam formaat, waardoor het document voor een langere termijn leesbaar blijft en de raadpleging wordt vereenvoudigd;
- Bij het toevoegen van structuur aan een 'platte' tekst, om hierdoor een gestructureerd document of gestructureerde database te verkrijgen, die gemakkelijker te raadplegen is.

Conversie naar een standaard vorm. Het belangrijkste voordeel van het converteren van het format van elektronische informatie is de garantie van toegankelijkheid en leesbaarheid.

Wanneer elektronische informatie wordt overgezet van het ene naar het andere formaat, dient gewaakt te worden voor het verlies van gegevens. De eigenschappen van programmatuurgebonden en standaard-formats corresponderen niet altijd met elkaar. De twee onderstaande voorbeelden maken dit duidelijk.

Standaardformats voor het behoud en toegankelijk houden van elektronische informatie worden behandeld in de hoofdstukken 5 en 6.

verlies van informatie

voorbeeld: verlies van informatie uit een document

Stel dat een document de volgende zin bevat: "Aanbevolen wordt om het project te beëindigen!". De voetnoot aan het einde van de pagina zegt echter "1 tenzij de aanvankelijk toegezegde financiering alsnog wordt toegekend".

Het document is in een bepaald tekstverwerkings-bestandsformaat opgesteld en wordt nu opgeslagen in een nieuw standaardformaat. Als er geen voorzorgsmaatregelen worden genomen en het conversieproces negeert de voetnoten, dan kan de strekking van de aanbeveling volstrekt anders worden!

Voorbeeld: verlies van informatie uit een gegevensbank

Stel dat een database, gemaakt met een bepaald rekenprogramma, wordt opgeslagen in een formaat dat specifiek is voor dat programma. Omdat het programma uitsluitend werkt onder een operating systeem op een specifieke computer, wordt besloten om de data op te slaan in een onafhankelijk formaat. Als er geen voorzorgen worden genomen, verdwijnen bepaalde verbanden tussen cellen. Formules kunnen verdwijnen, getallen verliezen hun waarde en de database wordt onbetrouwbaar.

Het structureren van informatie. Aan platte tekst kan structuur worden toegevoegd om zo een gestructureerd document of een gestructureerde database te verkrijgen.

Het tweede geval waarin het zinvol is om digitale informatie van het ene formaat naar het andere te converteren is om structuur toe te voegen aan een ongestructureerde, platte tekst. Dit is vaak het geval met gescande tekst wanneer de toevoeging van gestructureerde kenmerken (inhoud, indexgegevens etc.) de raadpleging vereenvoudigt. Er bestaan programma's die helpen om een document te structureren, bij voorbeeld door een standaard-opmaak voor te schrijven. Er worden dan verschillende soorten tekst (tekenreeksen, tekstfonts) gebruikt voor verschillende niveaus in de tekstindeling.

Bij het herstructureren van gescande documenten is het belangrijk om elementen te combineren die zijn onderbroken door pagina- of kolommenindeling, door het invoegen van een illustratie, tabel enz.

Datzelfde geldt voor databases die zijn opgezet als platte tekstbestanden (een directory of formulieren die zijn gescand, oudere databases die zijn opgeslagen in een eenvoudig formaat). De structuur ligt hier vaak onveranderbaar vast en velden zijn van elkaar gescheiden door scheidingstekens, bij voorbeeld tabs of kolomindeling of door de positie van de velden op de pagina (kolom, positie van een veld op een formulier).

Het laatste voorbeeld is vooral belangrijk omdat er geen *de jure* standaard is voor gestructureerde gegevensformaten en gegevens dikwijls moeten worden opgeslagen in een eenvoudig formaat (zie sectie 5.2.4 voor gegevensformaten).

4.3.6 De waardering van elektronische informatie

De controle van records. De meeste problemen doen zich voor gedurende de conversie of het verzenden van informatie.

Het bewaren van informatie is zinloos, tenzij de informatie kan worden geraadpleegd als hieraan behoefte is. Wanneer gegevens als een record in het systeem moeten worden opgenomen om een activiteit te bewijzen, is de waardering van die informatie belangrijk. Authenticiteit, integriteit en waarde moeten zorgvuldig worden afgewogen.

Het selecteren voor vernietiging van records die geen verdere waarde hebben is een belangrijke taak in het waarderingsproces.

Selectie

Voor gevorderden

Een belangrijke taak bij de selectie van records is het identificeren van records die voor vernietiging in aanmerking komen. Het is niet zinvol deze informatie te bewaren. Records dienen direct wanneer bekend is dat zij niet meer van nut zijn en geen verdere waarde meer hebben, te worden vernietigd. Dit houdt de waardevolle records beter toegankelijk.

Ditzelfde kan gelden voor items in databases, als deze items ontleend zijn aan andere gegevens of slechts gedurende een beperkte periode waarde hebben. Het verwijderen van items en het daarna opnieuw indexeren van de database kan leiden tot een snellere toegang en bewerking. De kennis van de archivaris om records te selecteren voor bewaring en vernietiging zou hierbij moeten worden gebruikt.

De beste manier om tot selectie van records over te gaan

Voorbeeld

De Handleiding Opbouw Bestuursarchief Gemeenten geeft richtlijnen om de selectie van archiefbescheiden eenvoudiger te doen verlopen. De Handleiding is gebaseerd op de Vernietigingslijst voor archiefbescheiden van (inter)gemeentelijke organen. Het systeem bevat drie belangrijke uitgangspunten:

1. Selectie aan de basis. Documenten met een kortstondige informatiewaarde en documenten die de status van archiefbescheiden niet hebben, dienen te worden opgenomen in een zogenaamde voorselectie voor onbelangrijke stukken.
2. Stukken betreffende de vaststelling van beleid en stukken betreffende de uitvoering dienen in afzonderlijke dossiers te worden ondergebracht
3. Aan de dossiers worden zo mogelijk al direct bij de aanmaak vernietigingstermijnen gekoppeld, zodat vanaf het begin bekend is hoe lang bepaalde informatie voor bewaring in aanmerking komt en na welke termijn deze dus wordt vernietigd.
4. De lijst van stukken met een beperkte informatiewaarde bevat bescheiden die doorgaans na een periode van een jaar kunnen worden vernietigd. :

De Selectielijst maakt het mogelijk om duplicaten, stukken met een kortstondige informatiewaarde en routinestukken die niet bedoeld zijn als een formele vastlegging in een werkproces binnen een korte termijn te vernietigen.

Controle van gegevens

Opties

Op elektronische gegevens kunnen verschillende controles worden uitgevoerd:

- Voldoet het bestand aan de vastgestelde standaarden;
- Is het beschermd tegen overschrijven; (write protection);
- Zijn de media leesbaar;
- Komen de gegevens overeen met de gegevens die in de documentatie zijn opgenomen (bij voorbeeld een test door enkele records af te drukken);
- Is de inhoud compleet

Als regel zijn tests niet uitputtend. Soms komen problemen pas naar boven wanneer via de gewone weg toegang tot een record wordt verkregen (bij voorbeeld gedurende het gebruik in de levenscyclus). Wanneer de informatie door een andere organisatie of -onderdeel werd verstrekt, kan gevraagd worden om de oorspronkelijke informatie te reconstrueren. De meeste problemen doen zich eerder voor bij de conversie, zoals beschreven in deze sectie, of het verzenden naar een andere organisatie (zie sectie 4.5) dan bij het bewaren van de informatie.

Wat te doen wanneer een applicatie gegevens genereert in een leveranciersafhankelijk formaat

Als de gebruikte applicatie gegevens genereert in een merkgebonden formaat dan kan het noodzakelijk zijn om -teneinde de informatie te bewaren- het totale systeem te bewaren. Dit houdt in: de applicatie, het IT-platform, de documentatie en mogelijk zelfs de kennis van het personeel dat destijds de applicatie en het platform heeft gebruikt. De kosten van een dergelijke oplossing dienen zorgvuldig afgewogen te worden tegen de intrinsieke waarde van de informatie.

Als alternatief zou de informatie kunnen worden overgebracht naar een formaat van een lagere orde, als platte tekst of als plat, volgordelijk bestand. Er kan dan wel informatieverlies optreden, vooral de eigenschappen van de structuur van documenten raken snel verloren.

Een derde mogelijkheid is om informatie te converteren naar een standaardformaat. Het ontwikkelen of aanschaffen van een conversiehulpmiddel brengt uiteraard kosten met zich mee. Als de gehele structuur van de gegevens ook noodzakelijkerwijze dient te worden opgeslagen, is de oplossing echter het overwegen waard. Natuurlijk dient het formaat waarnaar de gegevens moeten worden overgebracht zoveel mogelijk overeen te komen met maatschappelijk geaccepteerde standaards om de duurzaamheid van de informatie te garanderen.

Alhoewel de kosten van het converteren van elektronische informatie naar een nieuw formaat hoog kunnen zijn, kan het achterwege laten hiervan leiden tot latere hogere kosten.

***Opties** Er zijn verschillende oplossingen voor de opslag op lange termijn van gegevens. Welke de beste is, hangt af van verschillende criteria.*

Leveranciersgebonden formats

Opties

Houdt de applicatie rekening met standaard formats? JA: gebruik het standaard formaat in plaats van het eigen formaat

Indien nee:

Bestaat een conversiehulpmiddel? JA: gebruik dit hulpmiddel om het formaat te converteren

Indien nee:

Is verlies van de structuur acceptabel tot op zekere hoogte? JA: converteer naar een plat tekstformaat

Indien nee:

Is het goedkoper om een conversiehulpmiddel te ontwikkelen dan het oude systeem te behouden? JA Ontwikkel een hulpmiddel om naar een standaard formaat te converteren.

Indien NEE:

Behoud de applicatie en het platform met de records tot een oplossing is gevonden.

De kosten van het converteren van elektronische gegevens naar een nieuw formaat kunnen hoog zijn. Het niet converteren kost echter op de lange duur wellicht nog meer!

4.4 Beheer en de classificatie van elektronische gegevens

Beheer van elektronische informatie:

- *registratie van de elektronische informatie*
- *benoemen van verantwoordelijken*
- *coördineren van de bewerking*
- *klasseren van elektronische informatie*
- *beslissen over de overdracht (transfer)*

De doelstelling van het beheer van elektronische gegevens is om de betrouwbaarheid, authenticiteit, integriteit en toetsbaarheid van informatie gedurende de gehele levenscyclus te waarborgen. Hiertoe dient de context van de informatie goed te zijn gedefinieerd. Als de inhoud, context en structuur voldoende zijn om een bewijs te leveren van een bepaalde activiteit, dan wordt informatie een record.

Het beheer van de levenscyclus van elektronische informatie houdt ook andere taken in. De verantwoordelijkheid voor het beheer van een bepaald gegevensbestand kan van de ene organisatie naar een andere worden overgeheveld. Sectie 4.5 handelt over deze overdracht.

Het beheer van elektronische informatie houdt de volgende taken in:

- Het registreren van de momenten waarop en door wie elektronische informatie is gebruikt. Dit houdt in: het steeds bijhouden van het audit trail van de gegevens;
- Toewijzing wie verantwoordelijk is voor het beheer: wanneer meerdere organisaties of afdelingen zijn betrokken bij een bepaald proces, dan moet de elektronische informatie worden toegewezen aan de juiste organisatie of -afdeling en dient een beheerder te worden aangewezen;
- Follow-up: dit houdt de coördinatie in van verschillende stadia van bewerking (de ontvangst van elektronische informatie, conversie, het waarborgen van het behoud, gebruik, transport naar andere organisatie);
- Classificatie: het toekennen van zoek sleutels aan de hand waarvan het terugvinden van de specifieke elektronische informatie wordt vergemakkelijkt (zie hier onder);
- Besluit om informatie te verplaatsen naar een andere organisatie of organisatie-onderdeel.

Bij het beheer van elektronische informatie zijn steeds meer groepen mensen betrokken. Workflow-hulpmiddelen kunnen een goede dienst bewijzen om de uitwisseling binnen een bepaalde werkgroep te beheersen.

Classificatie van informatie: het ontsluiten van de informatie is één van de belangrijkste beheerstaken

Een van de belangrijkste taken, en wellicht de meest gecompliceerde, is het classificeren van informatie. Het classificatiesysteem moet zo helder zijn dat andere organisaties het begrijpen (vooral wanneer de verantwoordelijkheid voor elektronische informatie wordt overgeheveld naar een andere organisatie of organisatie-onderdeel).

De structuur van een document of gegevensbestand maakt het mogelijk om een specifiek informatie-item in een document of database terug te vinden, terwijl de classificatie het mogelijk maakt het document of de database terug te vinden tussen alle overige opgeslagen informatie. Het classificatieschema dient gebruikt te worden, onafhankelijk van de doelstelling van de informatie. Aanvullende criteria voor het indexeren van informatie kunnen worden toegevoegd om daarmee het specifieke karakter van elektronische informatie aan te geven.

Europese commissie: voorbeeld. Enkele onderdelen bij de Europese Commissie gebruiken een classificatieschema dat is gebaseerd op de taak van de organisatie, de (interne) administratie, het personeel en de financiën⁶.

Er kunnen verschillende criteria worden gebruikt voor het classificeren en indexeren van elektronische informatie. Deze criteria zijn vooral zinvol bij elektronische berichten.

Enkele voorbeelden van criteria zijn:

Type document

- Datum van productie en vervaldatum;
- Auteur (individueel of afdeling)
- Afzender, auteur
- Bestemming (individuele personen, afdelingen)
- Kopiehouders (individuele personen, afdelingen)
- Nummer van de elektronische boodschap, versienummer;
- Plaats van bewaring (datum, dossier, afdeling)
- Onderwerp
- Project of activiteit
- Trefwoorden
- Taal
- Aantal pagina's
- Status (officieel, onofficieel), vertrouwelijkheid
- Aangehechte documenten, verbanden met andere documenten
- Andere categorieën die kunnen worden gedefinieerd door de gebruiker.

Archivistische beschrijving van een record

Opties

De International Council on Archives heeft internationale standaards geproduceerd voor archiefdefinities (ISAD/G), die in overweging dienen te worden genomen wanneer een strategie voor het classificeren van documenten wordt ontwikkeld.

De identificatie van het archiefstuk wordt het "identificatiegebied" genoemd. Aanbevolen wordt om de wijze van identificatie niet gedurende de levenscyclus van het record te wijzigen.

Elektronische post dient -afgezien van de inhoud van de mededeling- beheerst te worden als . elektronische brieven, of . informele gegevens

Elektronische post

Voor gevorderden

Elektronische post kan op verschillende manieren worden beheerst. Dit is afhankelijk van de inhoud van de post.

- Een elektronische boodschap dient te worden beschouwd als een elektronische brief wanneer de inhoud gezien kan worden als een officiële verklaring.

⁶ Noot van de vertaler: het betreft hier een afgeleide van de Nederlandse Basis Archiefcode, die in de jaren vijftig binnen de EEG werd geïntroduceerd en onlangs op verzoek van de Europese Raad is geactualiseerd.

- Wanneer er geen toekomstig nut van de boodschap valt af te lezen, dient deze behandeld te worden als een informeel stuk. Dit kan bij voorbeeld bij een automatische boodschap wanneer iemand een mededeling doet dat hij momenteel niet aanwezig is en zijn boodschappen zal lezen zodra hij teruggekeerd is.
- Omdat elektronische post zowel een formeel als een informeel karakter kan hebben, is het soms moeilijk dit te beheersen. De archivaris kan met zijn vermogen om informatie te selecteren een belangrijke bijdrage leveren om nieuwe regels op het gebied van het beheer van elektronische post te definiëren.

4.5 Overdracht (transfer)

Als de actieve levensfase van een record is beëindigd, kunnen records worden overgebracht naar de archieven. Niet alle records hebben archiefwaarde. De documenten zonder archivistische waarde dienen te worden verwijderd met hulp van de archivaris op basis van vooraf vastgestelde selectiecriteria. (zie hiervoor het kader over de selectieve vernietiging in sectie 4.3.6)

Twee typen van overdracht

- *Fysieke overdracht*
- *Overdracht van verantwoordelijkheid*

Bij elektronische gegevens zijn er twee typen van overdracht:

- Fysieke overdracht van elektronische informatie
- Overdracht van verantwoordelijkheid.

Informatie wordt gewoonlijk verplaatst in de vorm van een record (nadat dit als record is ingenomen, vastgelegd). Het kan worden verzonden naar een andere organisatie of organisatie-onderdeel of naar een archiefdepot.

De verantwoordelijkheid voor elektronische informatie (vooral het management) en het fysieke beheer hoeven niet noodzakelijkerwijze hand in hand te gaan. Elektronische informatie kan door de afdeling, waar zij is ontstaan worden bijgehouden of door een daartoe gespecialiseerd organisatie-onderdeel of instituut. Het gemak waarmee vandaag de dag elektronische kopieën kunnen worden gemaakt of verzending hiervan wordt gerealiseerd maakt het mogelijk de twee rollen te scheiden. '

Hoe dan ook, er moeten altijd controles worden uitgevoerd voordat elektronische informatie wordt overgedragen.

Succesvolle fysieke overdracht

Opties

De sleutel tot een succesvolle fysieke overdracht ligt erin dat rekening wordt gehouden met enkele eenvoudige regels.

- Informatie moet worden gecontroleerd op compleetheid (ook de contextuele informatie moet aanwezig zijn);
- Verantwoordelijkheden moeten helder zijn gedefinieerd bij zowel de verzendende als de ontvangende organisatie;
- Wanneer elektronische informatie wordt overgebracht, moeten beide organisaties (zowel zender als ontvanger) er zeker van zijn dat de informatie niet wordt veranderd zonder voorafgaande toestemming.

Een voorbeeld uit het Verenigd Koninkrijk

Het Public Record Office in het Verenigd Koninkrijk heeft een nieuwe strategie voor de overdracht van elektronische records voorgesteld.

Overdracht vindt uitsluitend plaats via een beveiligd elektronisch netwerk.
Dit betekent dat de records onafhankelijk zijn van het medium, dat door de aanbieder wordt gebruikt.

De instantie die de records ontvangt om ze te bewaren (in dit geval: het Public Record Office) kiest een bepaald medium voor de opslag van de records. Deze technologie wordt gecontroleerd om de duurzaamheid te garanderen.

De overgedragen informatiebundel bestaat uit verschillende elementen, die samen zijn opgeslagen:

- Het elektronisch record zelf;
- Een controlerapport;
- De metadata van het record;
- Een viewer, die gebruikt kan worden op een PC (waarmee het document kan worden gelezen in de oorspronkelijke vorm);
- Een browser die ook op een PC gebruikt kan worden (waarmee documenten gezocht kunnen worden).

Men verwacht dat PC's in de toekomst voldoende mogelijkheid zullen hebben om compatibel te zijn, zodat de records in de toekomst blijvend gelezen kunnen worden. Elk record is hierdoor autonoom; problemen met het formatteren komen niet voor.

Controle van de records. Controles waarborgen een kwalitatief goede overdracht

Door een controle voorafgaande aan de overdracht wordt verzekerd dat de records goed leesbaar zijn.

Checks voorafgaande aan de overdracht

De volgende lijst is een voorbeeld van de stappen die door de overdragende organisatie genomen moeten worden bij het controleren van de informatie voor overdracht. Dit zou gedaan moeten worden in nauwe samenwerking met de ontvangende organisatie (bij voorbeeld de archiefdienst).

1. Maak twee kopieën van de gegevens
2. Vergelijk de gegevens met de geleverde documentatie, gebruik hierbij een statistisch programma
3. Identificeer en documenteer fouten in de documentatie
4. Voer cross-checks uit, zoals tussen records en variabelen om de stabiliteit van de gegevens te verifiëren
5. Raadpleeg de leverancier van de elektronische informatie wanneer het moeilijk blijkt om codes te identificeren of wanneer er fouten en inconsistentie in de gegevens blijken te zijn
6. Documenteer de fysieke bestanden en geef daarbij aan welke moeilijkheden zich voordeden tijdens het controleproces.

5. BEWARING VAN GEGEVENS VOOR KORTE EN LANGERE TERMIJN

Media en bestandsformaten. Twee sleutelfactoren bij bewaring gedurende lange termijn

Het spreekt vanzelf dat het fysieke medium waarop de elektronische informatie is opgeslagen, een zo lang mogelijke levensduur moet hebben. Dat geldt ook voor de technologie. Het heeft immers geen zin om records fysiek te bewaren als de apparatuur en programmatuur de gegevens niet langer kunnen bewerken of weergeven.

We hebben al gezien dat medium en inhoud bij elektronische records niet hetzelfde zijn. We zullen nu kijken naar standaard-typen van media en bestandsformaten, waarbij we ons richten op de levensverwachting van de media en de volwassenheid en duurzaamheid van standaards.

5.1 Media voor de opslag van gegevens

Drie typen medium

Naast papier:

- *Microfilm*
- *Magnetische media*
- *Optische media*

Er kunnen verschillende typen media worden gebruikt. Sommige zijn beter geschikt voor korte termijnopslag, terwijl anderen meer geschikt zijn voor de opslag op langere termijn.

Naast papier kunnen opslagmedia worden verdeeld in drie grotere families: microfilm, magnetische media en optische media, met een grote hoeveelheid subtypen binnen elke groep. We zullen hier niet op minder gebruikelijke typen ingaan, zoals de papieren tape.

Opslagmedia

basisconcepten

Microfilm: dit type opslag wordt veel gebruikt in archieven. Microfilm leent zich echter niet tot het eenvoudig hergebruik of zoeken van records. Er zijn vastgestelde standaarden en de levensverwachting is uitstekend.

Magnetische media: de technologie is redelijk oud, waarbij gebruik wordt gemaakt van polarisatie van magnetische deeltjes in de ene of de andere richting, om zo een bit op te slaan (nul of een). Voor deze opslagvorm gebruikt men gewoonlijk tapes die een volgorde van toegankelijkheid van de data mogelijk maakt.

Optische media: dit is het meest recente type opslagmedium. De technologie van compact discs gebruikt een afbuiging van een lichtstraal, die het oppervlak van een medium aftast om veranderingen in de waarde van bits te lezen. Deze vorm van opslag gebruikt gewoonlijk disks, waardoor een directe toegang tot de informatie wordt gegeven die sneller is dan tape. De intensiteit van opslag is hoog.

Tegenwoordig worden magnetische tape en microfilm veel gebruikt voor lange-termijn opslag, maar optische media beginnen gemeengoed te worden omdat ook zij hiervoor bijzonder geschikt zijn.

Voor korte-termijnopslag kan een brede variëteit aan media worden gebruikt, omdat de duurzaamheid en potentiële levensduur minder belangrijk zijn.

Magnetische media komen gewoonlijk voor in de vorm van tapes of cassettes

Magnetische media

Opties

De verschillende typen magnetische media zijn redelijk gestandaardiseerd, met variërende levensduur. De belangrijkste typen zijn hierna beschreven.

Diskette : 3,5 inch floppy disks zijn gestandaardiseerd en kunnen worden gebruikt voor veel systemen (PC's, Macintosh en Unix). Zij kunnen een relatief beperkte hoeveelheid informatie bevatten (gewoonlijk 1.44 MB) en omdat hun potentiële levensduur beperkt is, worden ze uitsluitend gebruikt voor zeer korte-termijn opslag en voor het overdragen van bestanden van de ene computer naar de andere.

Magnetische cassettes: cassettes worden veel gebruikt voor de opslag van data op middelgrote systemen. IBM maakt kwart-inch en half-inch cartridges.

Magnetische tape : 1 600-bpi tape is leesbaar op praktisch elke tape drive en is aanvaard door X-open als standaard formaat voor de uitwisseling van gegevens. 6 250-bpi tape met een capaciteit van 112.5 MB wordt in oudere archieven vaak gebruikt. Tapes moeten elke twee jaar opnieuw worden opgewonden en elke tien tot vijftien jaar worden herschreven op een nieuwe tape (van hetzelfde of een verschillend type).

8 mm video cassette : alhoewel gestandaardiseerd, is er maar één leverancier. De capaciteit op dit moment is doorgaans 2.3 GB. Cassettes moeten elke twee jaar worden overgeschreven.

DAT (digitale audio tape): DAT is gestandaardiseerd en alhoewel het hoofdzakelijk wordt gebruikt voor audio-opnamen, wordt er ook wel gebruik van gemaakt voor de opslag van andere informatie. De capaciteit op dit moment is doorgaans 2,3 GB. Elke twee jaar moeten de tapes worden herschreven.

Er zijn verschillende formats die tot nu toe, volledig leveranciers-afhankelijk zijn. Zo zijn er floppy disks van hoge intensiteit en verwisselbare harde disks. Aanvullend kunnen geheugenkaartjes van de omvang van een credit card, hoewel deze een beperkte capaciteit bezitten, een belangrijk opslagmedium worden in de toekomst.

De diskette, de magnetische cartridge en magneetband zijn stabiele, erkende standaards. De 8 mm videocassette en DAT-tape zijn standaards die worden gebruikt door enkele leveranciers, maar nog niet algemeen erkend.

De ontwikkeling van optische media gaat erg snel. De verwachting is dat deze media de magnetische media gaan verdringen voor lange-termijnopslag. Op dit moment wordt de snellere toegangstijd van magnetische media belangrijker gevonden dan de massa-opslagmogelijkheid van optische media.

Optische schijven hebben grote opslagcapaciteit en in tegenstelling tot tapes, een directe toegang.

Optische media

Opties

De belangrijkste typen optische media worden hieronder beschreven.

CD-ROM: een gestandaardiseerd medium. Dit maakt deze techniek tot een goed medium voor de opslag van archiefmateriaal. Er is een verschil tussen geperste CD-ROM's (deze zijn geschikter voor het maken van grote hoeveelheden kopieën en betrouwbaarder) en individueel vervaardigde CD-ROM's (goedkoper voor kleinere hoeveelheden, maar minder betrouwbaar tenzij zij een vernislaag krijgen nadat ze zijn aangemaakt). CD-ROM's moeten elke tien tot twintig jaar worden herschreven. Alhoewel het niet altijd nodig is om ze op te slaan in een geklimatiseerde omgeving (zoals dit dient te gebeuren met magnetische tapes), zorgt het wel voor een langere levensduur. De duurere Cd-rom's, die zijn ingebrand op glas, hebben een veel langere levensverwachting.

DVD (digital versatile disk). Dit nieuwe medium zou wel eens een uitstekend medium voor archivering kunnen worden. DVD's hebben een grote capaciteit (4.7 tot 18 Gb) en zullen spoedig wijdverbreid zijn. De meer recente DVD-spelers kunnen ook Cd-rom's inlezen.

WORM (write once, read many): er bestaat geen internationale standaard en WORM-disks worden in niet zo veel organisaties toegepast. Zij dienen elke tien tot twintig jaar te worden herschreven.

Erasable optical disk: de meeste van deze disks gebruiken magneto-optische technologie, die optische en magnetische opslagmedia combineert. Het medium kent een combinatie van snelheid, hoge opslagcapaciteit en de mogelijkheid om gegevens verschillende keren te herschrijven. Ook de gegevens op deze disks dienen elke 10 tot 20 jaar te worden geconverteerd.

CD-ROM en DVD zijn algemeen geaccepteerde standaards. De erasable optical disk is een standaard die door enkele leveranciers wordt gehanteerd. WORM is als standaard onbetrouwbaar.

Merkfafhankelijke software wordt dikwijls gebruikt om toegang te krijgen tot opgeslagen gegevens. Het is belangrijk om zorgvuldig na te gaan of de bestandsformaten en de zoekpaden eenvoudig kunnen worden benaderd.

Gezien de snelle ontwikkeling van de technologie en de beperkte levensverwachting van de elektronische media is het raadzaam om digitale archieven periodiek te herschrijven. Alhoewel dit leidt tot verhoogde kosten voorkomt het veel problemen die worden opgeworpen door niet-gestandaardiseerde formaten en de veranderende technologie. De meeste magnetische en optische media gebruiken foutherkenning en hebben correctiefaciliteiten die het automatisch herstel van gegevensfouten mogelijk maken.

Elke keer dat een archiefbestand wordt herschreven dient de gebruiker te kiezen tussen:

- Het houden van de oude bestanden in hun huidige vorm, of
- Het overzetten van het bestand naar een modern medium en/of format (zie ook sectie 4.3.5 over het converteren van het ene digitale formaat naar het andere).

Niet alleen dient aandacht te worden besteed aan het financiële aspect, maar ook aan dat van de toegankelijkheid, duurzaamheid en het behoud van de authenticiteit.

Het behoud van elektronische records

Het is belangrijk om bepaalde regels in acht te nemen op het gebied van de omgeving waarbinnen de gegevens worden opgeslagen

Lange termijnopslag

Voorbeeld

Er is discussie gaande binnen ISO, ANSI en ICA⁷ over de beste manier om elektronische records op te slaan. De volgende cijfers geven enkele praktijkvoorbeelden:

- Gemiddelde temperatuur: +18° C/-5° C
- Relatieve luchtvochtigheid: +40%/-5%
- Interval tussen herschrijven: 10 jaar.

⁷ ISO: International Standards Organization
ANSI: American National Standards Institute;
ICA: International Council on Archives

5.2 Bestandsformaten

Er is een groot aantal standaards, afhankelijk van het type gegevens dat dient te worden opgeslagen. Het beste is om te kiezen voor een beperkte groep standaards om hierdoor de circulatie van informatie te vergemakkelijken. Voor opslag op korte- en voor langere termijn dienen bij voorkeur dezelfde formats moeten worden gebruikt.

Verschillende formaattypen:

- *Grafische bitmaps*
- *Grafische vectors*
- *Tekst*
- *Data*
- *Programmatuur*
- *Video en audio*
- *Andere*

De verschillende formaattypen

basisbegrippen

De bestaande formats kunnen worden gegroepeerd tot verschillende grotere families, afhankelijk van hun inhoud.

❶ bitmapped graphics (raster grafieken). Deze bestaan uit een stippenpatroon. Wanneer bij voorbeeld een document wordt gescand, dan ontstaat dit type van grafisch beeld. Bitmapped graphics nemen veel ruimte in beslag en worden gewoonlijk gecompriemd voordat zij worden opgeslagen. Zij kunnen worden gebruikt als een bron voor latere codering (via OCR-reading, de computer herkent dan de tekst of de vector en deze kan opnieuw worden gebruikt en gemanipuleerd) of opgeslagen als vorm van het document, als elektronische fotokopie. Het type compressie bepaalt het type van grafisch formaat. Fax formats zijn bij voorbeeld een bepaalde vorm van gecompriemde beelden.

❷ Gevectoriseerde grafieken: bij kaarten, grafieken of andere grafische afbeeldingen die uitsluitend bestaan uit lijnen, blokken en figuren, kan veel ruimte worden bespaard door een vectorstructuur te gebruiken, waarbij uitsluitend de coördinaten van de vector die de grafische afbeelding bepalen (kromme en rechte lijnen, hoeken, driehoeken, cirkels etc.) worden opgeslagen. Dit type formaat is niet bruikbaar voor de opslag van foto's.

❸ Tekst: in zijn algemeenheid zijn bij tekst drie verschillende aspecten betrokken:

- De platte tekst zelf, die bestaat uit een set gecodeerde karakters met een bepaalde betekenis en samenhang;
- de tekststructuur (bij voorbeeld titel, hoofdstukken, termen die moeten worden benadrukt, lijsten en indexen etc.)
- De lay-out (termen worden benadrukt door gebruik van andere lettertypen, vet of een bepaalde kleur etc.).

❹ Data: de mogelijkheid om de gegevens te bewaren met de oorspronkelijke functionaliteit kunnen erg verschillen. Voor spreadsheets bestaat geen mogelijkheid om de rekencapaciteit te behouden zonder dat het programma wordt bewaard bij de spreadsheet. Voor toepassingen die zijn gebaseerd op informatie en procesmodellen nemen de mogelijkheden van behoud toe. De beste manier is om de gegevens uit de database op te slaan volgens een bepaalde afgesproken standaard.

⑤ Programmatuur: programma's zijn afhankelijk van het IT-platform. Ze geraken steeds sneller in onbruik

⑥ Audio, video en alle andere objecten die onderdeel kunnen uitmaken van een record. Hierbij denken we aan multi-mediatoepassingen, waarbij bij voorbeeld stemgeluid of muziek wordt gecombineerd met tekst en bewegend beeld.

De Toekomst

Voor gevorderden

In de toekomst zullen documenten en databasebestanden steeds vaker samengestelde documenten of zelfs object-georiënteerde documenten worden. In andere woorden: documenten zullen bestaan uit verschillende onafhankelijke elementen die met elkaar worden verbonden (tekst, bewegende en stilstaande beelden, geluidsfragmenten).

De tot nu toe ontwikkelde standaards (Microsoft's OLE, Open Doc van IBM en Apple en zelfs Java van Sun) zijn nog niet stabiel genoeg voor elektronische records die hun eigen bewerkingsinstructies met zich mee dragen (lezen, navigeren etc.).

5.2.1 Compressie van bitmaps

Compressie van bitmaps kan mét en zonder verlies van informatie

Er worden twee compressietechnieken gebruikt om de omvang van bitmaps te verkleinen.

1. **compressie waarbij geen verlies optreedt:** Nadat het document is gecomprimeerd en gedecomprimeerd is het beeld exact gelijk aan het origineel. Dit vraagt lage compressiewaarden van om en nabij 2:1.
2. **Compressie waarbij verlies optreedt:** hierbij wordt de minder belangrijke informatie van een document niet opgeslagen. Het menselijk oog onderscheidt bepaalde delen van een document beter dan andere. Afhankelijk van de hoeveelheid acceptabel verlies kunnen hierdoor hoge compressiewaarden worden bereikt.

Er zijn veel verschillende vormen voor bitmap graphics. Het is belangrijk om een volwassen en blijvende standaard te kiezen

Bitmap-formaten

Opties

De volgende beschrijvingen geven de belangrijke grafische bestandsformaten (inclusief fax- en videoformaten) weer.

TIFF (tag image file format): dit formaat wordt vaak gebruikt voor gescande bestanden. Er zijn verschillende opties, die afhankelijk zijn van het aantal kleuren. Bij dit formaat treedt geen verlies op, waardoor slechts een beperkte compressie wordt bereikt. De laatste versie van TIFF (6.0) ondersteunt multi-page opties. Om er zeker van te zijn dat het TIFF-formaat volledig uitwisselbaar blijft, dient over het formaat heen geen ander compressiemiddel (zoals Packbits 32 733 bij ITU-T, LZW of JPEG) te worden gebruikt.

GIF (graphics interchange format): dit format is ontwikkeld door CompuServe en wordt vrij algemeen gebruikt, vooral op Internet. Er zijn twee specificaties: GIF 87A en GIF 89A. Browsers kunnen meestal beide formats lezen.

JPEG (joint photographic experts group): deze internationale standaard wordt steeds meer gebruikt (inclusief op Internet). Het is een format waarbij gegevens verloren gaan tijdens compressie, wat een hoge compressiewaarde oplevert. Het is onbetwist een goede keuze in termen van opslagruimte en duurzaamheid.

Faxen: er zijn twee faxformaten, afhankelijk of een analoge- of een ISDN-lijn worden gebruikt.

Video. Er bestaan twee videoformaten:

MPEG-1, voor computer- en multimedietoepassingen;

MPEG-2, een recentere versie, voor digitale televisie (inclusief geluid).

Andere grafische formaten: het kan gevaarlijk zijn om andere formaten te gebruiken voor grafische afbeeldingen, zoals BMP en PCX, omdat er geen garantie is voor hun duurzaamheid. Het foto CD-formaat van Kodak wordt bij de opslag van documenten nog weinig gebruikt.

JPEG, fax en video zijn stabiele, erkende standaards. GIF en TIFF zijn geen standaard, maar worden wel door enkele leveranciers toegepast.

5.2.2 Vectorbestanden

Vectorbestanden maken het mogelijk om de structuur van een grafiek op te slaan. De volgende opties zijn aanwezig:

CGM (Computer Graphic Metafile): dit gestandaardiseerd formaat voor vectors biedt een betrouwbare garantie voor duurzaamheid van de gegevens.

Formats die specifiek zijn voor bepaalde typen applicatie: sommige applicaties vragen een specifieke aanpak omdat ze eigen standaards hebben, bij voorbeeld geografische informatiesystemen (GIS) of computer-aided design (CAD).

Andere grafische formaten vermijd andere niet-gestandaardiseerde formaten (zoals PICT bij Macintosh, Microsoft Windows Metafiles of andere specifieke applicaties), omdat er geen garantie is voor de houdbaarheid van de gegevens op lange termijn.

CGM is een betrouwbare standaard, voor CAD- en GIS-bestanden zijn nog geen standaards beschikbaar.

5.2.3 Tekstbestanden

Een tekstbestand inclusief de structuur maar zonder de lay-out, is hardware-onafhankelijk

Er zijn verschillende typen tekstbestanden, afhankelijk van hoe de structuur en/of de lay-out worden bewaard.

- **Een plat tekstbestand** is een bestand van laag niveau dat de tekst bevat in niet meer dan een rij opeenvolgende karakters. Het is moeilijk om binnen het tekstbestand te navigeren, omdat er geen structuur aanwezig is.
- **Een gestructureerd platte-tekstbestand**: hierbinnen kan wel worden genavigeerd. Het bestand is onafhankelijk van de hardware.
- **Een volledig geformatteerd tekstbestand** bevat zowel de karakters, structuur als de lay-out. Het bestand is niet onafhankelijk van de hardware die nodig is om het te lezen.

Karaktersets

Voor gevorderden

Er zijn drie hoofdfamilies van codes waarmee karakters worden weergegeven:

- ISO 646 is (bijna) gelijk aan de ASCII karakterset. Deze 7-bit standaard kent geen bijzondere Europese karakters (zoals diacritische tekens).
- 8-bit karaktersets zijn een superset van ISO 646. Er zijn twee standaards belangrijk voor de Europese unie: ISO/IEC 8859-1 voor west-Europa en ISO/IEC 8859-7 voor Griekenland. Er bestaan ook karaktersets voor Arabisch, Hebreeuws en Cyrillisch schrift.
- De universele karaktersets (UCS) standaard (ISO/IEC 10646) bevat de meeste karakters en symbolen die in de wereld worden gebruikt. Een karakter wordt gepresenteerd door 2 bytes (UCS-2, Unicode) of 4 bytes (UCS-4). Een andere code (UTF) gebruikt een variabel aantal bytes en is ontwikkeld om het transport van multi-byte karakters tussen verschillende machines gemakkelijker te maken. Dat gebeurt door het gebruik van 8-bit controlekarakters in een multi-

byte karaktersvolgorde te vermijden (de controlecode om het einde van een karakterstring aan te geven is bij voorbeeld vaak 00. In UTF wordt de waarde 00 in geen enkele van de andere bytes gebruikt, terwijl bij voorbeeld in de UCS-2 de letter "A" wordt weergegeven door 00 64.

Lay-out. Het is beter om uitsluitend de structuur van een tekst te bewaren dan de stijl die met elk element van dat document samenhangt.

Veel mogelijkheden van de lay-out van een tekst zijn afhankelijk van het platform dat wordt gebruikt. Het heeft bij voorbeeld geen zin om de cursor te specificeren voor een papieren uitdraai van een tekst. Met de snel voortschrijdende techniek zal het mogelijk zijn om met bepaalde hulpmiddelen de teksten die vandaag -in welke vorm dan ook- worden opgeslagen, in de toekomst te lezen.

Het vastleggen van de lay-out voor lange-termijnopslag (karaktersets zoals bij voorbeeld Century Schoolbook of Courier, of tekstmacro's) roept problemen op wanneer de tekst opnieuw gebruikt wordt. De beste oplossing is om uitsluitend de structuur van de tekst op te slaan (bij voorbeeld: dit is een kop van het eerste niveau) in plaats van ook nog de daarmee gepaard gaande stijl (bij voorbeeld: kop eerste niveau, 16 punts Arial/Helvetica). Het is dan aan het programma waarmee de tekst weer wordt opgeroepen om een bepaalde stijl voor de verschillende onderdelen te kiezen.

Gestructureerde tekst. Er zijn verschillende standaards en veel merkgebonden formats

Gestructureerde tekst

Opties

Er zijn verschillende standaards om gestructureerde tekst op te slaan.

SGML (standard generalized markup language): strikt genomen is dit niet een structuur, maar een manier waarop tekens worden vastgelegd. Het is een internationale standaard en wordt algemeen toegepast. SGML kan worden gebruikt om de tekst en zijn structuur vast te leggen, maar zonder de lay-out. Er zijn verschillende aanvullende standaards, zoals:

- DSSL: document style semantics and specification language;
- SPDL: standard page description language;
- SDIF: SGML document interchange formats;
- Font information interchange;
- DTD: document type definition (er zijn verschillende types, afhankelijk van het type document)

HyTime : dit is een uitbreiding van SGML die gebruikt kan worden om multimedia-materiaal in samengestelde documenten te verwerken

HTML (hypertext markup language): dit is een vereenvoudigde uitvoering van SGML en is redelijk standaard op Internet websites. Het is niet erg stabiel en minder goed te gebruiken voor omvangrijkere documenten. Voor lange-termijnopslag heeft SGML de voorkeur.

ODA (Office document architecture): dit is een internationale standaard die kan worden gebruikt om de tekst, structuur en lay-out van een bepaald document te integreren. ODA is ontwikkeld om te gebruiken binnen de elektronische kantooromgeving. Het is echter niet platform-onafhankelijk.

RTF (rich text format): dit formaat wordt hoofdzakelijk gebruikt bij alle Microsoft Office software. Er is geen garantie voor stabiliteit en duurzaamheid. Het is aan te bevelen om exportfilters te gebruiken of conversie software om bestanden die zijn aangemaakt onder RTF te bewaren.

PostScript: deze beschrijvingsstandaard van Adobe wordt ruim toegepast om tekst met lay-out af te drukken of te verzenden. Het is nodig om PostScript te vervangen door een meer open standaard.

PDF (portable document format): hiermee kunnen documenten op verschillende platforms worden afgebeeld.

Er bestaan ook een aantal merkgebonden formats die afhankelijk zijn van een bepaalde leverancier.

SGML en HyTime zijn standaards. HTML en ODA zijn een onderlinge standaard die door enkele leveranciers is afgesproken. RTF, PostScript en PDF zijn merkgebonden applicaties (alhoewel PDF al redelijk algemeen wordt toegepast).

EDI en elektronisch berichtenverkeer. Twee specifieke typen qua toepassing die verschillende structuren gebruiken

Specifieke applicaties

Tekststructuur kan extra velden bevatten om het indexeren te vergemakkelijken. Dit heeft vooral betrekking op elektronische berichten.

Standaards vereenvoudigen de uitwisseling van gestructureerde informatie, vooral bij zakelijk gebruik. Deze standaards zijn EDI (electronic data interchange) en vooral Edifact. In aanvulling op de bekende EDI-definities, zijn in verschillende sectoren van de handel en industrie EDI-standaards ontwikkeld, zoals in het bankwezen en de autoindustrie. Hoewel EDI erg aantrekkelijk is voor de formele uitwisseling van gegevens tussen organisaties (waarbij volledige arbeidsprocessen worden geautomatiseerd, inclusief de administratieve afhandeling) zijn de EDI-standaards traag en wordt hierbij niet altijd rekening gehouden met de snelle ontwikkeling van de technologie (zoals deze plaatsvindt bij grafische afbeeldingen, audio en video).

5.2.4 Gegevens en programma's

Gegevens en programma's vormen een groot probleem omdat er geen algemeen geaccepteerde standaard is vastgesteld

Er is op dit moment geen standaard format voor gegevensbestanden, zoals in de vorm van spreadsheets en databases. Om er zeker van te zijn dat de gegevens over een langere periode nog te raadplegen zijn, moeten gebruikers een hulpmiddel hebben waarmee zij het oude formaat kunnen lezen -of de software zelf bewaren.

Het is belangrijk om de interface tussen het programma en de gegevens (bij voorbeeld SQL bij een database) afzonderlijk op te slaan van het bestandsformaat. Programma's zijn niet zo onafhankelijk van hardware als de gegevens dat zijn.

Wanneer er geen standaardformat voor een bepaald gegevenstype is, dan is het raadzaam om

- een formaat te gebruiken dat veel wordt toegepast en gelezen kan worden door veel programma's, of
 - een conversie-strategie te ontwikkelen (of een strategie om ook de software te bewaren).
- Hoofdstuk 3.4 geeft enkele suggesties om een dergelijke strategie te ontwikkelen.

Het bewaren van programmatuur roept dezelfde problemen op als boven geschetst. De gebruiker dient het bronprogramma te bewaren en eventueel ook de hardware waarop het programma kan worden gebruikt.

Verskillende typen gegevens betreffen verschillende typen bestanden:

- *spreadsheets*
- *gegevensbestanden*
- *reken- en zakelijke toepassingen*
- *formulieren*
- *digitale handtekeningen*

Gegevensbestanden

Opties

Gegevens kunnen worden bewaard in verschillende bestandstypen. De belangrijkste zijn:

① **Spreadsheets** : het format van EXCEL kan worden beschouwd als een *de facto* standaard. De mogelijkheid om tabellen in SGML op te nemen kan worden gezien als een eerste stap naar een open standaard.

② **Database-bestanden** hoewel steeds meer databases SQL gebruiken als zoektaal, bestaat er geen format van hoog niveau om databases op te slaan en te bewaren. De beste oplossing is om een algemeen toegepast databaseprogramma te gebruiken of om de gegevens als platte tekst op te slaan met scheidingstekens tussen de velden, zodat de database door een ander programma weer opnieuw op te bouwen is. De structuur van de database moet in dat geval wel ook worden vastgelegd.

③ **Edifact (electronic data interchange for administration, commerce and transport)**: verschillende berichttypen maken het mogelijk om facturen en andere boekhoudkundige bestanden uit te wisselen.

④ **STEP (standard for the exchange of product data)**: deze set van standaards maakt de uitwisseling van gegevens over het vervaardigen van producten mogelijk. De 'Express-language for product data representation and interchange' maakt hiervan deel uit. Deze standaard wordt op veel gebieden gebruikt.

⑤ **IDEF (function and information modelling)**: IDEF0 is een methodologie die wordt gebruikt voor elk applicatie protocol in de STEP-standaard. Het wordt veel toegepast in processen waarbij bedrijven worden heringericht, bij business process re-engineering. IDEF3 richt zich op het verbeteren van bedrijfsprocessen en geeft enkele standaard grafische formats voor workflow.

⑥ **Formulieren** hierbij hoeft uitsluitend de in het formulier ingevoerde informatie te worden vastgelegd samen met een kopie van het blanco formulier. Er worden pogingen ondernomen om tot een standaard te komen (FIMS, HTML 3.0 enz.).

⑦ **Digitale handtekeningen** er zijn twee formats voor het versleutelen van handtekeningen: DES en RSA. Deze kunnen worden gebruikt om de authenticiteit van documenten te controleren. Encryptietechnieken zijn beschreven onder 6.3.1.

6. HET VERKRIJGEN VAN TOEGANG TOT - EN DE VERSPREIDING VAN INFORMATIE

6.1 Naar een informatiemaatschappij

De informatiemaatschappij plaatst elektronische informatie in het centrum

De Europese Commissie heeft het idee van de informatiemaatschappij omarmd, een maatschappij waarin elektronische informatie in het middelpunt wordt geplaatst

De informatiemaatschappij

Voor gevorderden

In januari 1994 presenteerde Jacques Delors, toenmalig voorzitter van de Europese Commissie, een White Paper over groei, concurrentie en werkgelegenheid.

Op hetzelfde moment vroeg hij Martin Bangemann, als lid van de Europese Commissie verantwoordelijk voor de portefeuille Industrie, om een rapport voor te bereiden, samen met een werkgroep bestaande uit prominente deskundigen op het gebied van de media, telecommunicatie en informatietechnologie. Dat rapport "Europa en de mondiale informatiemaatschappij" handelt niet uitsluitend over infrastructurele voorzieningen, maar ook over diensten, toepassingen en maatschappelijke aspecten.

In februari 1995 hield de Group of Seven haar eerste bijeenkomst over de informatiemaatschappij. De groep stelde voor om 11 pilotprojecten te starten, inclusief een model voor een "on-line government", oftewel een digitaal bestuur. Het gebruik van nieuwe technologieën moest het bestuur dichter bij de mensen brengen.

De Commissie gaat verder met het ontwikkelen van bouwstenen voor de informatiemaatschappij binnen Europa en heeft inmiddels aandacht verkregen bij publiek en beleidsbeslissers. Hiervoor heeft de commissie diverse nieuwe initiatieven ondernomen, waaronder het "Rolling Action Plan" dat een opsomming bevat van alle belangrijke acties, in het bijzonder de wettelijke maatregelen, die noodzakelijk zijn om de Informatiemaatschappij in Europa verder tot ontwikkeling te brengen. Hiervoor moeten vier belangrijke functies worden vervuld:

- het dienen als leidraad voor de lidstaten en andere Europese instituties (helderheid en transparantie);
- het leveren van gedetailleerde informatie over de ontwikkeling van een regulerend raamwerk;
- het leveren van open informatie aan alle geïnteresseerde partijen, vooral het publiek, over EU-beleid, waarbij een constructieve dialoog tussen de Commissie en alle relevante onderdelen van de maatschappij wordt gestimuleerd;
- het leveren van een handig management-hulpmiddel voor de Commissie.

Het is niet alleen van vitaal belang dat elektronische informatie wordt bewaard, maar ook dat deze zo eenvoudig mogelijk toegankelijk is. Informatie circuleert steeds efficiënter binnen organisaties. Niet alleen officiële rechtspersonen, maar ook burgers hebben eenvoudiger toegang tot informatie die voor hen is bedoeld of waarop zij recht hebben. Zelfs wanneer elektronische informatie niet meer dagelijks wordt gebruikt, kan het hergebruikt worden voor statistisch, wetenschappelijk of historisch onderzoek.

De archivaris en de informatiemaatschappij. Archivarissen zorgen voor het behoud van het collectieve geheugen van de informatiemaatschappij

Het beroep van archivaris vormt samen met andere beroepen binnen de informatiesector een essentieel en actief onderdeel van de informatiemaatschappij. Van hen wordt verwacht dat zij het collectieve

geheugen van de informatiemaatschappij behouden. Het DML-forum op 18-20 december 1996 in Brussel was de eerste interdisciplinaire benadering van dit vraagstuk.

Privacy en openbaarheid: *conflicterende belangen die door nationale en internationale wetgeving worden beschermd.*

Een van de problemen bij het verspreiden van elektronische informatie is om de juiste informatie te onderscheiden die moet worden behouden of waaruit een selectie moet worden gemaakt. Informatie kan worden gebruikt op een manier die nooit vooraf door de opsteller werd verwacht. De items in een database kunnen bijvoorbeeld gebruikt worden voor demografisch onderzoek, gebaseerd op specifieke selectiecriteria.

Hierbij bestaat steeds een spanningsveld tussen privacy en openbaarheid. De scheidslijn is door wetgeving bepaald. Het is belangrijk om de regels te raadplegen die de Europese en landelijke wetgeving op dit gebied stellen.

Visueel gehandicapten

Voor gevorderden

Visueel gehandicapten gebruiken twee verschillende methoden voor het lezen van elektronische informatie

1. Blinden en personen met een sterk verminderd gezichtsvermogen kunnen elektronische informatie gebruiken door een standaard-computerprogramma (tekstverwerker, rekenschema, begevensbestand enz.) te combineren met een extra programma dat op de achtergrond werkt. Voor personen met een verminderd gezichtsvermogen wordt (de tekst op) het scherm vergroot en voor blinden wordt extra output geleverd zoals braille of een kunstmatige stem. Deze faciliteiten bestaan voor alle computerplatforms. Het lezen van omvangrijke ongestructureerde documenten blijft moeilijk omdat een visueel gehandicapte niet de volledige pagina van het document kan overzien.
2. Beide groepen kunnen veel sneller lezen door gestructureerde documenten te gebruiken en speciale leesprogramma's (= driedimensionale browsers) die geluids-, gesproken- of braille-informatie direct doorzenden naar de hulpmiddelen voor de uitvoer van de gegevens (bijvoorbeeld een synthesizer, waarmee een kunstmatige stem en ander geluid kan worden gemaakt, of een braille-leeslijn). De tekst op het scherm wordt opnieuw geformatteerd om het lezen voor mensen met beperkt gezichtsvermogen te vergemakkelijken. Deze browsers zijn zowel voor SGML als voor HTML beschikbaar.

6.2 Standaards voor de toegang tot informatie

Verspreiding van informatie *Er moet gekozen worden voor een standaard om gemakkelijk toegang tot de gegevens te verkrijgen*

Sommige standaards zijn geschikter dan andere om informatie voor een zo groot mogelijke groep mensen toegankelijk te maken. Enkele hiervan zijn al genoemd in hoofdstuk 5 dat handelde over standaards voor de opslag van informatie. Hieronder zullen we een selectie geven van de standaards die het meest geschikt zijn voor het selecteren van informatie. Waar dit van toepassing is, kunnen voor de opslag en de verspreiding van informatie verschillende formats worden gebruikt.

Standaarden voor verspreiding van informatie

Opties

Media voor off-line distributie :

- 3 ½ inch 1.44 MB high-density disk;
- standaard CD-ROM;
- in de toekomst zou DVD (digital video disk) kunnen uitgroeien tot een uiterst geschikt medium.

Internet protocols voor het uitwisselen van documenten en het aanbieden van documenten via een netwerk:

- TCP/IP (transfer control protocol, Internet protocol): deze twee protocols maken een uitwisseling via het netwerk mogelijk. Ze worden algemeen toegepast op Internet en Intranet netwerken.
- HTTP (hypertext transport protocol) voor Web servers;
- FTP (File Transfer Protocol) voor file servers;
- SMTP (simple mail transfer protocols) voor elektronisch berichtenverkeer;
- MIME (multipurpose Internet multimedia extensions) maakt het mogelijk om in een boodschap verschillende formats te gebruiken (bij voorbeeld tekst met verschillende karaktersets, beeldmateriaal).

Document formats:

- SGML for portable documents. Het HTML-formaat is beter geschikt voor korte documenten zoals de home pages op het World Wide Web. Langere documenten kunnen meteen in SGML-format op webservers worden geplaatst.
- Tekstverwerkings formats die leveranciersgebonden zijn (Microsoft Word, Wordperfect). Hoewel deze niet worden aangeraden voor bewaring op langere termijn, zijn het wel *de facto* standaards die algemeen worden gehanteerd in de door Windows geregeerde PC-wereld.

Database-interfaces

- SQL-2 (structured query language) voor relationele databases;
- ISAM (indexed sequential access method) voor geïndexeerde sequentiële bestanden (dit is een zogenaamde low-level interface).

6.3 Beveiliging

6.3.1 Toegangsrechten

Aan gebruikers kunnen verschillende toegangsrechten tot informatie worden verleend:

- toegang tot de voorpagina van een document (met of zonder de meta-gegevens);
- toegang tot een deel van het document, of het document als geheel;
- het recht tot inzage, met de mogelijkheid tot afdrukken;
- andere rechten.

Een bestand anoniem maken is minder eenvoudig dan het lijkt

Over toegangsrechten dient zorgvuldig te worden nagedacht om ieders individueel recht op privacy en anonimiteit te waarborgen. Er zijn twee manieren om een bestand te anonimiseren voordat het wordt vrijgegeven voor openbaar gebruik:

- door het weglaten van bepaalde velden
- door het samenvoegen van gegevens tot statistische informatie

Anoniem of niet?

Voorbeeld

Er moeten voorzorgsmaatregelen worden genomen om er zeker van te zijn dat de procedure om een bestand anoniem te maken het onmogelijk maakt om uit bepaalde records een selectieve keuze te maken, waaruit toch weer persoonlijke gegevens kunnen worden afgeleid.

Stel dat de namen van personen uit een database zijn verwijderd. Een selectieve zoekactie zou gegevens kunnen opleveren over het beroep van een persoon, gecombineerd met een bepaalde woonplaats. Een verdere zoekactie zou hierdoor een bepaald individu kunnen identificeren en details kunnen onthullen over -bij voorbeeld- zijn of haar jaarsalaris of schuldenlast. Zo zal de hartspecialist in een dorp van 1000 inwoners vrij snel te identificeren zijn, ondanks dat zijn naam uit de database is verwijderd.

6.3.2 Encryptie en het bewijs van de authenticiteit

Het versleutelen van gegevens:

- *Om de vertrouwelijkheid te bewaren*
- *Om een record authenticiteit te verlenen*
-

Er zijn twee redenen om gegevens te versleutelen:

- Om te voorkomen dat onbevoegden een document lezen;
- Om er zichzelf van te verzekeren dat het document daadwerkelijk afkomstig is van een bepaalde persoon en niet door iemand anders is veranderd.

Bij het versleutelen van gegevens dient rekening te worden gehouden met specifieke nationale aspecten. De wetgeving op het gebied van encryptie varieert van land tot land. Er kunnen beperkingen zijn in de export van encryptie-software (zoals in de VS) of in het gebruik (zoals in Frankrijk). Dit onderwerp kent meer aspecten dan uitsluitend het garanderen van een bepaald niveau van geheimhouding of de toegankelijkheid van informatie. Het is bij voorbeeld ook van belang dat de verzender zowel als de ontvanger ervan verzekerd kunnen zijn dat de juiste informatie is ontvangen. Het vaststellen van de authenticiteit van een document is een belangrijke factor in het verkrijgen van informatie met een hoge graad van betrouwbaarheid.

Encryptie-algoritmen

We kennen twee belangrijke families van algoritmen op het gebied van de encryptie.

single-key (symmetrische) algoritmen

DES (data encryption system) type (NIST FIPS 46-1, benadert ISO 8227-DEA) Hierbij wordt dezelfde encryptie-sleutel gebruikt voor het coderen en decoderen van de data.

- Double-key (asymmetrische) algoritmen

RCA-type (Rivest, Shamir en Adelman, namen van de uitvinders). Hierbij heeft elke partij twee sleutels, waarvan de ene geheim blijft (de privé-sleutel) en de andere algemeen beschikbaar is (de openbare sleutel). Een boodschap die is gecodeerd met een van de sleutels kan met de andere worden gedecodeerd.

Voor een boodschap waarbij de authenticiteit moet worden vastgesteld zendt de verzender de gecodeerde boodschap met de openbare sleutelgegevens van de ontvanger. Alleen de ontvanger kan deze boodschap lezen, door gebruik te maken van zijn of haar privé-sleutel in combinatie met deze openbare sleutel.

Het PGP (Pretty good privacy) programma van Philip Zimmermann is een public domain (dus: gratis toegankelijk) encryptieprogramma van het RSA-type dat op Internet beschikbaar is (dit ondanks het verbod van de VS op de export van versleutel-algoritmen).

Er bestaan servers voor het bewijzen van de authenticiteit van documenten. Zij geven de mogelijkheid om een publieke sleutel te verkrijgen via een procedure die de authenticiteit waarborgt.

6.4 Toegang tot gegevens

Verschillende manieren om gebruikers toegangsrechten tot data te verlenen

Er zijn verschillende wegen om mensen toegang te verlenen tot elektronische informatie.

- De elektronische leeskamer, dit zijn faciliteiten die open staan voor het publiek waarbij met behulp van computers elektronische informatie kan worden gelezen;
- Het kopiëren van elektronische informatie om thuis te gebruiken;
- Het maken en onderhouden van een werkkopie van elektronische informatie in een verschillend, meer gebruikersvriendelijk formaat (toegang kan ter plaatse worden verleend, online, of door de gebruiker een kopie te geven op een bepaald medium)
- Het gebruik van een algemeen toegankelijk model dat is gebaseerd op metadata die automatisch toegang biedt tot de elektronische informatie die zinvol is voor de gebruiker (de toegang kan ter plekke worden verleend, online, of door de gebruiker een kopie te geven op een bepaald medium, zoals diskette of CD).

De laatste twee oplossingen zijn zinvol als het beleid erop gericht is om gegevens via Internet te verspreiden.

Duitsland

Voorbeeld

Bij gegevensbestanden in Duitsland wordt onderscheid gemaakt tussen "research copies" die beschikbaar zijn en geraadpleegd kunnen worden door een grote groep onderzoekers en "archiefkopieën", die worden opgeslagen als platte tekst om problemen met standaard-formats te voorkomen.

Verspreiden van informatie:

- *informatie beschikbaar maken (passieve verspreiding)*
- *distributie van informatie naar doelgroepen (actieve verspreiding)*

Het verspreiden van informatie vraagt om een beleid op het gebied van een bewustwording die noodzakelijk is met betrekking tot potentiële gebruikers. Er zijn twee mogelijke elkaar aanvullende strategieën:

- De informatie beschikbaar maken en het overlaten aan de gebruiker om on-line navigatiehulpmiddelen te gebruiken (passieve beschikbaarstelling);
- Informatie distribueren naar een bepaalde groep gebruikers (actieve beschikbaarstelling).

Deze twee strategieën kunnen worden gecombineerd met het aan een bepaalde doelgroep distribueren van details over de manier waarop informatie beschikbaar is.

Het formuleren van een beleid op het gebied van de verspreiding van informatie is essentieel. Hierdoor wordt aan potentiële gebruikers toegang tot de informatie verleend.

7. Conclusies

Deze richtlijnen pogen om enkele voorbeelden te geven van de manier waarop elektronische informatie momenteel wordt gebruikt en geven suggesties om een strategie te definiëren.

Er bestaat geen standaard-oplossing voor alle landen en elke situatie. Het is daarom aan u om uw eigen beleid op het gebied van elektronische informatie te formuleren, samen met iedereen die daarbij betrokken is.

Wanneer deze richtlijnen daarbij behulpzaam zijn, dan hebben zij hun doelstelling bereikt.

8. Annexen

- 8.1 Terminologie
- 8.2 Lijst van nog op te lossen problemen
- 8.3 Het waarom van standaardisatie
- 8.4 Checklist voor een strategie op het gebied van elektronische informatie
- 8.5 Prototype: welke metadata moeten worden ontwikkeld?
- 8.6 Prototype: hoe de goede standaarden te selecteren
- 8.7 Bibliografie
- 8.8 Vervolgactiviteiten van het DLM-Forum

8.1 Terminologie

De volgende definities zijn gebruikt om te komen tot deze richtlijnen. De nationale wetgeving in verschillende landen bevat ook definities die mede in overweging dienen te worden genomen⁸.

Informatie

Een kenniselement dat kan worden verspreid

Gegevenselement:

De weergave van een basis-element informatie in een zodanige vorm dat het mogelijk is dit te verwerken.

Machineleesbare gegevens (machine-readable data):

Gegevens in een vorm die het mogelijk maakt deze terug te zoeken, te bewerken en te communiceren met behulp van een digitale computer.

Document:

Een samenhangende en begrijpelijke hoeveelheid gegevens waarmee een bepaalde redenering wordt weergegeven of een activiteit wordt gerapporteerd.

Database:

Een bepaalde samenhangende hoeveelheid gestructureerde gegevens met de bedoeling informatie te kunnen terugvinden en te bewerken.

Record:

Een samenhangende hoeveelheid gegevens die is vastgelegd op een opslagmedium.

"Een specifieke hoeveelheid vastgelegde informatie die is gegenereerd, verzameld en ontvangen bij de start, de voortgang of het beëindigen van een institutionele of individuele activiteit en die voldoende inhoud, contextuele gegevens en structuur bevat om bewijs te leveren van een bepaalde activiteit (dit is een vertaling van de ICA-definitie: a specific piece of recorded information generated, collected or received in the initiation, conduct or completion of an institutional or individual activity and that comprises sufficient content, context and structure to provide proof or evidence of that activity").

Elektronisch record

Een record waarbij de informatie is vastgelegd in een vorm die bruikbaar is voor het terugzoeken, het bewerken en het communiceren met een digitale computer (ICA: "A record where the information is recorded in a form that is suitable for retrieval, processing and communication by a digital computer").

Medium

Het fysieke materiaal waarop records kunnen worden vastgelegd, waarop zij kunnen worden opgeslagen en waarvan zij kunnen worden teruggehaald.

8.2 Lijst van nog op te lossen problemen

1. De juridische waarde van elektronische informatie

Het concept van een origineel, zoals dit wordt toegepast op papieren documenten, geeft problemen wanneer we dit toepassen voor elektronische informatie. Het is zo gemakkelijk geworden om records te kopiëren dat het origineel niet meer van de kopie te onderscheiden is. Een elektronisch record kan verder nog verbindingen (links) bevatten met andere elektronische records.

⁸ De terminologie die nu volgt, is discutabel. Naar de mening van de vertaler zijn er in Nederland betere begripsformuleringen voorradig. Wij nemen echter deze definities integraal vertaald over, zodat zij eventueel ook kunnen worden geciteerd als afkomstig uit de Guidelines.

Dit probleem kan worden opgelost door encryptie te gebruiken met openbare en privé-sleutels (zie hiervoor sectie 6.3.2 over de data-encryptie). De auteur versleutelt een record met zijn persoonlijke sleutel en stuurt dit naar de server, die bestemd is voor geautoriseerde documenten. Iedereen die de openbare toegangscode voor deze server heeft, kan het record lezen in de wetenschap dat het niet veranderd is sinds de auteur het document op de server heeft geplaatst.

2. De betekenis van terminologie kan van land tot land verschillen

De vertaling van verschillende termen op zich is moeilijk, maar vaak niet voldoende omdat het probleem met de vocabulaire complexer is dan alleen dit aspect. Een oplossing kan zijn om een lijst met termen samen te stellen, uitgaande van de terminologie die in elke taal wordt toegepast.

8.3 Het waarom van standaardisatie

Om vrije concurrentie toe te staan en verzekerd te zijn van transporteerbare gegevens die pakket- en leveranciersonafhankelijk zijn, is het noodzakelijk dat software interfaces heeft en dat producten hierdoor uitwisselbaar zijn. Dit maakt het mogelijk dat verschillende leveranciers producten maken die op elkaar aansluiten, wat leidt tot grotere duurzaamheid van gegevens en toepassingen.

De mate van volwassenheid van de technologie kan variëren tussen merkgebonden producten en open standaards.

De facto standaards: wanneer een product sterk in de markt is vertegenwoordigd, wordt de uitwisselbaarheid met andere applicaties en gegevens gemeten door te refereren aan dat product. De uitwisselbaarheid is dan afhankelijk van de producent, bij voorbeeld Microsoft Word op het gebied van tekstverwerking.

Publiek beschikbare specificaties (PAS): soms werken leidende partijen in de markt als consortium samen om te komen tot een interface standaard. De definitie van een interface maakt het dan mogelijk om uitwisselbare producten te ontwikkelen (bv de X/Open of IETF-specificaties).

De jure standaards. Officiële instanties kunnen overeenkomen om een standaard te hanteren die bepaalde garanties biedt en daarmee een officiële standaard wordt (bij voorbeeld de ISO standaard karaktersets).

Met het volwassener worden van de technologie ontstaat eerst een de facto standaard, daarna Pas en uiteindelijk een de jure standaard.

Er bestaan verschillende officiële organisaties op het gebied van standaardisatie, zoals:

- **ISO**: the International Organisation for Standardisation. Deze is op tal van terreinen actief;
- **IEC**: the International Electrotechnical Committee.

Deze twee organisaties hebben een gezamenlijk comité opgericht om te komen tot standaards op het gebied van het bewerken van informatie.

- **ITU**: the International Telecommunication Unit. Het comité van de ITU-T (nieuwe naam voor het CITT) houdt zich specifiek bezig met standaards op het gebied van telecommunicatie.

Nieuwe standaards passeren verschillende stadia voordat zij definitief worden vastgesteld. Soms wordt een standaard al gebruikt wanneer deze het op één na laatste stadium als ontwerp-standaard (Draft international standard, of DIS) heeft bereikt.

Enkele organisaties die zich bezig houden met standaardisatie zijn:

- **CEN**: European Standardisation Committee (Comité européen de normalisation électrotechnique)
- **Cenelec**: European Electrotechnical Standardisation Committee (Comité européen de normalisation électrotechnique);
- **ETSI**: European Telecommunications Standards Institute

De eerste twee hebben te maken met informatietechnologie. Ze stellen Europese standaards op (Euronorm- EN) en Europese pre-standaards (ENV). Deze komen dikwijls overeen met internationale standaards. Er is echter een verschil in juridische terminologie: onder besluit 87/95/EEC van de Raad van de Europese Gemeenschap van 22 december 1986 is het toepassen van Europese standaards bij aanbestedingen verplicht, terwijl het volgen van internationale standaards facultatief is gesteld.

Naast het vaststellen en ontwikkelen van standaards stellen deze organisaties ook zogenaamde profielen op. Dit zijn sets van standaards met een keuze uit opties om de uitwisselbaarheid te vergemakkelijken. De internationale profielen die geproduceerd worden door de ISO/IEC staan bekend als international standardised profiles (ISP).

Verschillende andere organisaties produceren specificaties, zoals:

- de Open Group publiceert haar specificaties in de X/Open Portability Guide (XPG);
- IETF (Internet Engineering Task Force) produceert Internet specificaties nadat hiertoe een Request for Comments is uitgegeven;
- NIST: het US National Institute for Standards and Technologies ontwikkelt profielen die bekend staan als Federal Information Processing Standards (FIPS);
- Veel andere organisaties werken ook aan een keur van aspecten van standaards op het gebied van informatieverwerking.

De onderstaande tabel geeft de referenties van alle standaards die in deze richtlijnen zijn geciteerd.

Naam	Internationale Standaard of profiel	Europese standaard of profiel	Andere specificatie	Opmerkingen
Opslagmedia (sectie 5.1 en 6.2)				
3½ inch floppy disk	ISO/IEC 9529-1 ISO/IEC 9529-2	EN 29529-1 EN 29529-2		
.....				
½ inch cartridge	ISO 8462-1			
.....				
1 600-bpi tape	ISO/IEC 3788:1976			
.....				
6 250-bpi tape				
.....				
8 mm cassette	ISO/IEC 11319 ISO/IEC 12246			
.....				
DAT-cassette				
.....				
CD-ROM	ISO 9660 ISO 10149			
.....				
TMO				
.....				
DVD				in voorbereiding
.....				
bitmapped grafische afbeeldingen en vector grafieken (sectie 5.2)				
TIFF graphics				
.....				
FIF graphics				
.....				
JPEG graphics				
.....				
Groep 3 faxen	ITU-T groep III			vroeger: CCITT
.....				
Groep 4 faxen	ITU-T groep IV			vroeger: CCITT
.....				
MPEG-1 video				
.....				
MPEG-2 video				
.....				
CGM graphics	ISO 8632			
.....				
CAD graphics				
.....				
GIS graphics				
.....				
Karaktersets (sectie 5.2.3)				
7-bit	ISO 646			
.....				

8-bit west Europa	ISO/IEC 8859-1		
8-bit Grieks	ISO/IEC 8859-7		
Multi-byte	ISO/IEC 10646		
Gestructureerde tekst (sectie 5.2.3)			
SGML	ISO/IEC 8879	EN 28879	
DSSSL	DIS 10179		
SPDL	ISO/IEC 10180		
SDIF	ISO/IEC 9069		
Font information			
Interchange	ISO/IEC 9541		
Standard DTD	ISO/IEC 12083		
HyTime	ISO/IEC 10744		
HTML			W3C HTML 3.0
ODA/ODIFF	ISO 8613	EN 41509	
	FOD 26	EN 41515	
Gegevensformats (sectie 5.2.4)			
EDIFACT	ISO/IEC 9735	EN 29735	
STEP/Express	ISO 10303		
IDEF0&3			IDEF
FIMS			
Uitwisselingsprotocollen (sectie 6.2)			
HTTP			IETF RFC
FTP			IETF RFC
Database query (sectie 6.2)			
HTTP	ISO/IEC 9075		versie II
ISAM			
Encryptie algoritmen (sectie 6.3.2)			
DAS	ISO 8273		
DES			NIST FIPS 46-1; lijkt op DAS
RSA			

8.4 Checklist voor het ontwikkelen van een strategie op het gebied van elektronische informatie

Dit annex geeft een overzicht van vragen waarmee u te maken krijgt bij het definiëren van een strategie voor elektronische informatie. De nummers tussen haakjes verwijzen naar de secties in de Richtlijnen die het onderwerp in kwestie behandelen. Toekomstige revisies zullen cases behandelen waarin ervaringen worden verwerkt.

Niet iedere organisatie zal elke vraag moeten beantwoorden of elk onderwerp willen behandelen. De lijst van onderwerpen die nog niet zijn opgelost, waarover nog geen beslissing is genomen of die zijn verworpen geeft ook belangrijke informatie die onderdeel kan uitmaken van de strategie.

I. Algemene strategie

- A. Bepalen wie bij het project betrokken wordt (4.1). Het samenstellen van een multidisciplinair team.
- B. Lijst van algemene termen en concepten opstellen (8.1; 8.2)
- C. Gebruikseisen vaststellen en controleren (4.1)
- D. Vaststellen van een beleid voor de bepaling van de juridische waarde van records (8.2)
- E. Vaststellen van informatiebeleid en training en opleiding voor de afdelingen (4.1)
- F. Hoe om te gaan met verouderde technologie (media, documenten etc.) (5.1; 5.2)

II. Het beheer van elektronische informatie

- A. Vaststellen van verantwoordelijkheden (voor ieder moment van informatie-overdracht) (3.2)
 - 1. Verantwoordelijkheid voor het beheer van elektronische informatie
 - 2. Verantwoordelijkheid voor het behoud van elektronische informatie
- B. Vaststellen en vastleggen van belangrijke elektronische informatie (4.1; 4.2)
(omvang en afkaderen ten opzichte van andere elektronische informatie)
- C. Beleid op het gebied van contextuele documentatie voor elektronische informatie (2.1; 4.2)
- D. Vaststellen van een schema voor de selectieve bewaring en vernietiging van informatie (4.3.6)
 - 1. Procedure voor elektronische informatie van twijfelachtige afkomst (4.2)
 - 2. Procedure voor de goedkeuring voor vernietiging of wijziging van informatie (4.2)
- E. Vaststellen van regels voor het classificatie-plan (4.1; 4.4)
 - 1. Bouw van een trefwoordensysteem (thesaurus)
- F. Strategie voor de versleuteling (encryptie) van informatie (6.3.2)
 - 1. Vertrouwelijkheid en geheimhouding
 - 2. Waarborgen van de authenticiteit
- G. Beleid op het gebied van het fysieke transport van elektronische informatie (4.5)
 - 1. Lijst van berichten die moeten worden getransporteerd
 - 2. Lijst van controles die voor elke overdracht moeten worden uitgevoerd
 - 3. Bepaling van het medium dat wordt gebruikt voor het fysieke transport, het originele medium waarop de berichten zijn opgeslagen, een beveiligd netwerk etc.

III. Het behoud van elektronische informatie

- A. Opties voor het behoud van originele elektronische documenten (4.2)
Uitsluitend elektronische records, of ook papieren kopieën
- B. Opties voor media voor gegevensopslag (5.1)
 - 1. Korte-termijn opslag
 - 2. Lange-termijn opslag
 - 3. Bewaarcondities (temperatuur, luchtvochtigheid, bepalen termijn waarbinnen informatie op een ander medium moet worden overgebracht etc.)
- C. De opties om opgeslagen bestanden te bewaren in verschillende formats (5.2)
 - 1. Bitmapped graphics (5.2.1)
 - a. Type compressietechniek (5.2.1)
 - b. Behoud van faxen (4.3.4; 5.2.1)
 - 2. Vector graphics (5.2.2)
 - a. Specifieke grafische bestanden (CAD, GIS)

3. Audio, video en multimediatebestanden (5.2.1)
 4. Tekstbestanden (5.2.3)
 - a. karaktersets (standaard, geaccepteerd)
 - b. gestructureerde tekstformats
 - c. het al dan niet behouden van de lay-out.
 5. Gegevensbestanden (5.2.4)
 - a. Databases (al dan niet leveranciergebonden)
Uitwisselbaarheid en het beheer van bestanden
 - b. Gestructureerde tekstformats
 - c. Lay-out al dan niet bewaren
 6. Programma's (5.2.4)
- D. Beleid voor het bewaren van verouderde systemen en programmatuur (4.3.4)
Onderhoud, documentatie, kennis over de systemen etc.

IV. Strategie voor het behoud van papieren documenten

- A. Opties voor het behoud van papieren documenten (4.2)
Het vaststellen welke documenten voor scanning in aanmerking komen
- B. Kwaliteitskaart voor documenten die gescand moeten worden (4.3.3)
- C. Strategie voor optische karakterherkenning (4.3.4)
 1. Opties bij gebruik van OCR
 2. OCR-procedure (4.3.4)
- D. Strategie voor het vectoriseren van grafische afbeeldingen (4.3.4)
- E. Procedure om verschillende elementen van elektronische informatie samen te voegen (tekst, vectorgrafieken, bitmaps, etc). (4.3.4)

V. Het converteren of behouden van gegevensformats (4.3.3)

- A. Beleid voor het documenteren van formats van gegevens die worden gebruikt bij systemen en software (2.4.2)
- B. Opties voor de conversie en het behoud van verouderde formats (4.3.6)
- C. Opties voor het toevoegen van structuur aan herstelde of teruggehaalde tekst en gegevens (4.3.4)
- D. Opties voor het anonimiseren van een bestand (6.3.1)
 1. Procedure voor het waarborgen van de anonimiteit
- E. Procedure voor het analyseren van verlies aan informatie door het toepassen van conversie (4.3.3)

VI. Toegang en gebruik van gegevens

- A. Beleid voor de toegang tot gegevens (inzage, printautorisaties) (6.3.1)
- B. Informatie beschikbaar maken (passieve verspreiding) (6.4)
 1. Elektronische leeskamer (6.4)
 - a. Standaards voor inzage (6.2)
 2. Het kopiëren van elektronische informatie voor distributie (6.4)
 - a. Medium waarop wordt gekopieerd (6.2)
 3. Het kopiëren van elektronische informatie in een bepaald formaat om het daarna te kunnen verspreiden (6.4)
 - a. Media, protocollen, talen en formats (6.2)
 - b. Internet strategie (6.2)
 4. Toegangsmodel voor het automatisch genereren van een gebruiksvriendelijk formaat (6.4)
 - a. Media. Protocollen. Talen en formats (6.2)
 - b. Internet strategie (6.2)
- C. Bevorderen van de toegankelijkheid (actieve verspreiding) (6.4)

8.5 Prototype: welke metadata zou moeten worden toegevoegd aan een record?

Dit annex geeft een voorbeeld van mogelijke metadata, gebaseerd op het 'Dublin Core Metadata Proposal' van December 1996. Meer informatie kunt u vinden bij http://www.purl.org/metadata/dublin_core.

Buiten dit prototype zijn er natuurlijk nog andere mogelijkheden. Het kan de lezer echter helpen bij het vaststellen van een eigen strategie.

De 15 elementen zijn alle optioneel en uit te breiden. Zij beschrijven de context van een specifieke bron.

TITEL

de naam die aan een bepaalde bron is gegeven door de auteur of de ontwerper/afzender

ONTWERPER (auteur of steller)

De persoon of organisatie die primair verantwoordelijk is voor de intellectuele inhoud van de bron. Een auteur is dit bij voorbeeld bij een geschreven document; de artiest, fotograaf of illustrator bij visuele documenten.

ONDERWERP (subject en trefwoorden)

Het onderwerp waarover de bron handelt, of trefwoorden of zinsneden die de inhoud van de bron beschrijven. De bedoeling van de specificatie van dit element is om het gebruik van gecontroleerde trefwoorden en tabellen te bevorderen. Dit element kan zowel classificatiegegevens bevatten (bij voorbeeld een classificatienummer van de Basis-Archiefcode of Dewey-classificatienummers) als thesaurustermen.

BESCHRIJVING

Een tekstuele beschrijving van de inhoud van de bron, bij een document bestaat dit uit een uittreksel en bij visuele bronnen uit een inhoudsomschrijving. Toekomstige metadataverzamelingen kunnen mogelijk beschrijvingen bevatten van bronnen die direct door de computer worden toegevoegd. Indien deze informatie niet in het huidige netwerksysteem is in te voegen, dan dient een bepaald veld een verwijzing (link) te bevatten naar zo een beschrijving. Dat is beter dan de beschrijving zelf bij de metadata op te nemen.

UITGEVER

Hier wordt beschreven wie verantwoordelijk is voor het beschikbaar stellen van de bron in zijn huidige vorm, zoals bij voorbeeld een uitgever, een universiteitsafdeling of een bedrijf. De bedoeling van deze specificatie is om de entiteit te kunnen aangeven die toegang biedt tot de bron.

MEDEWERKERS

Personen of organisaties in aanvulling op het ontwerpersveld die een belangrijke intellectuele bijdrage hebben geleverd aan de bron, maar waarvan de bijdrage secundair is aan individuen of entiteiten die zijn genoemd in het ontwerpersveld (bij voorbeeld: redacteur, vertaler, illustrator)

DATUM

De datum waarop de bron in zijn huidige vorm beschikbaar is gemaakt. Aanbevolen wordt om een acht-cijferig nummer te gebruiken in de vorm JJJJMMDD (jaar-maand-dag) zoals gedefinieerd in ANSI X3.30-1985. Volgens dit schema zou het datumelement voor de dag van publicatie van de Guidelines worden geschreven als 19961203, of 3 december 1996. Er zijn andere mogelijkheden maar indien een bepaald model wordt gebruikt zou de organisatie eraan vast moeten houden om verwarring over datums te voorkomen.

TYPE

De categorie van de bron, zoals: home page, novelle, brochure, gedicht, werkdocument, technisch rapport, essay, woordenboek. Het type bron zou moeten worden gekozen uit een limitatieve lijst van typen. Een voorlopige set van deze types kan worden gevonden op het volgende URL:

<http://www.roads.lut.ac.uk/Metadata/DC-ObjectTypes.html>

FORMAAT/FORMAT

De gegevensweergave van de bron, zoals tekst/html, ASCII, PostScript, executable bestand, JPEG-image. De bedoeling van het weergeven van dit element is om informatie te leveren die nodig is om mensen of machines beslissingen te laten nemen over de gecodeerde gegevens (welke hard- of software nodig is om de gegevens weer te geven bij voorbeeld).

Zoals bij de bepaling van het type van de bron dient het format te worden toegekend aan de hand van een lijst die in het systeem is opgenomen. In principe kunnen formats fysieke media bevatten, zoals boeken, seriewerken, of andere niet-elektronische media.

UNIEK KENMERK (identifier)

Een bepaalde gegevensstring of een nummer dat wordt gebruikt om een bepaalde bron een uniek kenmerk te geven, waardoor het zich van andere bronnen onderscheidt. Voorbeelden bij bronnen op het gebied van netwerken houden in URLs en URNs, wanneer deze zijn ingevoerd. Andere identifiers die uniek zijn op de wereld zijn bij voorbeeld het International Standard Book Number.

BRON (source)

Het werk, zowel gedrukt als elektronisch, waaraan de beschreven bron is ontleend. Een html-weergave van een sonnet van Shakespeare kan bij voorbeeld de papieren versie identificeren van het sonnet, waarvan de elektronische versie afkomstig is.

TAAL

Taal van de intellectuele inhoud van de bron. Waar mogelijk en praktisch zou de inhoud van dit veld overeen dienen te komen met de Z39.53 drie-karaktercode voor geschreven taal. Zie hiervoor

<http://www.sil.org/sgml/nisoLang3-1994.html>

RELATIE

Relatie tot andere bronnen. De bedoeling van het specificeren van dit element is om een mogelijkheid te scheppen om relaties aan te geven tussen bronnen die een formele relatie hebben tot andere bronnen, maar die wel als afzonderlijke bron bestaan. Bij voorbeeld, foto's in een boek, hoofdstukken in een boek of onderdelen in een collectie. Een formele specificatie van hiervan wordt momenteel ontwikkeld. Daarom is dit onderdeel momenteel experimenteel.

BEREIK (Coverage)

De ruimtelijke en tijdelijke karakteristieken van de bron. Ook dit onderdeel is in ontwikkeling.

RECHTEN (Rights)

Een link naar een copyright notitie, naar een verklaring van de rechten, naar een dienst die informatie geeft over voorwaarden waaraan moet worden voldaan wil men toegang te krijgen tot de bron. Onder bewerking.

8.6 Prototype: hoe de juiste standaards te kiezen

Dit annex levert een voorgestelde standaard voor elk gegevenstype. Het is mogelijk om welke keuze dan ook te maken, gebaseerd op de informatie uit deze richtlijnen. Het is echter goed gebruik om een aanbevolen standaard te definiëren. Ook is het mogelijk om een lijst van aanvaardbare standaards per bestandstype te noemen. Uiteraard zijn er ook andere mogelijkheden die de lezer kunnen helpen om tot een eigen strategie te komen.

Gegevenstype	Aanbevolen standaard	Opmerkingen
Karaktersets	ISO/IEC 8859-1	Voor Westeuropese landen is de Unicode een andere mogelijkheid (ISO/IEC 10646) als aanvullende karakters nodig zijn.
.....
Gestructureerde tekst	SGML	
.....
Bitmap graphics	JPEG	
.....
Faxen	ITU-T groep III	
.....
Vector grafieken	CGM	
.....
Audio en video	MPEG II	
.....
CAD/CAM	STEP	
.....
Accounting/facturen	EDIFACT	
.....
Andere databases	Plat bestand, comma scheidingstekens. Er bestaat geen standaard database format. Het platte bestand levert behoud op lange termijn, indien de database goed is gedocumenteerd.	
.....
Versleutelde bestandsprogramma's	RSA Bron- of Pc-uitwisselbare versie Ook hier geen standaard voor het samengestelde programma. De Java Byte Code, die platformonafhankelijk is kan behulpzaam zijn bij het behoud van Java-applicaties.	
.....
Media voor lange-termijn bewaring	DVD. Alhoewel DVD betrekkelijk nieuw is, is de verwachting dat het in de toekomst ruim gebruikt zal worden. De grote opslagcapaciteit en het gebruiksgemak kunnen helpen om DVD te maken tot een archiefmedium. Er zijn al verschillende organisaties die een keuze hebben gemaakt voor het medium. Het is aan te bevelen om slechts een medium te kiezen voor de toekomst, of de hoeveelheid beperkt te houden om het gebruik in de toekomst beheersbaar te houden.	
.....

8.7 Bibliografie

Deze bibliografie geeft uitsluitend enkele referentiewerken die gebruikt zijn bij het samenstellen van deze richtlijnen. Een grote hoeveelheid interne documenten van nationale en EU-organisaties zijn ook geraadpleegd, maar niet hier weergegeven.

Archives in the European Union, Report of de Group of Experts on the Coordination of Archives, European Commission, Secretariat-General, Brussels. Luxembourg, 1994
ISBN 92-826-8233-1, Cat. no CM-83-94-741-FR-1

Proceedings of the DLM-Forum on electronic Records, Brussels 18 to 20 December 1996, INSAR-European Archives News, Supplement II, 1997, EUR-OP, Luxembourg, 1997, p. 376, ISBN 92-828-0111-X. Cat. No CM-AC-97-S01-EN-C (DE, EN, FR)
On Internet: <http://www.echo.lu/dlm/en/home.html>

European procurement handbook for open systems (EPHOS) Reference: EUR 14021

ICA Guide on electronic records, international Council on Archives, Committee on Electronic Records, Paris, 1997 (EN, FR).
On Internet: <http://www.archives.ca/ica/p-er/english.html>

B. Bauwens, F. Evenepoel, J. Engelen, 'Standardisation as a prerequisite for accessibility of electronic text information for persons who cannot use printed material' *IEEE transactions on rehabilitation engineering*, Vol. 3, No 1, pp. 84-89, 1995

Chr. Reeves, T. Wesley, *Guidelines for accessible Web page design*, Brochure published by The Harmony consortium (1997). Also available on the Web
<http://www.esat.kuleuven.ac.be/teo/harmony/guidelines>

8.8 De DLM-vervolgactiviteiten

De vervolgactiviteiten van het DLM-Forum werden vastgelegd in de zogenaamde "Tien Punten". Deze zijn gerelateerd aan de gehele productie- en onderhoudsketen van elektronische records. Zij worden gecontroleerd door het DLM-Monitoring Committee dat werd ingesteld als één van de vervolgmaatregelen. Deze punten houden in:

1. Publicatie van het verslag van het DLM-Forum in de werktalen van het Forum: Duits, Engels en Frans. Dit leverde een boekwerk op van 376 pagina's, dat verkrijgbaar is bij het Office for Official Publications of the European Communities, 1997, ISBN 92-828-0111-X.
2. Het uitvoeren van een allesomvattende studie naar de relatie tussen overheidsadministraties en archieven;
3. Het promoten van de beste manier om elektronische documenten te beheren door het uitbrengen en actueel houden van Guidelines;
4. Het verzamelen en verspreiden van informatie over onder handen zijnde projecten en andere relevante acties op het gebied van elektronisch documentenbeheer;
5. Het instellen van nationale contactfunctionarissen, de zogenaamde Focal Points die het nationale correspondentie-adres van het Forum vormen door informatie-uitwisseling, het vaststellen van functionele vereisten voor elektronisch documentenbeheer en records management in de overheidsadministratie en de private sector, door samenwerking tussen verschillende disciplines op landelijk niveau te verbeteren en zo maximaal profijt te behalen op het gebied van standaardisatie, de creatie, de verzending, migratie en gemeenschappelijke toegang tot elektronische records;
6. De lidstaten in de EU worden opgeroepen om hun betrokkenheid in nationale en internationale organisaties die standaardisatie bevorderen te vergroten (ISO, CENELEC, CEN). De multidisciplinaire aanpak zou in het bijzonder moeten worden benadrukt.
7. Op basis van de ervaringen in de verschillende lidstaten zou een opleidingsprogramma moeten worden opgesteld over elektronisch documentenbeheer voor beheerders en archivariissen
8. De Europese Commissie zal in samenwerking met de lidstaten de juridische consequenties op het gebied van elektronisch recordsmanagement vastleggen. Een rapport of studie over dit onderwerp dient aan de lidstaten te worden gepresenteerd.
9. De lidstaten dienen samen met de Europese Commissie de directe toegang tot informatie met behulp van nieuwe technologie voor de burger via archiefdiensten en andere informatieverwerkende lichamen te stimuleren.
10. Om de follow-up van het DLM-Forum voor elektronische records te bewaken wordt een multidisciplinaire werkgroep ingesteld die periodiek rapporteert.