

Ministerie van OCW
Facilitair Bedrijf
Werkprogramma 'Digitaal werken bij OCW
1 augustus 2003



De Bron ontsloten

*Haalbaarheidsonderzoek
auteurs-gegenereerde metadata*

Verantwoording

Dit rapport is het eindresultaat van het haalbaarheidsonderzoek van 'Auteurs-gegenereerde metadata' bij het ministerie van Onderwijs, Cultuur en Wetenschap. Het onderzoek heeft plaatsgevonden in opdracht van Chantal Zeegers, programmamanager 'Digitaal werken bij OCW 2003-2006' en is uitgevoerd door Wil Rombout, werkzaam als adviseur Informatiediensten bij de directie Facilitair Bedrijf (FacB).

Electronische bereikbaarheid.....

De basisschool van je zoontje kampt met overlast van hangjongeren. Het schoolplein is vies en onveilig. Het school-bestuur probeert er van alles aan te doen, maar het blijft dweilen met de kraan open. Het is tijd voor actie! Je wilt weten of de overheid dit probleem serieus oppakt en –mocht dat niet zo zijn- de overheid attenderen op de ernst van het probleem.

E-government, elektronische bereikbaarheid...je besluit de overheid langs digitale weg te bevragen. Je surft op internet naar Postbus 51, je zoekt met de zoekmachine op: 'onveiligheid' 'schoolpleinen' en 'overlast'. Zonder dat je het merkt kom je terecht op de website van het ministerie van Onderwijs, Cultuur en Wetenschap. Je vindt daar een hele rubriek met allerlei stukken over scholen, fysieke en sociale onveiligheid, degelijke speeltoestellen en de wijze van bekostiging van de basisschool. Beleidsstukken, kamervragen een WOB-verzoek over dit onderwerp: alles is eenvoudig terug te vinden en te raadplegen.

Toch wil je weten of je mee kan praten over dit onderwerp en zo ja, met wie dan. Gelukkig kun je via de website deze vraag naar het ministerie mailen. En dat doe je dan ook.

Een dag later ontvang je een mailtje van het ministerie met daarin een bevestiging dat de vraag in behandeling is genomen door de directie Primair Onderwijs met de naam en het telefoonnummer van de contactpersoon. Ter referentie wordt er een code meegestuurd waarmee eenvoudig ingelogd kan worden op de website van het ministerie OCW om de status van afhandeling te volgen.

Je voelt je in ieder geval serieus genomen. Je bent benieuwd naar het antwoord.

Toekomstmuziek?

Het kabinet wil dat vanaf 2006 50 % van alle overheidsdienstverlening elektronisch verloopt. Informatietechnologie geeft ons immers de mogelijkheid ons werk sneller te doen en onze beleidsproducten en onze werkprocessen trans-paranter te maken.

Het gebruik van metadata blijkt een kritische succesfactor. Metadata: gegevens over gegevens. De taal waarmee je informatie kunt terugvinden. Als je bijvoorbeeld zoekt op 'onveiligheid' en 'schoolplein' vind je ook informatie over veiligheidsvignetten.

Voor de burger als informatiezoeker of subsidie-aanvrager is het van belang dat de overheid één taal spreekt, één metadata-standaard hanteert. Daarom houden interdepartementale samenwerkingsverbanden als Interlab en de Commissie Nieuwe Media zich bezig met standaardisatie. Ook in het buitenland denkt men er zo over. Zo is in Engeland één standaard verplicht gesteld voor de publicatie van alle overheidsinformatie.

Wil Rombout geeft in zijn rapport een helder overzicht van theorie, toegevoegde waarde, knelpunten en haalbaarheid van het gebruik van metadata. Dit rapport wordt gepubliceerd te midden van een levendige discussie over de functie van de toekomstige documentaire informatievoorziening. Ik hoop dat dit rapport helpt inzicht te verwerven in deze discussie en een extra impuls geeft aan de goede initiatieven op dit vlak. Ik wil de auteur, Wil Rombout, hartelijk danken voor zijn bijdrage aan dit complexe onderwerp.

Diane Keizer-Mastenbroek
Plaatsvervangend Secretaris-Generaal van het ministerie van Onderwijs, Cultuur en Wetenschap

Managementsamenvatting	7
1. Inleiding	11
1.1. Leeswijzer	11
2. De aanleiding en context van het project	13
2.1. Organisatiebeschrijving	14
2.1.1. <i>Missie van OCW</i>	14
2.1.2. <i>Organisatiestructuur</i>	14
2.2. Knelpuntenanalyse (oriëntatiefase)	16
2.3. Toename digitale documentaire informatie	17
2.3.1. <i>Intern</i>	17
2.3.2. <i>Extern</i>	17
2.4. Tendensen in het overheidsbeleid	17
2.4.1. <i>Transparante overheid</i>	17
2.4.2. <i>Kwaliteitsslag OCW</i>	18
2.4.3. <i>Taakstelling OCW</i>	19
2.5. Initiatieven en ervaringen tot nu toe	19
2.5.1. <i>Standaarden en richtlijnen</i>	19
2.5.2. <i>Onderzoek naar metadata (desk research)</i>	19
2.5.3. <i>Intern onderzoek</i>	20
2.5.4. <i>Extern onderzoek</i>	20
3. Opdrachtformulering	23
3.1. Achtergrond project	23
3.1.1. <i>Project doelstellingen</i>	24
3.1.2. <i>Omvang van het project</i>	24
3.1.3. <i>Projectresultaten</i>	25
3.1.4. <i>Randvoorwaarden & beperkingen</i>	25
3.1.5. <i>Business Case</i>	25
3.1.6. <i>Kwaliteitsverwachtingen</i>	26
4. Diepteonderzoek	27
4.1. Onderzoeksdoelstelling	27
4.2. Onderzoeksvraagstelling	28
4.3. Inrichting pilots bij de directies	29
4.3.1. <i>Evaluatie van de pilot</i>	30

4.4.	Interviewronde bij pilotmedewerkers en anderen	33
4.4.1.	<i>Resultaten van de interviewronde (samenvatting)</i>	35
4.5.	Zoekmachine Verity	39
4.6.	Kwantitatieve informatie: zoekgedrag, volume G:\schijven, Intranet/OCW-plein	40
4.7.	Externe oriëntatie bij ministerie AZ	41
4.8.	Webcontent-managementsysteem	42
4.9.	Literatuuronderzoek	43
4.9.1.	<i>Metadata</i>	44
4.9.2.	<i>Toepassing van metadata</i>	44
4.9.3.	<i>Metadata Formats</i>	46
4.9.4.	<i>Dublin Core metadatastandaard</i>	46
4.9.5.	<i>Trefwoordsysteem</i>	51
4.9.6.	<i>De trefwoordenlijst van OCW</i>	51
4.9.7.	<i>Thesaurus</i>	52
4.9.8.	<i>De OCW - ADIONthesaurus</i>	54
4.9.9.	<i>Rubrieken</i>	55
4.9.10.	<i>Metadata-repository</i>	55
5.	Conclusies en aanbevelingen	59
5.1.	Interne en externe informatievoorziening	59
5.1.1.	<i>Gemeenschappelijk dataserver (G:\schijf)</i>	60
5.1.2.	<i>Organisatie van de informatievoorziening</i>	61
5.1.3.	<i>Instrumentarium</i>	62
5.1.4.	<i>Trefwoordsysteem versus thesaurus</i>	63
5.2.	Metadata-repository	63
5.3.	Implementatiestrategie	64
5.3.1.	<i>Communicatie</i>	65
6.	De bron ontsloten met Dublin Core metadata	67
6.1.	Voorbeeld uitwerking metadata (html en Dublin Core metadatatags)	68
7.	Bronvermeldingen	71
7.1.	Literatuur en publicaties	71
7.2.	Internetbronnen	72

Managementsamenvatting

Het rapport is het eindresultaat van een haalbaarheidsonderzoek dat als onderdeel van het departementsbrede werk-programma 'Digitaal werken bij OCW 2003-2006' bij het ministerie van OCW is uitgevoerd.

Het werkprogramma heeft als doelstelling dat medewerkers hun taken beter moeten kunnen uitvoeren en dat de informatievoorziening hen daar efficiënt in ondersteunt. Als uiteindelijk resultaat betekent dit, dat alle informatie met één 'user-interface' plaats- en tijdonafhankelijk benaderbaar moet zijn. Om dat te kunnen bereiken is een eenduidige ontsluiting van de documentaire informatie noodzakelijk.

Onderzoeksmethodiek

In dit kader is een haalbaarheidsonderzoek uitgevoerd naar de mogelijkheden van ontsluiting bij de bron (auteurs-gegenereerd). Daarvoor is een knelpuntenanalyse uitgevoerd en zijn de belangrijkste probleemgebieden, de tendensen in het overheidsbeleid en de initiatieven en ervaringen met betrekking tot de informatievoorziening tot nu toe beschreven.

Daarnaast zijn een aantal deelonderzoeken uitgevoerd om de benodigde kennis en inzichten te vergaren over bestaande Storage & Retrievaltechnieken (S&R). Zo is er met gebruikers nieuwe invoer- en zoekfunctionaliteit voor metadata ontwikkeld en getest. Er is literatuuronderzoek gedaan naar de rol en functie van standaard S&R-technieken en zijn medewerkers uit verschillende doelgroepen uit verschillende disciplines geïnterviewd.

Externe informatievoorziening

Beschikbaarheid en beschikbaarstelling zijn de kernbegrippen in de wereld van de documentaire informatievoorziening. In toenemende mate worden er digitale (eind)producten geproduceerd. Deze digitale documenten zijn van belang voor het afleggen van verantwoording, de reconstructie van het gevoerde beleid en op termijn voor het cultureel erfgoed.

Van de overheid wordt ook verwacht dat zij op een efficiënte (met minder mensen) en transparante wijze haar taken voor de maatschappij uitvoert. Dit staat hoog op de politieke agenda.

Zowel de OCW-medewerker als een burger moet makkelijk en snel kunnen beschikken over alle gewenste informatie. Uiteraard is de informatie betrouwbaar en voorzien van actuele statuskenmerken. Hier is een goed georganiseerde informatievoorziening voor nodig.

Deze ontwikkelingen maken het noodzakelijk dat de toegankelijkheid en beschikbaarheid van de documentaire informatie anders georganiseerd moeten worden.

De inrichting en ontsluiting van relevante informatiesystemen is daarom actueel onderwerp van gesprek. Er is nadrukkelijker aandacht voor de ontwikkeling van instrumenten voor kennisdeling en ontsluiting van alle beschikbare relevante bronnen.

De functie van metadata in dit alles is evident want ontsluiten is metadata. Het is belangrijk om aan te sluiten bij reeds ontwikkelde standaarden voor ontsluiting en beschikkingstelling van documentaire informatie. Voor de externe informatievoorziening kan de internationale metadatastandaard van de Dublin Core Initiative¹ een belangrijke rol vervullen. De internationale ontwikkelingen bevestigen dit.

Maar het ministerie staat niet op zich zelf. Zij is als onderdeel van de rijksoverheid een schakel in de keten. Een gemeenschappelijke rijksoverheidsbrede keuze voor een metadatastandaard is hierdoor gewenst, maar helaas niet beschikbaar. OCW kan, door de interdepartementale samenwerking op dit gebied te intensiveren, hier een belangrijke impuls aan geven.

Interne informatievoorziening

Voor de interne informatievoorziening is de functie 'Zoeken en vinden' de functie waar het allemaal om gaat. Het is van belang om in de immer groeiende omvang van digitale documenten sneller en preciezer de relevante documenten te vinden.

De pas geïmplementeerde ordeningsstructuren vertonen al weer dezelfde problemen als voor de herstructurering van de gemeenschappelijke dataservers (G:\-schijven). De opslag van de documenten is niet eenduidig en er worden weer veel dezelfde documenten op verschillende plekken op de G-schijf aangetroffen. Bestandbenamingen worden gebruikt om de toegankelijkheid te vergroten maar dat is erg persoonsgebonden. Tijd blijkt hier een belangrijke rol in te spelen. Het ontbreken of niet uitvoeren van (de)centrale beheerprocedures en -richtlijnen completeert de string aan problemen. Kennis gaat verloren of dreigt verloren te gaan.

Gezien de groeiende hoeveelheid digitale documenten is het organisatorisch wenselijk om het toekennen van metadata bij de maker (de bron) te leggen. Het (achteraf) toekennen van metadata door bijvoorbeeld DIV-afdelingen zou een te groot beslag op capaciteit leggen. Het kiezen voor toekennen aan de bron vereist echter wel voorzieningen en instructies waarmee iedereen overweg moet kunnen.

Binnen de organisatie worden verschillende kunsttalen (trefwoordensysteem en thesaurus) voor de ontsluiting van de documentaire informatie gebruikt. Hoewel de thesaurus op dit moment het minst in de belangstelling staat is het een ontsluitingsinstrument dat bij uitstek geschikt is voor het gebruik in een geautomatiseerde documentaire omgeving. De verschillende representatiemogelijkheden bieden goede mogelijkheden om gebruikers te helpen bij het vinden van relevante informatie. Hier wordt niet alleen de OCW-medewerker bedoeld maar ook de informatiezoekende burger. Ook de multilinguïstische mogelijkheden (meertaligheid) zijn in het licht van de Europese éénwording een belangrijk voordeel.

¹

ISO 15836:2003(E), Information and documentation - The Dublin Core metadata element set.

OCW beschikt al over voldoende technische hulpmiddelen (sjabloontechologie Smart Documents, Verity-zoekmachine, etc.) om de eerste stappen in het komende metadata-beleid te zetten. Er lijkt voldoende kennis en kunde in huis zijn om een toegankelijke digitale informatievoorziening te realiseren.

Implementatiestrategie

Het kiezen van de juiste implementatiestrategie (stap voor stap) en een juiste communicatie zijn belangrijke succesfactoren. Door op het juiste moment, op passende wijze, tussen alle betrokkenen een heldere communicatie tot stand te brengen, wordt de kans op het welslagen van de innovatie groter. De veranderingsbereidheid wordt vergroot door de gebruikers in een vroeg stadium bij het proces te betrekken.

Conclusie en aanbevelingen

Auteurs-gegenereerde metadata is geen 'utopie' is de eindconclusie van het onderzoek mits aan een aantal voorwaarden wordt voldaan. Belangrijke aanbevelingen hiervoor zijn:

- ⋮ sluit aan bij internationaal geaccepteerde standaarden zoals de metadatastandaard van de Dublin Core Initiative;
- ⋮ intensiveer interdepartementale samenwerking en aansturing m.b.t. de ontsluiting van overheidsinformatie;
- ⋮ draag er zorg voor dat de toekenning van metadata zoveel mogelijk geautomatiseerd kan plaats vinden;
- ⋮ doe nader onderzoek naar de toegevoegde waarde van een metadata-repository als ontsluitings-instrumentarium en als zoekleiding voor de interne en externe informatievoorziening.

1 Inleiding

OCW bevindt zich in de fase 'ochtendgloren' van het werkprogramma 'Digitaal werken bij OCW'. Een werkprogramma met als doelstelling dat medewerkers hun taken beter moeten kunnen uitvoeren en dat de informatievoorziening hen daar efficiënt in ondersteunt. Als uiteindelijk resultaat betekent dit, dat alle informatie met één 'user-interface' plaats- en tijdonafhankelijk benaderbaar moet zijn. Om dat te kunnen bereiken is een eenduidige ontsluiting van de documentaire informatie noodzakelijk.

De discussie over de rol en het karakter van de toekomstige documentaire informatievoorziening (DIV) speelt daarom dagelijks. Een informatievoorziening die in eerste aanzet zowel toegang tot elektronische informatie-bronnen biedt als toegang tot de papieren informatiebronnen (de hybride DIV). Hoe die omgeving er over vier jaar uit zal zien is niet te voorspellen. Maar één ding staat vast: de rol die digitale informatie in de documentaire informatievoorziening in de naaste toekomst speelt, zal groter en groter worden.

Naast de verandering in het medium (van papier naar digitaal) waarmee de primaire processen van de documentaire informatievoorziening belast worden, vinden er ook ontwikkelingen plaats in de rol van de DIV. Deze zal steeds sterker verweven raken met de primaire processen van de OCW-organisatie. De documentaire informatievoorziening in traditionele zin wordt doorgaans uitgevoerd door professionals met een archivistische en / of documentaire achtergrond. Daarbij ligt een groot accent op de nabewerking van de informatie. Met de groei van het digitaal communicatieverkeer verschuift het accent in toenemende mate van de achterkant van het proces naar de voorkant. Al bij de ontvangst of creatie van (nieuwe) kennis zal de toegankelijkheid en beschikbaarheid geborgd moeten worden.

1.1 Leeswijzer

In hoofdstuk 2 worden de organisatie van het kerndepartement OCW en de resultaten van de knelpuntenanalyse beschreven. In deze knelpuntenanalyse worden de belangrijkste probleemgebieden, de tendensen in het overheidsbeleid en de initiatieven en ervaringen tot nu toe beschreven. Hoofdstuk 3 beschrijft de achtergrond van digitaliseringstraject en de plaats die dit onderzoek daar in inneemt en worden de projectdoelstellingen, projectresultaten en de kwaliteitsverwachtingen van de klanten uiteen gezet.

Aan de hand van een onderzoeksdoelstelling is in hoofdstuk 4 de onderzoeksvraagstelling (de centrale vragen) gedefinieerd. Om de benodigde kennis en inzichten te vergaren zijn een aantal deelonderzoeken uitgevoerd. Zo is er gedurende een periode van twee maanden nieuwe invoer- en zoekfunctionaliteit voor metadata met gebruikers ontwikkeld en getest. Daarnaast is er literatuuronderzoek gedaan en zijn 20 medewerkers uit de verschillende doelgroepen van het werkprogramma 'Digitaal werken bij OCW' geïnterviewd. De resultaten van deze deelonderzoeken zijn in hoofdstuk 4 verder uitgewerkt.

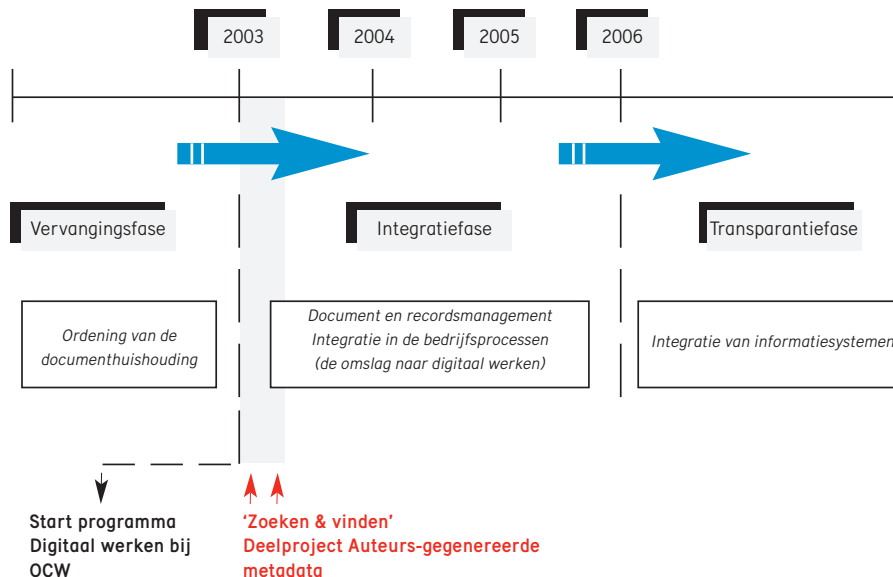
De conclusies en aanbevelingen zijn weergegeven in hoofdstuk 5.

2 De aanleiding en context van het project

Dit project is onderdeel van het departementsbrede werkprogramma Digitaal werken bij OCW. Een belangrijk element van het werkprogramma 'Digitaal werken' is dat medewerkers hun taken beter kunnen uitvoeren en dat de informatievoorziening hen daar efficiënt in ondersteunt. Zij moeten daartoe documentaire informatie op een goede en toegankelijke wijze kunnen opslaan en vooral ook terug kunnen vinden.

Als uiteindelijk resultaat betekent dit, dat alle informatie met één user-interface plaats- en tijd-onafhankelijk benaderbaar moet zijn – ongeacht de plek van opslag – via op elkaar afgestemde en marktconforme 'Storage & retrievaltechnieken'. Om dat te kunnen bereiken is een eenduidige ontsluiting van de documentaire informatie noodzakelijk.

In onderstaand figuur is schematisch de positie van het project 'Zoeken en Vinden' in het werkprogramma 'Digitaal werken bij OCW' weergegeven.



2.1 Organisatiebeschrijving

2.1.1 Missie van OCW

OCW werkt voor mensen. Vanuit dat besef is de OCW-missie ontstaan:

'OCW biedt mensen kansen hun talenten te ontdekken, te ontwikkelen en te gebruiken'.

Diverse groepen OCW-medewerkers hebben met elkaar dit gemeenschappelijke doel geformuleerd en hebben in de toelichting ook de taak en werkhouding verwoord:

'Kennis, kunde en creativiteit zijn essentieel voor de manier waarop mensen met elkaar samenleven en zich voorbereiden op hun toekomst. Daarom stellen wij, als OCW, mensen in de samenleving in staat te ontdekken wat ze kunnen en die vermogens verder te ontplooien en te gebruiken. Bovendien zetten wij ons in voor het bewaren, doorgeven en toepassen van de waardevolle resultaten die dat oplevert. Dat is creatief, inventief en veelzijdig werk dat een open, betrokken en collegiale manier van samenwerken vergt. Met oog voor de politieke verhoudingen spelen wij in op ontwikkelingen in de samenleving. Wij doen dat met respect voor de vrijheden die nodig zijn voor goed onderwijs, bloeiende cultuur en excellente wetenschap.'

2.1.2 Organisatiestructuur

De minister en de twee staatssecretarissen geven het ministerie van OCW politiek-strategisch leiding. De Bestuursraad (secretaris-generaal, plaatsvervangend secretaris-generaal en directeuren-generaal) zorgt ervoor dat de bewindspersonen beleidsvoorstellen krijgen voorgelegd die passen in de politiekstrategische kaders en goed uitvoerbaar zijn. Binnen OCW werken met elkaar samen:

- ⌘ een bestuursdepartement met 25 directies;
- ⌘ drie rijksdiensten;
- ⌘ twee agentschappen;
- ⌘ twee inspecties.

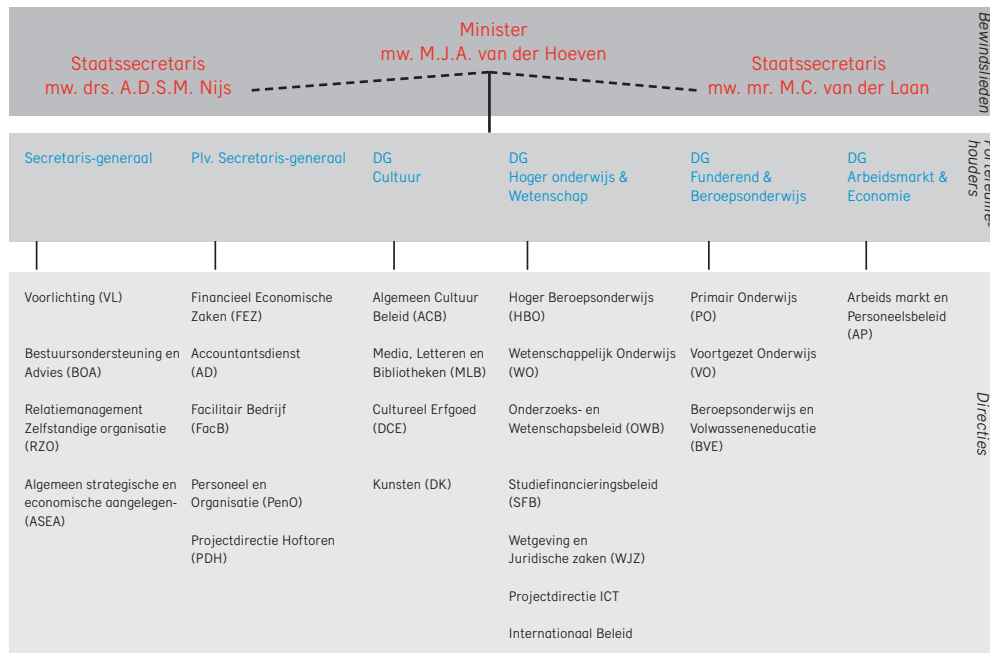
Ook horen bij OCW drie adviesraden: Onderwijsraad, Raad voor Cultuur en de Adviesraad voor Wetenschap- en technologiebeleid.

Beleid en uitvoering

Tien directies ontwikkelen beleid voor bepaalde onderwijssectoren, cultuursectoren of voor wetenschap. Daarnaast zijn er directies die voor het hele gebied aspecten onder hun hoede hebben, zoals de directies ICT, Wetgeving en Juridische Zaken of Arbeidsvoorwaarden en Beroepskwaliteit. Verder zijn er directies die andere directies ondersteunen bij de beleidsontwikkeling, zoals de directies

Personeel en Organisatie, Voorlichting en het Facilitair Bedrijf. De uitvoering van het beleid wordt georganiseerd door relatief zelfstandige organisaties, zoals de twee agentschappen, maar ook door procesmanagement en zelfstandige bestuursorganen (ZBO's).

Organisatieschema



De medewerkers

Bij het bestuursdepartement werken ongeveer veertienhonderdvijftig mensen. Daarnaast werken bij de agentschappen achthonderd mensen, bij de rijksdiensten vijfhonderd en bij de inspecties vijfhonderd.

Bij OCW werken bijna evenveel vrouwen als mannen. Vergeleken bij andere Nederlandse organisaties werken bij OCW veel academici en hbo'ers, vanuit diverse achtergronden: rechten, sociologie, politicologie, (kunst)geschiedenis, economie, informatiekunde, wiskunde, biologie et cetera. In hun loopbaan zijn zij sterk gericht op verder leren en trainen. Interne mobiliteit is bij OCW actueel. De laatste jaren zijn veel jonge mensen bij OCW komen werken en is OCW gast voor veel trainees en stagiaires.

Het budget

OCW werkt met een budget van circa 45 miljard gulden. Het geld wordt geïnvesteerd in:

- ⌘ primair onderwijs en voortgezet onderwijs;
- ⌘ beroepsonderwijs en volwasseneneducatie;
- ⌘ hoger beroepsonderwijs en wetenschappelijk onderwijs;
- ⌘ onderzoek en wetenschapsbeleid;
- ⌘ studiefinancieringsbeleid;
- ⌘ kunst, cultuur en media.

De helft van het totale budget is bestemd voor de salarissen van leraren, universitaire docenten, onderzoekers en andere personeelsleden.

2.2 Knelpuntenanalyse (oriëntatiefase)

Voor consistentie in de organisatie van informatie is het belangrijk om aan te sluiten bij reeds ontwikkelde standaarden voor de ontsluiting van digitale informatiebronnen. Data wordt hierbij in de meest uitgebreide zin gehanteerd: tekst, afbeelding, muziek, grafisch materiaal of bewegende beelden: het zijn data. Evenzeer als voor niet-tekstdata is het, door het groeiend volume, ook belangrijk metadata aan tekstdata toe te voegen. Daarnaast kunnen metadata ook een belangrijke rol spelen voor het administratief beheer van de informatiebronnen (versiebeheer, houder).

Keuze voor gestructureerde metadata

De rol die metadata spelen bij het vinden van informatie met zoekmachines is nog niet helder. Er is inmiddels wel een aantal zoekmachines (waaronder de SURFnet Search Engine) die het mogelijk maken om specifiek te zoeken met metadata. Bij de keuze voor de Dublin Core Metadata set gaat men uit van de verwachting dat zoekmachines op termijn deze standaard zullen ondersteunen. De bestaande zoekmachine van het ministerie, de Verity Information Server, kan de Dublin Core Metadata aan.

Met de directie Voorlichting wordt een nieuwe procedure rond webpublicaties besproken. Daarbij staat ook de bestaande trefwoordenlijst ter discussie. Moet er gebruik gemaakt worden van een geautoriseerde trefwoordenlijst?

Er is een thesaurus aanwezig, maar het onderhoud daarvan is een zwak punt. Bovendien wordt door gebruik te maken van trefwoorden voorbij gegaan aan de toegankelijkheid voor de burger: die zal niet de thesaurusterm 'schoolverzuim' gebruiken maar bijvoorbeeld 'spijbelen'. De eindgebruiker heeft behoefte aan eenvoudige zoekfunctionaliteit, de meer professionele gebruiker zal vaker gebruik willen maken van geavanceerde zoekfaciliteiten.

2.3 Toename digitale documentaire informatie

2.3.1 Intern

Het ministerie is een informatie- en kennisintensieve organisatie. Papier is (nog) de voornaamste drager van informatie en kennis, maar in toenemende mate wordt gewerkt aan de vervanging van papier door digitale media.

OCW heeft ervoor gekozen om iedereen zijn eigen digitale teksten te laten opslaan en beheren via gemeenschappelijke schijven, de G:\-schijven. De toegankelijkheid tot die informatie vereist een specifieke aanpak. Het eindproduct is nog altijd een papieren document. De verwachting is wel dat in toenemende mate digitale (eind)producten tot stand komen. Ook deze digitale documenten zijn van belang bij het afleggen van verantwoording en reconstructie van het gevoerde beleid. Deze ontwikkeling vraagt om een andere manier van omgaan, toegankelijk maken en opslaan van documenten.

2.3.2 Extern

Behalve een toename in interne digitale documenten neemt ook de beschikbaarheid van externe elektronische documenten, zoals tijdschriften, rapporten en beleidsnota's, toe. Het beheer van de inhoud (content) wordt daarbij steeds belangrijker. Er is nog geen allesomvattende standaardisatie voor elektronisch archiveren en beschikbaar houden van elektronische publicaties. Ook de elektronische collectievorming op het intranet van OCW staat nog in de kinderschoenen. Evident is de noodzaak tot borging van toegankelijkheid van deze documenten.

2.4 Tendensen in het overheidsbeleid

2.4.1 Transparante overheid

Beschikbaarstelling en toegankelijkheid zijn de kernbegrippen voor de informatievoorziening vanuit de documentaire invalshoek. In 2001 verschenen twee rapporten die elk de noodzaak ervan onderschrijven en waarin de overheid, de burger en informatie in relatie met elkaar worden gebracht. Deze relatie is niet nieuw maar is door de ontwikkelingen in de informatie- en communicatie-technologie (ICT) dringend aan herziening toe.

Beide commissies keken vanuit verschillende gezichtspunten naar informatievoorziening vanuit de overheid aan de burgers. De conclusies en aanbevelingen echter vertonen duidelijk raakvlakken en overlappingsen.

Op 7 augustus 2001 verscheen het rapport 'In dienst van de democratie' van de Commissie Toekomst Overheidscommunicatie onder voorzitterschap van J. Wallage waarin gesteld wordt dat technologische ontwikkelingen het mogelijk maken de communicatie tussen overheden en burgers

te verbeteren met als doel het draagvlak voor overheidshandelen te verbreden. Burgers hebben recht op communicatie en maximale transparantie. Een 'actieve publicatieplicht' houdt in, dat overheidsdocumenten op internet moeten worden gepubliceerd, ook de plannen waarover nog geen besluit is genomen.

Op 6 september 2001 verscheen het rapport 'Burger en overheid in de informatiesamenleving'; de noodzaak van institutionele innovatie' van de Eenmalige Adviescommissie ICT en Overheid onder voorzitterschap van A. Docters van Leeuwen. In dit rapport worden een aantal aanbevelingen gedaan die er toe zullen moeten leiden dat de rol van de overheid in de informatiesamenleving wordt versterkt. Die 'nieuwe' overheid moet de toegang en de toegankelijkheid tot de elektronische overheid en / of haar diensten garanderen.

Gegeven de vraag wat de consequenties zijn van ICT op de rol van ons ministerie in de informatie-verstrekking aan de burger kunnen de aanbevelingen van beide rapporten niet los van elkaar gezien worden.

2.4.2 *Kwaliteitsslag OCW*

Een belangrijk element van de Kwaliteitsslag OCW is dat medewerkers hun uit te voeren taken beter kunnen sturen en beheersen en dat zij de daarbij gebruikte gegevens en documenten gestructureerd kunnen opslaan en beter kunnen terug vinden.

Medewerkers moeten plaats- en tijdonafhankelijk aan hun eigen documenten kunnen werken, maar ook moeten meerdere personen aan één document kunnen werken. Medewerkers dienen snel de juiste versie van een document te kunnen inzien. Managers willen inzicht in de afhandeling van documenten en de werkvoorraad van hun medewerkers. Medewerkers moeten zich kunnen verantwoorden, evenals de organisatie. Natuurlijk worden de beveiliging en toegang goed geregeld en kunnen de directeuren en beleidsmedewerkers daar zelf keuzes in maken.

Door het integreren van document- en recordmanagement in de werkprocessen wordt de adequate digitale documenthuishouding vorm gegeven en wordt het werk van de OCW- medewerkers, het werken met documenten, optimaal ondersteund. Deze integratie vindt plaats in een zogenaamd Document Management Systeem en Records Management Applicatie (DMS / RMA). Dit heeft verregaande consequenties voor de werkwijze van alle medewerkers van OCW.

Het behelst dus veel meer dan het implementeren van een applicatie. Het vergt organisatorische, instrumentele en procedurele aanpassingen waaronder:

- ⋮ de invoering van een eenduidige taal (metadata) o.a. voor de ontsluiting en benaming van documenten;
- ⋮ het inrichten van een beheerorganisatie;

- ⋮ het analyseren, definiëren en vastleggen van de verantwoordelijkheden organisatiebreed;
- ⋮ het opleiden van de medewerkers.

Ook zal er intensief gecommuniceerd moeten worden met de eindgebruikers en zullen ook deze medewerkers bij de implementatie begeleid moeten worden

2.4.3 *Taakstelling OCW*

Het kabinet Balkenende wil een slagvaardiger, slankere overheid. Met ministeries die minder regelen en meer aan het veld overlaten. Het ministerie van OCW heeft als opdracht de taakstelling uit het Strategisch Akkoord uit te voeren. Dit moet voor het departement een bezuiniging opleveren van 11%. Inmiddels is een nieuwe kabinet aangetreden die uitgaat van een hoger percentage (14%)

De Bestuursraad en bewindslieden hebben besloten de taakstelling op een kwalitatieve, beleidsrijke wijze in te vullen. Dus niet via de zogenoemde 'kaasschaafmethode' (overal een beetje van af halen), maar doorgaan op de ingeslagen weg van de Kwaliteitsslag. Zo wordt er verder gewerkt aan de strategische herpositionering van het departement, aan de kwaliteits- en efficiencyverbetering van de organisatie en aan de verdere professionalisering van de medewerkers. Met andere woorden: samen op weg naar een departement dat de goede dingen doet op de goede manier en met de juiste mensen.

2.5 **Initiatieven en ervaringen tot nu toe**

2.5.1 *Standaarden en richtlijnen*

Internationaal zijn er verschillende initiatieven en werkgroepen bezig om standaarden en richtlijnen te ontwikkelen om metadata aan bronnen toe te kennen. Eén van de standaards is de Dublin Core Metadata. Deze basisset metadata elementen is ontwikkeld door experts uit verschillende landen, van IT-specialisten tot bibliothecarissen. Op basis van internationale en interdisciplinaire consensus zijn vijftien elementen vastgesteld, waarmee de inhoud van een document en een breed scala aan internetbronnen beschreven kan worden.

2.5.2 *Onderzoek naar metadata (desk research)*

In de interne notitie 'Ontsluiting van Internet, intranetten en Metadata'² zijn de bevindingen vastgelegd naar de stand van zaken rond metadata.

²

Seitzinger, Ruud: *Ontsluiting van Internet en intranet met metadata*. OCW: 13 april 2003

De belangrijkste conclusies in deze notitie zijn:

- ⋮ dat er veel technieken voor het terugvinden van papieren documenten zijn ontwikkeld zoals catalogisering, rubricering, indexering en inhoudelijke ontsluiting met thesaurustermen. Deze technieken kennen hun toepassing ook in de elektronische wereld, zoals op indexen gebaseerde of thematisch gerangschikte zoekmachines en gespecialiseerde wegwijzers (portals).
- ⋮ dat de ervaringen met zoekmachines duiden op beperkte zoekfunctionaliteiten. Vangst (recall) en precisie (precision) laten te wensen over en (andere) methoden zouden moeten worden ingezet om deze te verbeteren;
- ⋮ dat door het toepassen van een metadastandaard, als de Dublin Core Metadata Set, elektronische documenten en publicaties preciezer terug te vinden moeten zijn en dat aansluiting bij deze standaarden gewenst is.

2.5.3 Intern onderzoek

Binnen het bestaande publicatieproces van internet en intranet is gekeken naar het moment waarop metadata kunnen worden toegevoegd. Bij de publicatie van elektronische documenten worden metadata in het document opgenomen. Dat maakt dat het moment van ontsluiten niet - zoals voorheen bij papieren publicaties - achteraf plaats vindt maar tijdens of aan het begin van het publicatieproces.

Uit dit onderzoek volgt dat op het OCW-plein (internet) voor de ontsluiting van persberichten reeds op bescheiden schaal metadata (keywords) worden toegekend. Er wordt daarbij gebruik gemaakt van een niet geformaliseerde of gestandaardiseerde trefwoordenlijst.

Het publicatieproces op OCW-Net (Intranet) blijkt diffuus: er is sprake van een groot aantal lokale redacteuren van de verschillende directies. Er wordt niet gewerkt met gestandaardiseerde metadata.

2.5.4 Extern onderzoek

De Stichting 'Itafit' is een onderzoeksnetwerk dat als doelstelling heeft de verspreiding van kennis over bestuurlijke vernieuwing en ICT te bevorderen. Onder meer door het verrichten van casestudies. Eén van de casestudies die zij hebben uitgevoerd is het vooronderzoek 'De kunst van het vinden'³ met als doelstelling: het vinden van voorbeelden en ervaringen op het terrein van vindbaarheid.

³ 'De kunst van het vinden' Vooronderzoek naar mogelijkheden tot het verbeteren van de vindbaarheid van (rijks)overheidsinformatie: Stichting Itafit, April 2002

Uit de onderzochte cases blijkt dat vindbaarheid niet op zichzelf staat maar het hart raakt van het overheidsapparaat. Het vergroten van de vindbaarheid maakt deel uit van de aanleg van een landelijke overheidsinformatie-infrastructuur. Een overheidsorganisatie die geïsoleerd via een website haar beleidsinformatie als product aanbiedt, zal niet slagen. De website moet de verbinding vormen tussen verschillende partijen in de keten (zoals het Parlement, adviesorganen, et cetera). Samenwerking in de vorm van afstemming tussen overheidsorganisaties is dus noodzakelijk.

De meest belangrijke aanbevelingen zijn:

Standaarden als uitgangspunt

Het is wenselijk om minimaal op nationaal niveau afspraken te maken over hoe om te gaan met metadata. Het is beter om aansluiting te zoeken bij internationale afspraken zoals Dublin Core Metadata Initiative en deze eventueel om te zetten naar de eigen taal.

Vraaggerichte centrale overheidssite

Overheidswebsites bieden algemene ingang tot de overheid. Ze dienen min of meer als toegangspoort, waardoor online groepen gebruikers toegang hebben tot informatie en diensten. Een goede site heeft de logica van de klant als uitgangspunt. Vindbaarheid wordt pas echt een issue indien wordt gedacht vanuit de klant.

Intelligente zoekmechanismen

Het internet maakt het lastig om op eenvoudige wijze ordening in informatie aan te brengen omdat het bij uitstek een warrig netwerk is. Vindbaarheid is gebaseerd op de idee dat op een intelligente wijze informatie wordt ontsloten waarbij men leert van het zoekgedrag van de gebruiker. Amazon.com wordt hier als voorbeeld aangegeven.

Leren van anderen

Met leren van anderen wordt bedoeld dat overheidsorganisaties meer oog moeten hebben voor al bestaande oplossingen en niet telkens weer het wiel moeten willen uitvinden. Men moet veel meer gebruik maken van bestaande tools, toepassingen en standaarden. Daarnaast kunnen organisaties ook veel leren van de gebruiker door zijn zoekgedrag te analyseren.

3 Opdrachtformulering

3.1 Achtergrond project

In 2001 heeft de PSG van OCW het startsein gegeven voor het digitaliseringstraject. Het uiteindelijke doel van het digitaliseringstraject is dat alle informatie via één toegang tijds- en plaats-onafhankelijk leesbaar is. Om dit te bereiken is het traject opgedeeld in drie fasen: de vervangingsfase, de integratiefase en de transparatiefase⁴.

De vervangingsfase is inmiddels afgerond en blijkt de tussentijdse evaluatie van Het Expertise Centrum⁵ (HEC) is deze fase zeer voorspoedig verlopen. Het HEC heeft geadviseerd snel van start te gaan met de ontwikkeling en uitvoering van de volgende fase.

Beleidsplan Digitaal Werken

De informatiemanager van OCW heeft voor de integratiefase opdracht gegeven tot het realiseren van een beleidsplan en een werkprogramma. Dit beleidsplan Digitaal Werken (de nadere uitwerking van het beleidsplan van november 2000) is op 30 september 2002 door de directie FacB vastgesteld⁶. In dit beleidsplan is de doelstelling, dat moet leiden tot het volledig digitaal werken binnen het kerndepartement, nader uitgewerkt.

Aanbevelingen uit het beleidsplan:

- ⋮ vaststellen van de te ondersteunen (beleids)dossiers en werkprocessen en de daaruit voortvloeiende gegevens- en procesmodellen (metadatabeleid, ordeningsplannen, workflow, automatisaties);
- ⋮ invoeren van een zoekfunctionaliteit en een applicatie voor documentmanagement en recordmanagement. In dit instrumentarium worden alle interne documenten opgenomen, toegankelijk gemaakt, gerouteerd, opgeslagen en gearchiveerd. Documenten kunnen beveiligd worden, er vindt versiebeheer plaats en documenten kunnen worden hergebruikt. Met goede zoekfunctionaliteit kunnen alle opgeslagen documenten snel gevonden worden;
- ⋮ aanpassing van de administratieve organisatie op het gebied van de documenthuishouding;

⁴ Deze fasering is gebaseerd op het beleidsplan document digitalisering (FacB, nov 2000)

⁵ Projectevaluatie departementsbrede implementatie digitalisering documentopslag (HEC, april 2002)

⁶ Op 30-9-2002 is het beleidsplan over het management van de digitale documenten van het ministerie van onderwijs, cultuur en wetenschappen, 2002 – 2005 dd. 30 september 2002, versie 0.1, besluitnr. 2002/12 vastgesteld.

- ⋮ het bieden van ondersteuning aan alle OCW-medewerkers bij hun verantwoordelijkheid ten aanzien van het verantwoord omgaan met (digitale) documenten middels gedegen cursussen en instructies;
- ⋮ het begeleiden van de betrokken professionals in hun omslag van archiveren van afgedane documenten naar:
 - ⋮ het toegankelijk maken van documenten en;
 - ⋮ hun adviesrol naar medewerkers toe.

3.1.1 *Project doelstellingen*

Om uitvoering te geven aan het beleidsplan is het werkprogramma 'Digitaal Werken bij OCW' van start gegaan. Het doel van het werkprogramma is dat alle OCW-medewerkers in staat worden gesteld effectief en efficiënt de juiste versie van het goede document te vinden en te ontvangen. Welke directie een bepaald document heeft gemaakt of welke directie een document in behandeling heeft, mag hier niet direct een rol in spelen. De eerste fase in het programma is het project 'Zoeken en vinden' waar dit onderzoek 'Auteurs-gegenereerde metadata' onderdeel van uit maakt.

3.1.2 *Omvang van het project*

Het is noodzakelijk om, gelet op de groeiende omvang van digitale eindproducten, de duurzame bewaring, ontsluiting en beschikbaarstelling van de informatiebronnen op de gemeenschappelijke dataservers (de G:\-schijven) beter te beheersen en te beheren. Hiervoor moet iedere medewerker van OCW kunnen beschikken over adequate voorzieningen (inhoudelijk, technisch, organisatorisch en procedureel en vooral ook gebruikersvriendelijk) om:

- ⋮ relevante metadata-elementen vast te leggen;
 - ⋮ op basis van metadata sneller en preciezer relevante documenten c.q. informatie te vinden.
- De noodzakelijke basis (gewenning, toegankelijkheid van documentaire informatie) voor de verdere invoering van digitaal documentbeheer wordt hier gelegd.

Om de benodigde kennis hiervoor te vergaren worden bij twee directies voor maximaal 30 gebruikers een pilotomgeving ingericht. In deze omgeving worden de Storage & Retrieval instrumenten ontwikkeld en getest. Daarnaast wordt in een interviewronde de status van de toegankelijkheid van de documentaire informatievoorziening getoetst. De DMS (DocMan) valt buiten de scope van dit onderzoek.

3.1.3 *Projectresultaten*

- ⌘ Het doen van aanbevelingen voor:
 - ⌘ de vaststelling en invoering van een eenduidige 'standaardtaal' o.a. voor de inhoudelijk ontsluiting van digitale (eind)documenten op de G:\-schijf;
 - ⌘ de benaming van digitale (eind)documenten op de G:\-schijf.
- ⌘ Het realiseren van - op de gebruiker afgestemde - Storage & Retrieval-instrumenten voor:
 - ⌘ de toekenning / vastlegging van metadata;
 - ⌘ het zoeken documenten.
- ⌘ Het doen van aanbevelingen voor c.q. realisatie van een departementaal implementatieplan (inclusief gebruikersinstructie c.q. opleiding).
- ⌘ Realisatie van een adequate beheersorganisatie.
- ⌘ Literatuuronderzoek 'historie en functie metadata': centraal toekennen (achteraf) versus vastlegging bij de bron (auteurs-gegenereerd).

3.1.4 *Randvoorwaarden & beperkingen*

- ⌘ Het projectresultaat moet voldoen aan de eisen die gesteld worden in de Archiefwet, de Wet Openbaarheid van Bestuur, de Wet Bescherming Persoonsgegevens en het Voorschrift Informatiebeveiliging.
- ⌘ De pilots worden zoveel mogelijk met behulp van beschikbare instrumentaria uitgevoerd (sjablonentechnologie, ontsluitingstaal, Verity-zoekmachine, macro's in Office) en binnen de standaards die voor OCW zijn gedefinieerd.
- ⌘ In het licht van de verhuizing naar de Hoftoren (oktober) en de bevrozing van de netwerk-faciliteiten moeten de investeringen in de huidige ICT-infrastructuur beperkt blijven tot het hoogst noodzakelijke.
- ⌘ Voldoende betrokkenheid bij de gebruikersorganisatie (de pilotdirecties).
- ⌘ Voldoende financiële middelen.
- ⌘ Beschikbaarheid van een ICT-beheersomgeving voor de Sjabloontechnologie en Verity Information Server.
- ⌘ Professioneel projectmanagement volgens de Prince-2 methodiek.

3.1.5 *Business Case*

In het Beleidsplan komt naar voren dat er invulling gegeven dient te worden aan transparantie van informatie en gegevens, zowel intern als extern. Hiertoe dienen documenten professioneel, verantwoord en overzichtelijk te worden opgeslagen zodat het 'organisatiegeheugen' voor meerdere doeleinden adequaat is ontsloten.

Het doel van het programma 'Digitaal Werken bij OCW' komt hieruit voort. Een belangrijk element van dit digitaliseringstraject is dat medewerkers hun taken beter kunnen uitvoeren en dat de informatievoorziening hen daar efficiënt in ondersteunt. Zij moeten daartoe documentaire informatie op een goede en toegankelijke wijze kunnen opslaan en vooral ook terug kunnen vinden.

Als uiteindelijk resultaat betekent dit, dat alle informatie door middel van één 'user-interface' plaats- en tijdonafhankelijk benaderbaar moet zijn – ongeacht de plek van opslag – via op elkaar afgestemde en marktconforme 'Storage & Retrieval' technieken. Om dat te kunnen bereiken is een eenduidige ontsluiting van de documentaire informatie noodzakelijk

Het werkprogramma 'Digitaal werken bij OCW' is een meerjarenprogramma waarin de benodigde instrumenten gefaseerd worden geïmplementeerd. De resultaten van dit project vormt de noodzakelijke basis (gewinning, vergroting toegankelijkheid van documenten) voor de verdere invoering van digitaal documentbeheer binnen OCW.

3.1.6 *Kwaliteitsverwachtingen*

- ⋮ Door de gebruikers (pilotdirecties) geaccepteerde userinterface voor zowel de bewerkingsfase (vastleggen van metadata bij de bron) als in de zoekfase.
- ⋮ Een door de gebruikers (pilotdirecties) geaccepteerde standaard ontsluitingstaal.
- ⋮ Medewerkers van de betrokken pilotdirecties kunnen relevante informatie op hun G:\-schijf sneller en preciezer terugvinden.
- ⋮ Organisatiebrede (per directie) toegankelijkheid van de documentaire informatie op de G:\-schijf is beter georganiseerd.
- ⋮ De vastgelegde metadata-elementen zijn (her)bruikbaar voor het publicatieproces op het OCW-Intranet c.q. het OCW-plein (Internet).
- ⋮ De beheerorganisatie van het instrumentarium is in lijn met beheerorganisaties van overige applicaties en is eenduidig en helder met een juiste balans in centraal en decentraal. Hiertoe zal intensieve afstemming plaatsvinden met de betrokkenen in de bestaande beheerorganisatie en de pilotdirecties.

4 Diepteonderzoek

Het uiteindelijke resultaat van het werkprogramma 'Digitaal werken bij OCW 2003 - 2006' is, dat alle OCW-medewerkers relevante informatie en informatiebronnen met één 'user-interface' plaats- en tijd-onafhankelijk kunnen benaderen. Op elkaar afgestemde en marktconforme 'Storage & Retrieval' technieken zijn hierbij het uitgangspunt.

Het is daarom noodzakelijk om, gelet op de groeiende omvang van digitale (eind)producten op de gemeenschappelijke dataserver (de G:\-schijf), de duurzame bewaring, ontsluiting en beschikbaar- stelling van deze informatiebronnen beter te beheersen en te beheren. Immers deze digitale bronnen vormen in een zeer belangrijke mate de neerslag van het handelen van de OCW- organisatie. Om dit bereiken is een eenduidige ontsluiting van de documentaire informatie bij de bron nood-zakelijk. Iedere OCW-medewerker moet hiervoor beschikken over adequate voor- zieningen (technisch, organisatorisch en procedureel) om:

- ⋮ op een gebruikersvriendelijke wijze metadata vast te kunnen leggen;
- ⋮ sneller en preciezer relevante documenten c.q. informatie te vinden.

Om dit te realiseren moeten we antwoorden vinden op de volgende vraag:

'Wat kan c.q. moet de organisatie doen om, gebruikmakend van bestaande Storage & Retrievaltechnieken, de documentaire informatievoorziening te optimaliseren c.q. de (duurzame) toegankelijkheid van de documentaire informatie op de gemeenschappelijke dataserver (de G:\-schijf) te vergroten en te borgen?'

4.1 Onderzoeksdoelstelling

Het verkrijgen van actuele kennis en inzichten over bestaande Storage & Retrievaltechnieken om passende maatregelen te kunnen nemen om:

- ⋮ de organisatiebrede (duurzame) toegankelijkheid van digitale (eind)documenten op de G:\-schijf beter te borgen;
- ⋮ de veranderingsgezindheid van de medewerkers rond een nieuwe werkwijze positief te beïnvloeden.

- ⋮ Daarnaast moeten de resultaten van het onderzoek:
 - ⋮ een basis vormen voor een adequaat in te richten beheersorganisatie;
 - ⋮ bijdragen/ input leveren aan de gerelateerde projecten uit het werkprogramma 'Digitaal werken bij OCW 2003-2006':
 - de ontwikkeling en implementatie van de Verity-zoekmachine voor G:\-schijven;
 - de functionele analyse DMS / RMA.

4.2. Onderzoeksvraagstelling

Om de doelstellingen van het onderzoek, zoals genoemd in § 4.1., te realiseren zijn de volgende centrale vragen gedefinieerd.

- 1 Onder welke condities is de gebruiker bereid om gestructureerde metadata toe te kennen en vast te leggen bij de creatie van Officedocumenten?
- 2 Waarom is het belangrijk te kiezen voor ontsluiting van documentaire informatie bij de bron in plaats van dit centraal te organiseren?
- 3 Leidt het zoeken op basis van gestructureerde en kwalitatief goed toegekende metadata tot snellere en betere zoekresultaten dan bij het zoeken op vrije tekst op de gemeenschappelijke netwerkschijf (G:\-schijf)?

De meest essentiële vraag waar het onderzoek een antwoord op moet geven is de eerstgenoemde vraag.

Wanneer de condities goed zijn en voldoende aansluiten op de belevingswereld van de gebruikers zal de acceptatiedrempel laag zijn en daardoor het verandervermogen hoog.

Veranderen is het doorbreken van een status quo. Van een situatie waarin gebruikers weten hoe het gaat, wat er van ze verwacht wordt en waarin ze zich al dan niet kundig voelen naar een situatie waarvan ze niet weten hoe het gaat, wat er precies van ze verwacht wordt en waarin ze nog niet bekwaam zijn.

Om antwoord te kunnen geven de centrale vragen en deelvragen zijn de volgende activiteiten uitgevoerd:

- **Voorbereiding**
In overleg met de opdrachtgever is het onderzoekgebied en de onderzoeksopdracht nader afgebakend in de projectbrief 'Auteurs-gegenereerde metadata', 16 december 2002.
- **Literatuuronderzoek**
Door het lezen, analyseren en verifiëren van beschikbare documentatie is knelpuntenanalyse uitgevoerd en een globaal beeld gevormd van de huidige situatie en relevante ontwikkelingen op het gebied van metadata.

- **Interviews**
Door middel van een interviewronde met OCW-medewerkers uit verschillende disciplines is vanuit het gebruikersperspectief de mening gepeild rond ontsluiting en toegankelijkheid van informatie in brede zin (Internet, Intranet en G:\-schijf).
- **Pilots**
In twee pilotomgevingen is de technische realisatie van invoer- en zoekfunctionaliteiten in samenwerking met de gebruikers ontwikkelt en getest. Daarvoor is gebruik gemaakt van de nieuwe sjabloontechnologie Smart Documents en de Verity-zoekmachine.
- **Kwantitatieve analyse**
Door de volumetoename op de gemeenschappelijke dataservers te monitoren en te analyseren is de urgentie van de problematiek inzichtelijk geworden.

4.3 Inrichting pilots bij de directies

Om de toegankelijkheid van de snel groeiende hoeveelheid documentaire informatie op de G:\-schijven te vergroten is het gebruik van een eenduidige ontsluitings-/zoektaal onontbeerlijk. Gestructureerde en gestandaardiseerde metadata zijn hier een goed hulpmiddel bij.

Dit vergt evenwel organisatorische, instrumentele en procedurele aanpassingen zoals:

- ⋮ invoering een eenduidige taal (gestandaardiseerde metadata) voor het ontsluiten en terug kunnen zoeken van documenten;
- ⋮ afspraken over de benaming van documenten;
- ⋮ aanpassing / ontwikkeling van een technisch instrumentarium (formulier om metadata vast te leggen en te zoeken op basis van metadata);
- ⋮ een implementatieplan;
- ⋮ inrichting van een beheersorganisatie;
- ⋮ instructies / opleiding voor de OCW-medewerkers.

Naast deze meer instrumentele aanpassingen is het van groot belang te weten hoe dit wordt ervaren bij de gebruikers. Om dit te bewerkstelligen is bij twee directies een pilotomgeving ingericht met als doel:

- ⋮ het in samenwerking met een beperkt aantal gebruikers ontwikkelen en testen van metadata opslag- en zoekfunctionaliteiten;
 - ⋮ de benodigde kennis en inzichten te vergaren voor een mogelijke implementatietraject.
- Bij de inrichting van de pilotomgeving is bij de gebruikers de nieuwe sjabloontechnologie Smart Documents geïntroduceerd voor de vastlegging van een beperkte set metadata (titel, onderwerp, auteur, trefwoord en optioneel een samenvatting). Daarnaast is er voor de pilot een zoek- en

resultatenschermb gemaakt om het zoeken via metadata op het Intranet en de G:\-schijf mogelijk te maken.

4.3.1 Evaluatie van de pilot

Algemeen

- ⌘ Het traject 'Auteurs-gegenereerde metadata' wordt als een positieve volgende stap in het digitaliseringstraject gezien. De beperkte toegankelijkheid, van de informatie opgeslagen op de gemeenschappelijke dataservers (G:\-schijven) en de Intranetsserver, is al langer een onderwerp van zorg.
- ⌘ De input van de pilotdeelnemers, bij de ontwikkeling en testen van het nieuwe instrumentarium, is zeer beperkt geweest. Slechts een enkeling heeft zich de moeite getroost om er daadwerkelijk mee aan de slag te gaan. Het gelogde zoekgedrag van de pilotmedewerkers kende bijvoorbeeld een hoogtepunt vlak voor de evaluatiebijeenkomst. In de gehele pilotperiode daarvoor was volgens de lograpporten geen gebruik gemaakt van de zoekfunctionaliteiten.
- ⌘ Onderschatting van de tijdsbesteding, die de deelname aan een pilot met zich meebrengt, is als een van de oorzaken genoemd. Zowel aan de kant van de pilotdeelnemers als aan de projectkant moet met voldoende tijd aan het project kunnen worden gewerkt.
- ⌘ Gebruikersvriendelijkheid geldt uitdrukkelijk voor alle facetten van de nieuwe functionaliteit als één van de belangrijkste acceptatiecriteria.

Informatie (inhoud)

- ⌘ Bij een klein volume worden de documenten sneller gevonden met de Verkenner of via Word Office (bestand openen). Dit is mede bepalend voor de bereidheid om metadata toe te voegen voor het gebruik van zoekinstrumenten. Bij een gering volume zullen de gebruikers sneller geneigd zijn de standaard hulpmiddelen te gebruiken (Verkenner e.d.).
- ⌘ Een eenduidige en heldere definitie van de metadata-elementen wordt belangrijk gevonden. Hoewel onderwerp en titel vaak hetzelfde zullen zijn blijft de behoefte voor het kunnen vastleggen van zowel een titel als een onderwerp aanwezig. Bijvoorbeeld: de notitie met de titel *'Met de rug tegen de muur'* met als onderwerp *'Internationalisering van het hoger onderwijs'*.
- ⌘ Bij de zoekresultaten wordt het belangrijk gevonden dat:
 - de ingegeven zoekvraag bij de zoekresultaten wordt gepresenteerd (eis);
 - de zoekwoorden in de samenvatting gemarkeerd worden (wens).
- ⌘ De opslag en hergebruik van zoekopdrachten, de zogenaamde zoekprofielen is ook een wens.

- ⌘ Rubrieken en trefwoorden
 - De OCW-rubriekenlijst is onvoldoende helder. In de toelichtingen moet beter worden beschreven wat de scope van een rubriek is. De rubrieken zijn ook niet consistent: (zie rubrieken *Beleid en Arbeidsvoorwaardenbeleid*).
 - De verwijzingsystematiek van de trefwoordenlijst is niet consistent. Daar waar bij het ene begrip verwezen wordt van de afkorting naar het volledige begrip wordt bij een andere trefwoord verwezen van het volledige trefwoord naar de afkorting (zie onderstaand voorbeeld).

ADMINISTRATIEVE ORGANISATIE

zie: AO

AO

x ADMINISTRATIEVE ORGANISATIE

PUBLIC RELATION

x PR

PR

zie: PUBLIC RELATION

- Zowel de trefwoordenlijst als de rubriekenlijst hebben een dynamisch karakter. Naast de behoefte om als gebruiker zelf vrije trefwoorden toe te kennen moet het mogelijk zijn om optioneel de rubriekenlijst te wijzigen of uit te breiden met nieuwe rubrieken. De beheerorganisatie 'Trefwoorden en rubrieken' moet hier in voorzien.

Organisatie

- ⌘ Bij een innoverend project zoals 'Auteurs-gegenereerde metadata' wordt communicatie als één van de belangrijkste succesfactoren gezien. Hier wordt o.a. onder verstaan:
 - een goede voorbereiding door middel van instructiebijeenkomsten;
 - heldere en korte handleidingen/instructies met inhoudsopgave, goede voorbeelden, alfabetisch register, et cetera;
 - een goed bereikbare helpdesk.
Een laagdrempelige 1e lijns (DIV en lokale redacteuren) en goed bereikbare 2e lijns ondersteuning (centrale helpdesk).
- Geconstateerd wordt dat de bereidwilligheid van medewerkers afneemt als dit onvoldoende is georganiseerd. De bereikbaarheid van een helpdesk speelt hierbij een belangrijke rol.

- ⋮ Het gefaseerd invoeren van nieuwe functionaliteiten (hapklare brokken) op de werkplek wordt ook erg belangrijk gevonden. Medewerkers hebben tijd nodig om aan de nieuwe instrumenten en werkwijze te wennen. Het moet snel zichtbare resultaten opleveren.
Als de nieuwe functionaliteiten niet leveren wat verwacht wordt, ontstaat er een drempel om er verder mee aan de slag te gaan en vervalt men weer in oude werkwijzen/gewoontes.
- ⋮ Een dwingende vorm van registratie van metadata wordt daarom noodzakelijk geacht. Als mogelijkheid wordt bijvoorbeeld aangegeven dat je een document niet kan opslaan als er minimaal niet een trefwoord of rubriek is ingevuld. Ook de andere metadata-elementen zouden ingevuld moeten zijn om het document systematisch te kunnen opslaan: titel (combinatie van vorm en onderwerp), onderwerpsomschrijving en auteursgegevens. Dit stelt eisen aan de wijze waarop de nieuwe sjabloonttechnologie Smart Document is of moet worden ingericht.

Personeel

Voor wat betreft de participatie van de pilotmedewerkers is een keuze gemaakt voor medewerkers die dagelijks geconfronteerd worden met het zoeken van informatie. Gedurende de pilotfase is geconstateerd dat de bereidwilligheid afneemt als de communicatie (zoals in de voorgaande paragraaf al is aangegeven) en directe ondersteuning onvoldoende is. Daarnaast ervaart men een toename van de bereidheid tot het doen van inspanningen vooraf als die inspanningen de juiste resultaten opleveren.

Techniek

- ⋮ De koppeling met de nieuwe sjabloonttechnologie (Smart Document) is door de pilotdeelnemers positief ervaren. De deels geautomatiseerde vastlegging van metadata zoals onderwerp, auteur en trefwoorden (keuzelijsten) wordt hiermee voldoende ondersteund. Het invoerscherm metadata is voldoende helder en gebruikersvriendelijk.
Het zoveel mogelijk geautomatiseerd vastleggen van metadata (minimale inspanning bij de auteur) is een eis.
- ⋮ Zoeken is een kunst: intuïtief zoeken is mede afhankelijk van mogelijkheden van een zoekmachine. Een goede (zoek)instructie c.q. handleiding is hierbij zeer gewenst.
- ⋮ De verfijnfunctie van het huidige zoekformulier werkt volgens betrokkenen niet naar behoren. Daarnaast is niet helder tot op welke (sub)directoryniveau de zoekmachine indexeert.
Beide opmerkingen dienen nader onderzocht te worden.

- ⋮ Er is aparte zoekfunctionaliteit beschikbaar voor het Documentair Informatiesysteem (de Document Manager Browser), het Intranet en Smart Documents. Ook binnen de Officesuite en de Windows omgeving worden separate zoekmogelijkheden aangeboden. De gebruiker raakt hier het spoor bijster.
- ⋮ Vastgesteld moet worden welke zoekfunctionaliteit die nu naast elkaar worden aangeboden, geoptimaliseerd en geïntegreerd kunnen worden om in één zoekopdracht vollediger en betere zoekresultaten te verkrijgen. Uitgangspunt moet zijn dat de aangeboden zoekfunctionaliteit eenvoudig werkt en betere resultaten oplevert dan bijvoorbeeld de standaard zoekmogelijkheden van de Officesuite en de Windows Verkenner. Termen als snelheid en nauwkeurigheid van de zoekresultaten zijn basiseisen.
- ⋮ Ook wordt verwacht dat de zoekmachine intelligent om gaat met typerfouten, synoniemen en homoniemen.
- ⋮ Technische aanpassingen / wensen:
 - Uitbreiden van de filtermogelijkheden van het zoekformulier: alle documenten van de laatste week, laatste maand, laatste kwartaal of van het huidige jaar.
 - In de zoekresultaten is een optie 'meer van hetzelfde' of 'gelijkwaardige pagina's' gewenst. In Google wordt dit gerealiseerd door related search.
 - Formulier metadata: laatst gebruikte trefwoorden en rubrieken (15 of 20) onthouden.
 - De nieuwe zoekfunctionaliteit moet de medewerkers van OCW in staat stellen meerdere domeinen te kunnen bevragen.

4.4 Interviewronde bij pilotmedewerkers en anderen

De interviewronde met twintig medewerkers kende een aantal doelstellingen:

- ⋮ het, in brede zin (Internet, Intranet en g:\-schijven), vergaren van kennis en inzicht rond ontsluiting en toegankelijkheid van informatie;
- ⋮ in zijn algemeenheid toetsen van de computer- en documentaire geletterdheid;
- ⋮ het peilen van de mening van de gebruikers over mogelijk implementatiescenario's.

De vragenlijst is breder van opzet als dat voor het onderzoek noodzakelijk is. Dit komt omdat dit onderzoek 'Auteurs-gegenereerde metadata' onderdeel uitmaakt van het grotere project 'Zoeken en Vinden'.

De vragenlijst is daarvoor opgedeeld in acht onderdelen:

- **Algemeen**
De vragen richten zich op de kenmerken van de respondent (doelgroep, leeftijd, geslacht) en ervaring met de ontwikkelingen rond digitaal werken. Daarnaast een aantal op kwantiteiten gerichte vragen.
- **Toegankelijkheid**
Dit onderdeel bestaat volledig uit stellingen die gericht zijn op de centrale opslag en beschikbaarstelling van documentaire informatie en de rol van de auteur in deze.
- **Ontsluiting / onderwerpsanalyse**
In dit onderdeel wordt de (on)tevredenheid getoetst van de terugvindbaarheid c.q. relevantiebepaling van documenten (op basis van de nieuwe ordenings-structuur en bestandbenaming) en de wijze waarop de (documentaire) informatie toegankelijk wordt gemaakt.
De context van de vragen is gericht op de G:\-schijf.
- **Opslagmedium**
Hier richten de vragen zich op de plaats waar de digitale documenten worden opgeslagen en de wijze waarop overdracht c.q. schoning van deze informatie plaatsvindt.
- **Zoekproces**
Dit deel van de vragenlijst richt zich op de plaats waar informatie (welke bronnen) gezocht wordt en met welk instrumenten. De terugvindbaarheid van de eigen documenten maar ook die gemaakt zijn door anderen wordt hier getoetst. Daarnaast richt een deel van de vragen zich op de documentaire geletterdheid en voor een beperkt deel op de computergeletterdheid.
- **Zoekresultaten**
In dit onderdeel richten de vragen zich iets meer op de interface en de inhoud van het resultatenscherf. Maar ook de grotere zoekmogelijkheden van metadata worden hier getoetst.
- **Kennisdeling**
De bereidheid van de gebruikers tot het delen van informatie wordt in dit onderdeel getoetst.
- **Metadata**
Het laatste onderdeel van het interview is gericht op de toegevoegde waarde van metadata voor het zoekproces en de relevantiebepaling van de resultaten. Daarnaast wordt de respondenten bevraagd over de wijze waarop het vastleggen van metadata het beste ingevoerd kan worden.

4.4.1 Resultaten van de interviewronde (samenvatting)

Algemeen

In totaal hebben 20 medewerkers van OCW zich bereid getoond mee te werken aan het onderzoek. 14 van de 20 medewerkers hebben ook meegedaan aan de pilot 'Auteurs-gegenereerde metadata'. De respondenten vertegenwoordigen de vier van de vijf doelgroepen die als zodanig geïdentificeerd zijn in het Programma 'Digitaal werken bij OCW'. De (functioneel) applicatiebeheerders zijn in dit onderzoek niet meegenomen omdat zij een zeer indirecte rol spelen bij de ontsluiting van informatie.

Omdat de informatievoorziening ondersteunend moet zijn aan de primaire processen van OCW zijn de doelgroepen managers (3) en beleidsmedewerkers (9) het sterkst vertegenwoordigd. De leeftijd is evenredig verdeeld. Er waren geen medewerkers van 60 jaar of ouder bij. In totaal deden 14 mannen en 6 vrouwen mee aan het onderzoek.

Het programma 'Digitaal werken' wordt positief ervaren en gezien als een duidelijke verbetering van de informatievoorziening. Het is niet altijd even helder wat de samenhang van de verschillende deeltrajecten is. Dit zou vaker en duidelijker gecommuniceerd moeten worden door middel van nieuwsbrieven en presentaties.

Gemiddeld worden er per dag per persoon 4,8 Officedocumenten gemaakt. OCW-breed is dat 7000 documenten per dag. Ongeveer 4300 zijn daarvan archiefwaardig. Indicatief zijn dat op jaarbasis 1.450.000 documenten (920.000 archiefwaardig). Documenten worden aan DIV overgedragen als ze ondertekend, verstuurd of geregistreerd zijn. Maar ook als besluitvorming heeft plaatsgevonden. In de meeste gevallen pas als het proces is afgerond.

De toegankelijkheid

Met de stellingen over een centrale opslag, toegankelijkheid en beschikbaarstelling van de functionele documenten was het merendeel van de respondenten (70% en meer) het eens. Een duidelijk voorbehoud werd gemaakt ten aanzien van vertrouwelijke documenten. De status van vertrouwelijkheid die genoemd zijn:

- ⌘ privacyvertrouwelijkheid (burgerbrieven);
- ⌘ personeelsvertrouwelijkheid;
- ⌘ bedrijfs- of organisatievertrouwelijkheid (bijvoorbeeld financiële verslagen en / of gegevens van instellingen).

Men was het er duidelijk over eens dat de auteur niet alleen verantwoordelijk is voor het beschikbaar stellen van functionele informatie maar er ook voor moet zorgen dat die informatie toegankelijk is. Omdat tijd als een belangrijke factor wordt gezien is een adequate omgeving (technisch instrumentarium) voor betrokkenen een uitdrukkelijke voorwaarde.

De mate van toegankelijkheid van de documentaire informatie op de gemeenschappelijke data-servers (G:\-schijven) neemt af naarmate het volume toeneemt. De individuele ervaringen zijn zeer uiteenlopend van positief tot verre van optimaal. Het ontbreken van nadere (beheer)afspraken speelt een belangrijke rol bij de toegankelijkheid van de G:\-schijven.

De ontsluiting

Er is overeenstemming over het feit dat de auteur de meest aangewezen persoon is om de eigen documenten toegankelijk te maken. Juiste omgevingsfactoren worden hierbij wel als voorwaarde gesteld.

Naast opslag op een logische plek in de ordeningsstructuur worden elementen zoals een onderwerp, datum en in sommige gevallen een versie-aanduiding in de bestandbenamingen gebruikt om de toegankelijkheid van de documenten te vergroten. Meer dan de helft (60%) is tevreden over de terugvindbaarheid van de eigen documenten met de huidige ordeningsstructuur. De bestandbenamingen van de eigen documenten geeft voor 75 % voldoende informatie om de relevantie te kunnen bepalen.

Een negatiever beeld treedt op als dezelfde vraagstelling gericht is op documenten van anderen. Over de terugvindbaarheid van 'logisch' in de huidige ordeningsstructuur geplaatste documenten is slechts 40% tevreden. Ook de toegevoegde waarde van bestandsbenamingen gemaakt door anderen lijkt erg persoonsgebonden te zijn. 65% geeft aan dat de benamingen onvoldoende informatie bieden voor het bepalen van de relevantie van documenten.

Het opslagmedium

90% van de respondenten slaat de documenten op de G:\- of H:\-schijf. 15% gebruikt ook de C:\- of D:\-schijf. De I:\- en L:\-schijf worden gebruikt voor informatie-uitwisseling bij projecten en departementale vergaderingen.

De belangrijkste redenen om documenten niet op de G:\-schijf op te slaan zijn de vertrouwelijkheid en de persoonsgebondenheid van documenten. Gemiddeld gaat het hier om 35% van de documenten. De definitieve documenten worden in het algemeen wel op de G:\-schijf opgeslagen. Opslag van functionele e-mail is heel gedifferentieerd.

Centrale criteria voor schoningsacties zijn er niet. De eigen interpretatie van de relevantie van de documenten bepaalt ze bewaard blijven of worden verwijderd. Slechts 2 respondenten hebben aangegeven de DIV te betrekken bij schoningsacties. Behoudens de Triade-procedure (3) is men niet bekend met een procedure / richtlijn hoe de overdracht van documenten / informatie bij functiewisseling moet plaatsvinden. Het is het aan de medewerker zelf hoe hij / zij daar mee omgaat.

Het zoekproces

Over het algemeen zoekt iedereen (80%) zelf zijn informatie. Daar wordt per dag gemiddeld 35 minuten aan besteed. Een beleidsmedewerker besteedt aanzienlijk meer tijd (45 minuten) aan het zoeken van informatie dan een manager (13 minuten). Een derde van de zoektijd wordt door respondenten gevoelsmatig als inefficiënt ervaren. Bij het zoeken naar archiefdocumenten wordt over het algemeen de DIV ingeschakeld.

60% Heeft bij het zoekproces problemen met het vinden van de zowel eigen documenten als documenten van anderen. Ook de frequentie van de problemen die worden ondervonden - van zelden tot 1 á 2 keer in de week - is gelijk hierbij. Men heeft en gunt zichzelf in de bewerkingsfase onvoldoende tijd om documenten met een goede bestandbenaming en op goede plek in de ordeningsstructuur op te slaan.

Een extra element bij het niet kunnen vinden van documenten van anderen zijn:

- ⋮ de documenten zijn opgeslagen op de persoonlijke homedirectory (H:\schijf);
- ⋮ hoge mutatiegraad van medewerkers (geen goede overdracht);
- ⋮ schoning van digitale collecties.

Voor zover e-mail is opgeslagen op de G:\schijf moet deze met de zoekmachine gevonden kunnen worden.

De G:\schijf en het Intranet worden als de belangrijkste bronnen gezien voor het zoeken van informatie. Ook het Internet en het OCW-plein blijken belangrijk te zijn. In tegenstelling tot het Internet (14) wordt op Intranet nauwelijks gezocht met de zoekmachine (5). Hoewel 60% wel weet dat er een zoekmachine is wordt deze niet of nauwelijks gebruikt. Deels omdat de benodigde informatie op een andere wijze gevonden wordt (wegwijzer, via directiesites, attendering) maar grotendeels omdat het zoekproces als moeizaam ervaren wordt en de resultaten niet goed zijn. Voor het zoeken op de G:\- en H:\schijf wordt voornamelijk de Verkenner en de Word-functionaliteiten gebruikt.

Op Internet wordt veruit het meest gebruik gemaakt van Google als zoekmachine. De voorkeur gaat hier naar uit vanwege de gebruikersvriendelijkheid, de snelheid en de goede resultaten (trekzekerheid).

9 van 14 respondenten gebruiken alleen het eenvoudige zoekscherm en 5 gebruiken beide. Voor het vergroten van de zoekresultaten wordt in het algemeen andere zoekwoorden gebruikt. Voor het verkleinen worden er meer zoekwoorden in de zoekopdracht gebruikt. 25 - 35 % gebruikt ook booleaanse operatoren in de verfijning van de zoekopdracht. Zowel een instructie / handleiding voor de definiëring van een zoekstrategie en het gebruik van booleaanse operatoren is gewenst.

De zoekresultaten

Over de getoonde zoekresultaten op het Intranet is men niet tevreden. Het zijn er vooral teveel met een beperkte relevantie of veel van hetzelfde. Voor de bepaling van de relevantie worden alle in het zoekresultaat gepresenteerde gegevens gebruikt. De samenvatting lijkt het belangrijkste te zijn. De meest relevante documenten moeten binnen 20 treffers beschikbaar zijn.

De belangrijkste gegevens die getoond moeten worden in de zoekresultaten zijn: onderwerp, titel en (Google)samenvatting, trefwoorden, auteur. De Google samenvatting wordt 95% als een toegevoegde waarde gezien bij de relevantiebepaling. Verder is men het erover eens dat de presentatie van metadata een toegevoegde waarde heeft. Ook wil men graag kunnen doorzoeken op deze metadata.

Kennisdeling

Bijna niemand heeft bezwaar tegen het feit dat een directe collega zijn / haar documenten kan inzien. Dit bezwaar groeit wel als andere collega's (directie-overstijgend) zijn / haar documenten zouden kunnen inzien. Het bezwaar wordt minder als de status van documenten een meer definitief karakter hebben.

Metadata

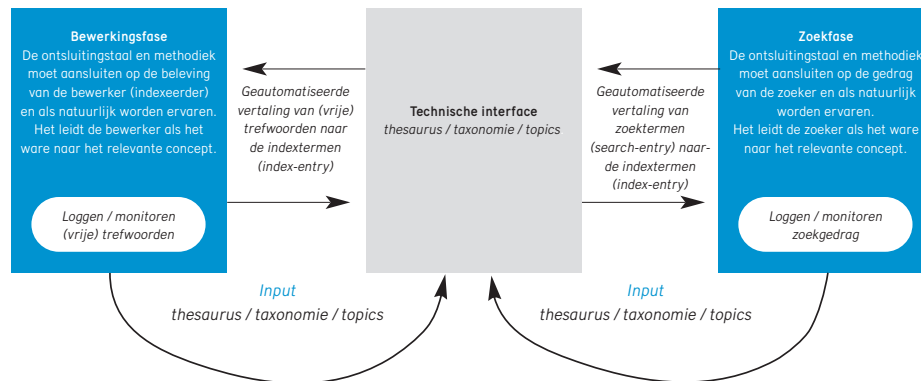
Zowel voor het zoekproces als bij het bepalen van de relevantie van de zoekresultaten wordt (gestructureerde) metadata als een belangrijk instrument gezien. Er worden geen overwegende bezwaren geuit om als auteur verantwoordelijk te zijn voor het vastleggen van metadata. 75% vindt zelfs dat het (onder bepaalde voorwaarden) dwingend opgelegd kan worden. Belangrijke voorwaarden zijn: uniformiteit, goede technische voorzieningen, grote mate van geautomatiseerde vastlegging, beschikbaarheid van adequate opleiding / instructies, etc. Het moment van opslag wordt door betrokkenen (70%) het meest geschikt geacht om metadata vast te leggen.

- Voor de implementatiestrategie zijn de volgende (belangrijkste) aandachtspunten genoemd:
- ⋮ goede en heldere communicatie: voordelen en de toegevoegde waarde presenteren (9);
 - ⋮ zoveel mogelijk geautomatiseerd (7);
 - ⋮ ontwikkeling van gebruikersvriendelijk toegankelijk formaat, veel keuzelijsten (6);
 - ⋮ (centraal) beleid / spelregels (5);
 - ⋮ adequate opleiding / instructie (3);
 - ⋮ door belangrijkste velden verplicht te stellen (kan document niet opslaan als) (4).

4.5 Zoekmachine Verity

Met de toenemende hoeveelheid digitale documenten spelen goede zoekfunctionaliteiten een steeds belangrijkere rol. Niets is zo frustrerend als weten dat het er moet zijn maar het niet kunnen vinden. Daarnaast vervult de techniek een belangrijke intermediaïrende rol tussen de bewerkingsfase en de zoekfase. Onderstaand grafische weergave illustreert dit.

Figuur 2



Wil een zoekmachine de juiste rol en functie kunnen vervullen dan is het van belang om naast de technische (on)mogelijkheden ook de beschikbare kennis om die mogelijkheden te benutten te toetsen binnen de organisatie.

Het is net als bij een geluidsinstallatie: je kunt investeren in de beste CD-speler en versterker maar als de boxen onder de maat zijn blijft het geluid slecht. Het omgekeerde is natuurlijk ook waar. De technische mogelijkheden en de kennis om die te benutten moeten in balans zijn met elkaar en met de gebruiker.

Hiervoor is een vragenlijst gemaakt om er achter te komen wat de (on)mogelijkheden van de Verity zoekmachine zijn en welke toegevoegde waarde deze in het mogelijke metadata-beleid zou kunnen betekenen.

4.6 Kwantitatieve informatie: zoekgedrag, volume G:\schijven, Intranet/OCW-plein

Doel van dit onderzoek was enerzijds gericht op het verkrijgen van meer feitelijke gegevens om de uitkomsten van de andere onderzoeken te verifiëren dan wel te onderbouwen. Immers met feitelijke gegevens (kengetallen) is de urgentie van een betere ontsluiting beter aan te tonen. Anderzijds was dit deel van het onderzoek er op gericht bruikbare informatie te krijgen rond het zoekgedrag van de gebruikers. Deze gegevens vormen een de nodige input voor het onderhoud van het vocabulairum en de verdere ontwikkeling van de zoekfunctionaliteiten.

Toename volume G:\schijf

De volumetoename van de gemeenschappelijke netwerkschijven van vier directies zijn gedurende de onderzoeksperiode gelogd en geanalyseerd. Het volume (aantal opgeslagen documenten) op de G:\schijf neemt per week gemiddeld met 3% toe. Het volume bestaat voor 90% uit Office-documenten (80 % = Word).

Peildatum	Directie 1			Directie 2		
	Aantal	Vershil	Toename (%)	Aantal	Vershil	Toename (%)
28-01-2003	14459			92126		
05-02-2003	17721	3262	18,4%	96526	4400	4,6%
13-02-2003 (incl. rft en txt)	17968	247	1,4%	127521	30995	24,3%
19-02-2003	18144	176	1,0%	130617	3096	2,4%
12-03-2003	19094	950	5,0%	160183	29566	18,5%
26-03-2003	22046	2952	13,4%	150331	-9852	-6,6%
10-04-2003	20233	-1813	-9,0%	150331	0	0,0%
01-05-2003	20652	419	2,0%	178034	27703	15,6%
15-05-2003	20553	-99	-0,5%	178124	90	0,1%
Eindtotaal / gemidd. toename pw in %	20553		2%	178124		4%

Publicaties op het OCW-Intranet en OCW-plein

Volgens de bijgehouden statistieken worden er jaarlijks tussen de 4500 en 5000 (html) documenten op het Intranet gepubliceerd. Voor het OCW-plein ligt dit tussen de 3500 en 4000 stuks.

Zoekgedrag

De zoekactiviteiten van de pilotdeelnemers zijn in de pilotperiode gelogd. Het doel was inzicht te krijgen in de wijze waarop de gebruikers zoeken naar informatie.

4.7 Externe oriëntatie bij ministerie AZ

Door het ministerie van Algemene Zaken wordt onderzoek gedaan in het kader van het gemeenschappelijk webbeleid voor de Rijksoverheid. De aanleiding voor dit onderzoek is de uitvoering van één van de aanbevelingen uit het door de Commissie Toekomst Overheidscommunicatie (Commissie Wallage) uitgebrachte rapport 'In dienst van de democratie'. In deze aanbeveling wordt opzetten van zogenaamde activiteitenindexen of anders gezegd een beleidsagenda door de bestuursorganen verplicht gesteld. De aanbeveling wordt als volgt toegelicht:

'Op deze beleidsagenda of activiteitenindex komen dan alle beleidsonderwerpen en beleidsverantwoordingen voor, die een bestuursorgaan onder handen heeft. Naast inzicht in de bij de overheid beschikbare informatie, biedt een beleidsagenda of activiteitenindex de burger tevens kijk op de democratische procesgang. In Nederland ontbreekt, in tegenstelling tot de ons omringende landen, een voor ieder raadpleegbaar register, waarin alle officiële documenten van de overheid en alle geregistreerde ingekomen stukken zijn opgenomen en dat voor iedereen

raadpleegbaar is. Een verplichte activiteitenindex van ieder bestuursorgaan, dat jaarlijks aan de hand van de begrotingsbehandeling wordt geactualiseerd en dat als bronnen ondermeer de jaarverslagen en plannen van directies en diensten hanteert, kan volgens de werkgroep op termijn tot een dergelijk register leiden. De werkgroep beveelt daarom aan om bij het aantreden van een nieuw kabinet hiervoor een pilot te starten, zodat snel met een dergelijke aanpak ervaring kan worden opgedaan.'

(In dienst van de democratie, Wallage 2001, p 38)

De resultaten van het onderzoek moeten leiden tot een effectievere ontsluiting van de ministeriële beleidsinformatie op basis van een beleidsagenda of activiteitenindex waarbij de volgende ontsluitingsmethodes aan bod komen:

- ⌘ toepassing en gebruik onderwerplijsten;
- ⌘ toepassing en gebruik taxonomieën;
- ⌘ toepassing en gebruik zoekmachines.

Eind juni 2003 wordt het eindrapport verwacht.

4.8 Webcontent-managementsysteem

Ten tijde van dit onderzoek wordt binnen de OCW-organisatie gewerkt aan de departementale invoering van een webcontent-managementsysteem. Het nieuwe systeem wordt per 01-09-2003 ingevoerd waarbij de bestaande content gemigreerd wordt. De lokale redacteurs zijn een belangrijke doelgroep als het gaat om de verrijking (inhoudelijke ontsluiting) van de webpublicaties met metadata op het OCW-Intranet en het OCW-plein.

Vooruitlopend op komend metadatabeleid zijn er pro forma afspraken gemaakt over vast te leggen metadata. Vooral snog worden worden voor het migratietraject de volgende metadata vastgelegd:

Titel	Verplicht, tenzij oude titel letterlijk kan worden overgenomen
Auteur	Optioneel: alleen invullen als auteur daadwerkelijk bekend is, anders geldt de directie als auteur (zie volgend veld)
Directie	Verplicht
Trefwoorden	Nog nader uit te werken
Beschrijving (korte beschrijving van de pagina)	Verplicht bij HTML-pagina's, niet bij downloads (word, pdf, excel)
Datum (datum laatste wijziging)	Verplicht bij HTML-pagina's, niet bij downloads (word, pdf, excel)

4.9 Literatuuronderzoek

In veel elektronische informatiesystemen kan men tegenwoordig gebruik maken van zogenaamde free text zoekfuncties. Dat wil zeggen dat er gezocht kan worden met woorden uit de natuurlijke taal in titels, samenvattingen of volledige teksten. Hierbij is een discussie ontstaan over de meerwaarde van het gebruiken van een gecontroleerd vocabularium, zoals een classificatieschema of woordsysteem (trefwoordenlijst, thesaurus).

Daar waar in handmatige systemen, kaartcatalogi en registers in boekvorm, het werken met natuurlijke taal niet mogelijk is (te veel verschillende termen), speelt in elektronische informatiesystemen het aantal mogelijke zoektermen een veel geringere rol.

In toenemende mate zijn er er informatiesystemen die zowel het zoeken met een gecontroleerd vocabularium ondersteunt als het zoeken met natuurlijke taal.

Zoekmachines zijn op basis van zoekwoorden in de natuurlijke taal **voldoende** in staat om relevante documenten c.q. informatie te vinden. *'Ik krijg altijd wat ik zoek'* is het argument.

Dat is helemaal juist. In een omgeving van het WWW waar miljarden primaire en secundaire bronnen worden aangeboden, zullen de zoekresultaten altijd wel relevante bronnen bevatten (en vaak ook in de eerste tien resultaten), maar de precisie is in hoge mate afhankelijk van hoe goed de gebruiker de zoekopdracht definieert. In de context van een overheidsorganisatie als OCW is dit een ander verhaal. De overheid moet immers verantwoording afleggen aan de maatschappij over de dingen die ze doet en gedaan heeft.

Met de verdergaande digitalisering wordt de afhankelijkheid daarvan steeds groter. Daar waar men in heden ten dage nog terug kan vallen op papieren dossiers, zal dit in de nabije toekomst minder of niet meer mogelijk zijn. Men zal volledig moeten vertrouwen op wat de digitale werkomgeving te bieden heeft.

Een volledige en juiste ontsluiting van de documentaire informatie is daarom noodzaak. Immers bij een goede en gerichte zoekopdracht moet je er op kunnen vertrouwen dat niet alleen de informatie relevant is maar vooral ook volledig.

4.9.1 *Metadata*

Metadata zijn gegevens die een beschrijving geven van een document. Metadata beschrijven de eigenschappen en kenmerken van bepaalde informatie. Zo vertellen ze wat over de herkomst en inhoud van een document, zoals auteur, onderwerp en datum.

In de bibliotheek is dit fenomeen bekend in de vorm van cataloguskaarten waarin de titel, naam van de auteur, datum van uitgifte en een samenvatting van de inhoud zijn opgenomen. Het belang van het gebruik van metadata op de G:\-schijf kan worden duidelijk gemaakt in het voorbeeld van een bibliotheek zonder catalogus: een enorme, maar ontoegankelijke opslag van informatie.

Gewoonlijk destilleert een zoekmachine trefwoorden uit de totale tekst van een document of webpagina en maakt daarbij geen onderscheid welk woord belangrijk is en welk er minder toe doet. Bij een standaard fulltext-indexering van een document hebben woorden als 'beroepskolom' bijvoorbeeld dezelfde status als 'metadata' en 'ook'. De getoonde hiërarchie van zoekresultaten is hierdoor ondoorzichtig. De kans is groot dat met een dergelijke zoekmethode een relevant document pas als laatste in de lijst getoond wordt. En dat is ongewenst.

Naarmate de documentenverzameling groeit wil de gebruiker de meest relevante documenten na een zoekopdracht bij de eerste resultaten terugzien. Door documenten te voorzien van metadata met daarin de belangrijkste beschrijvingselementen van het document, krijgen de opgeslagen documenten ineens een structuur waar een zoekmachine iets mee kan en dus ook een gebruiker wat aan heeft.

Metadata maken het dus voor een zoekmachine makkelijker om de belangrijkste informatie uit de documenten te plukken en in haar (index)database op te nemen.

Een optimaal gebruik van metadata is te behalen als een zoekmachine ook de optie biedt om specifiek te zoeken binnen de metadata-elementen. Een gebruiker kan dan in plaats van een algemeen trefwoord in te typen, aangeven dat hij wil zoeken naar bijvoorbeeld de auteur of de titel van het document. Ook is het denkbaar dat hij of zij een trefwoord geeft dat in een samenvatting zou kunnen voorkomen. Zo zal het aantal resultaten dat de zoekmachine geeft, drastisch worden gereduceerd en wordt de kans dat het juiste document wordt gevonden groter.

4.9.2 *Toepassing van metadata*

Om metadata inderdaad een oplossing te laten zijn voor het goed ontsluiten van informatie op de G-schijf, is het van groot belang direct bij de creatie van documenten metadata toe te voegen. Impliciet betekent dit dat de auteur de meest aangewezen persoon is om metadata toe te voegen. Hij / zij kan inhoudelijk het best bepalen waar het document over gaat.

In de bestandseigenschappen van Officedocumenten kunnen metadata elementen als titel, onderwerp, auteur, trefwoorden, samenvatting, datum, formaat, et cetera, binnen de eigenschap (tag)

worden geplaatst. Hiervoor kan gebruik worden gemaakt van tools die dit deels automatisch doen. Om ervoor te zorgen dat de makers zoveel mogelijk dezelfde soorten metadata hanteren en dus dezelfde 'taal' spreken, moeten afspraken worden gemaakt over welke standaarden gebruikt worden zoals bijvoorbeeld:

- ⌘ een (directiespecifieke) gecontroleerd vocabularium (trefwoordenlijst, thesaurus, etc);
- ⌘ de inhoud van titelveld;
- ⌘ de inhoud van het onderwerpveld;
- ⌘ de auteur;
- ⌘ bestandstypering: nota, memo, brief, notitie, plan van aanpak, et cetera;
- ⌘ mogelijk gebruik van elementen van het ordeningsplan;
- ⌘ het toevoegen van een samenvatting.

Een aantal elementen kan automatisch worden ingevuld, zoals:

- ⌘ auteur, maker of organisatie die primair verantwoordelijk is voor de intellectuele inhoud van het werk (medewerker, afdeling en / of directie);
- ⌘ uitgever die verantwoordelijk is voor het extern beschikbaar stellen van het document (Ministerie van OCW);
- ⌘ datum van de totstandkoming of beschikbaarstelling van het document;
- ⌘ formaat dat wordt gebruikt om de nodige hard- en software te identificeren die nodig is om het bestand te bekijken of uit te voeren;
- ⌘ taal waarin het document is beschreven;
- ⌘ en copyrightverklaring.

Een aantal van de bovengenoemde elementen zal 'verplicht' zijn, de andere zijn optioneel. Als alle gewenste elementen zijn ingevuld, kunnen de metadata worden geïndexeerd door de zoekmachine (Verity). Daarnaast wordt het mogelijk om met een zogenaamde metadatagenerator de documenten en bij publicatie op het Intranet en / of OCW-plein metadata te genereren voor opname in een HTML-document.

In dit onderzoek is gebruik gemaakt van de Qualified Dublin Core metadata. te weten de elementen title, description, author en subject. Deze elementen komen overeen met de simple tags van HTML. Voor de G:\-schijf omgeving is er een element toegevoegd: onderwerp. Dit element is zo vanzelfsprekend voor betrokkenen bij de dagelijkse werkzaamheden dat er niet van is afgeweken.

4.9.3 Metadata Formats

Metadata is ingedeeld in 3 formats. Deze indeling is tot stand gekomen in het project DESIRE⁷. De drie verschillende categorieën met hun kenmerken zijn:

- 1 *Simple formats (metatags voor zoekmachines zoals Lycos, AltaVista en Yahoo)*
 - ⋮ ongestructureerde data;
 - ⋮ worden gebruikt door zoekmachines zoals Altavista en Yahoo;
 - ⋮ worden door webrobots gegenereerd.
- 2 *Structured formats (Dublin Core, IAFA templates, RFC 1807, SOIF, LDIF)*
 - ⋮ gebaseerd op soms niet vastgelegde standaarden;
 - ⋮ bevatten veldbeschrijvingen;
 - ⋮ worden handmatig gemaakt door non specialisten.
- 3 *Rich formats (MARC, ICPSR, CIMI, EAD, TEI)*
 - ⋮ gebaseerd op internationaal vastgelegde standaarden;
 - ⋮ worden handmatig gemaakt;
 - ⋮ vereisen specialisten voor maken en handhaven (onderhouden).

4.9.4. Dublin Core metadatastandaard

De Dublin Core Metadata is een standaard set van 15 elementen waarmee het beschrijven en ontsluiten van elektronische bronnen mogelijk gemaakt wordt. In een workshop, georganiseerd door de OCLC en het National Centre for Supercomputer Applications (NCSA) in maart 1995⁸, zijn deze elementen vastgelegd om de ontwikkeling van een metadata set te stimuleren die gedistribueerde elektronische informatie kon beschrijven.

Ondanks het feit dat het begon als een middel om webdocumenten te beschrijven kreeg het snel aandacht van de formele bronbeschrijvings gemeenschappen zoals musea, bibliotheken, commerciële organisaties enz. Het is een van de belangrijkste standaarden geworden van vandaag de dag. Dublin Core kenmerkt zich door een aantal karakteristieken die het zeer geschikt maken voor het beschrijven van elektronische bronnen. De Dublin Core metadata verschaffen informatie over informatie en kunnen bijvoorbeeld objecten, documenten, afbeeldingen of geluid beschrijven.

⁷ *DESIRE (Development of a European Service for Information on Research and Education) ontwikkelt een visie op het totstandkomen van een netwerk van kruislings doorzoekbare informatie gateways: <http://www.desire.org>*

⁸ *OCLC/NCSA Metadata Workshop Report: <http://www.oasis-open.org/cover/metadata.html>*

Redenen om Dublin Core metadata te gebruiken zijn:

- ⌘ gemakkelijk te maken en te indexeren;
- ⌘ maakt nauwkeuriger zoeken mogelijk;
- ⌘ bevorderen van interoperabiliteit en standaardisatie;
- ⌘ internationaal toepassingsgebied;
- ⌘ multi-disciplinair karakter.

Eenvoud

De Dublin Core is te gebruiken door zowel professionele gebruikers als leken op het gebied van bronbeschrijving. De meeste elementen hebben een algemene, makkelijk te begrijpen semantiek. Dublin Core wordt erkend en ondersteund door meer dan 20 verschillende landen in Europa, Noord-Amerika, Azië en Australië.

Basiselementen

De basiselementen van Dublin Core worden in de volgende tabel beschreven. De structuur van de beschrijving per element is als volgt:

- ⌘ Dublin Core metadata-element (nederlands equivalent)
- ⌘ Dublin Core label van het element;
- ⌘ Omschrijving van het element.

Metadata-element	Dublin Core label	Omschrijving
Titel	DC:Title	De titel van de bron. Meestal gegeven door de auteur, maker of uitgever
Auteur of maker	DC:Creator	De auteur, maker of organisatie die primair verantwoordelijk is voor de intellectuele inhoud van de bron.
Onderwerp / trefwoorden	DC:Subject	Het onderwerp of de inhoud van de bron. Het onderwerp wordt bij voorkeur beschreven door middel van een gecontroleerd vocabulairum.
Omschrijving	DC:Description	Een omschrijving van de inhoud van de bron. Samenvattingen of inhoudsomschrijvingen voor tekstdocumenten en inhoudsomschrijvingen voor niet-tekstdocumenten.
Uitgever	DC:Publisher	De organisatie die verantwoordelijk is voor het publiceren en beschikbaarstellen van de bron in de huidige vorm.
Andere medewerkers	DC:Contributor	Een persoon of organisatie die een belangrijke, maar secundaire, bijdrage heeft geleverd aan de inhoud van de bron.
Datum	DC: Date	De datum van totstandkoming of beschikbaarstelling v.d. bron.
Bestandstype	DC: Type	Categorie of genre van de inhoud van de bron
Format	DC: Format	Het dataformaat van de bron.
Bestandsidentificatie	DC: Identifier	Een string of een nummer ter identificatie van de bron. Dit is bij voorkeur een URI (Internetadres zoals een URN of een URL). Andere voorbeelden zijn de International Standard Book Number (ISBN) of International Standard Serial Number (ISSN).
Bron	DC:Source	Informatie over een andere bron waar de huidige bron van afgeleid is.
Taal	DC: Language	De taal waarin de bron is geschreven.
Relatie	DC: Relation	Informatie over een andere bron die een bepaalde relatie met de huidige Internetbron heeft.
Dekking	DC: Coverage	Dekking of wel indicatie van ruimte en tijd die betrekking heeft op de inhoud van de bron. Inhoudelijk ontsluitingselement.
Copyright	DC: Rights	Een copyrightverklaring, een link naar een copyrightverklaring of een link naar een service die informatie geeft over copyright van de Internetbron.

Deze elementen zijn vervolgens te categoriseren naar inhoud, eigendom en fysieke weergave⁹.

Inhoud	Intellectueel eigendom	Fysieke weergave
Title	Creator	Date
Subject	Publisher	Type
Description	Contributor	Format
Source	Rights	Identifier
Language		
Relation		
Coverage		

De Dublin Core basiselementen (15) kunnen eventueel worden aangevuld met andere elementen: ook wel qualifiers genoemd. Deze qualifiers kunnen zelf worden gedefinieerd, of worden overgenomen van een lijst met aanbevolen elementen: Op 11 July 2000 publiceerde het Dublin Core Metadata Initiative (DCMI) een formele aanbeveling met betrekking tot het gebruik van deze qualifiers.

Het persbericht is te lezen op de url: <http://www.zotgroup.com/development/dcmi/dcqualifiers.htm> en de lijst met qualifiers op de url: <http://purl.org/dc/documents/rec/dcmes-qualifiers-20000711.htm>

Interoperabiliteit

In de bijlage 7 worden de Dublin Core metadata afgezet tegen de metadatatags van HTML en de documenteigenschappen van de Office-omgeving. Eveneens wordt in de tabel aangegeven een eerste aanzet gegeven hoe de ontsluiting van de documentaire informatie in de verschillende informatiedomeinen kan worden ingericht.

Meer informatie over de toepassing van de elementen zijn te vinden op de volgende websites:

- ⋮ Nederlandse vertaling van de Dublin Core Gebruikersgids
<http://www.kb.nl/coop/donor/project-nl-index.html?coop/donor/rapporten/-DCsimpleformat.html>
- ⋮ Engelse usageguide is te vinden op: <http://dublincore.org/documents/2000/07/16/-usageguide/>

⁹

Bron DONOR - Homepage - : <http://www.kb.nl/coop/donor/>

en in de volgende interne onderzoeksrapporten:

- ⋮ Metadata-onderzoek: Kaderstage IDM (Roald Maris, januari 2002)
- ⋮ Ontsluiting van elektronische bronnen d.m.v. Dublin Core metadata (Roald Maris, 18 september 2002).

Internationale ontwikkeling

Internationaal wordt Dublin Core steeds meer gezien als de ontsluitingstaal voor overheid-informatie. In mei 2001 heeft de Britse overheid, één van de toonaangevende landen in de informatierevolutie, met de publicatie 'Towards e - Government, The e-Government Metadata Framework'¹⁰ gekozen voor de Dublin Core Standaard als kader voor de ontsluiting van overheids-informatie. Met de volgende zinsneden, in het voorwoord van de publicatie, geeft de Minister of State, Rt. Hon Ian McCartney MP, het belang van metadata en het hanteren van een metadata-kader weer:

'Getting to the right information fast is the key to internal government efficiency and serving the citizen. Metadata helps us to do that.'

'The Metadata Framework is another important step in our campaign to deliver e-Government and make life simpler for all.'

Dit jaar is de E-Government Metadata Standaard¹¹ voor de Britse overheid, waar de Dublin Core metadata de kern vormen, vastgesteld en verplicht gesteld.

ISO-normering

Op 26 februari 2003 is versie 1.0 (RFC 2413) van de Dublin Core Metadata verheven tot de internationale standaard: ISO 15836:2003(E), Information and documentation - The Dublin Core metadata element set.

¹⁰ UK GovTalk - Metadata - : <http://www.govtalk.gov.uk/schemasstandards/metadata.asp>

¹¹ Office of the e-Envoy: <http://www.e-envoy.gov.uk/>

4.9.5 Trefwoordsysteem

Een informatietaal kan worden gedefinieerd als een kunstmatige taal (vocabulary) voor het ontsluiten van de inhoud en bepaalde formele aspecten van documenten. En als die informatietaal bestaat uit woorden of woordgroepen van een natuurlijke taal, wordt het een woordsysteem genoemd. Globaal kunnen twee soorten woordsystemen worden onderscheiden: de thesaurus en het trefwoordsysteem. De Nederlandstalige definitie omschrijft een trefwoordsysteem als:

'Een precoördinatieve informatietaal, die bestaat uit een geordende verzameling van uit natuurlijke taal gekozen termen, waarvan de vorm, en de onderlinge semantische en syntactische relaties zijn vastgelegd.'
(Magrijn et al., 1997, p. 117)

Het precoördinatieve karakter is het belangrijkste verschil met een thesaurus. Precoördinatie wil zeggen dat onderwerpen in de bouw en indexeerfase zoveel mogelijk worden uitgedrukt in één samengestelde term.

Een trefwoordsysteem is in de regel een gecontroleerd vocabularium. Dit houdt in dat de gebruikte termen aan de juiste voorwaarden moeten voldoen en er regels bestaan over bijv.:

- ⋮ woorddefinitie, woordredactie en woordkeuze;
- ⋮ het gebruik van de woordvorm (enkelvoud of meervoud);
- ⋮ het gebruik van semantische (synoniemen) en syntactische relaties;
- ⋮ het gebruik van bepaalde types termen, zoals homoniemen, eigennamen en samengestelde termen;
- ⋮ de betekenis waarin een bepaalde term gebruikt kan worden.

4.9.6 De trefwoordenlijst van OCW

De trefwoordenlijst van OCW is een van de ADION-thesaurus afgeleide alfabetische lijst van voorkeurstermen en niet-voorkeurstermen. De zoekleiding bestaat uit ZIE (gebruik) en in een beperkte mate van ZIE OOK verwijzingen. De terugverwijzingen (syntactische relaties) ontbreken. Terugverwijzingen zijn niet aangetroffen.

De trefwoordenlijst wordt gebruikt in het documentair management systeem Document Manager voor de onderwerpsontsluiting. Het vaststellen van het onderwerp van het document berust niet op een adequate analyse (object en bepalingen) van de inhoud van het document. Het zoeken van informatie op basis van de trefwoordenlijst vindt postcoördinatief plaats met booleaanse operatoren AND en OR. De trefwoordenlijst wordt beheerd in een Excel-spreadsheet.

4.9.7 Thesaurus

Voor het begrip thesaurus wordt in woordenboeken verwezen naar het Griekse woord 'Thèsaurus' dat vertaald wordt als schatkamer of ook wel verzamelwerk van de gehele woordenschat van een taal. De Nederlandstalige definitie omschrijft een thesaurus als:

'Een postcoördinatieve informatietaal, die bestaat uit een geordende verzameling van uit natuurlijke taal op basis van zoveel mogelijk enkelvoudige begripseenheden gekozen termen, waarvan de vorm, en de onderlinge semantische relaties zijn vastgelegd.'
(Magrijn et al., 1997, p. 95)

Het postcoördinatieve karakter is het belangrijkste verschil met een trefwoordsysteem. Postcoördinatie wil zeggen dat ieder onderwerpselement wordt uitgedrukt in afzonderlijke enkelvoudige of samengestelde term, met de mogelijkheid deze losse termen in de zoekfase te combineren.

Een thesaurus is een gecontroleerd vocabularium wat inhoudt dat de gebruikte termen aan de juiste voorwaarden moeten voldoen en er regels bestaan over bijvoorbeeld:

- ⋮ woorddefinitie, woordredactie en woordkeuze;
- ⋮ het gebruik van de woordvorm (enkelvoud of meervoud);
- ⋮ het gebruik van synoniemen: welke term (indexeerterm) wordt effectief gebruikt bij het indexeren? De andere synoniemen (verwijstermen) verwijzen naar deze voorkeursterm;
- ⋮ het gebruik van bepaalde types termen, zoals homoniemen, eigennamen en samengestelde termen;
- ⋮ de betekenis waarin een bepaalde term gebruikt kan worden.

Ook wordt de aard van de semantische relaties, met name equivalente, hiërarchische en associatieve relaties, in een thesaurus gestandaardiseerd weergegeven.

Equivalentierelatie

Een equivalente relatie (ook wel gelijkwaardigheidsrelatie genoemd) betreft de relatie tussen termen met een gelijke betekenis zoals synoniemen, quasi-synoniemen, afkortingen en samenstellingen. Het regelt verder het gebruik van een voorkeursterm (descriptor) en verwijstermen (non-descriptor). Met het gebruik van de verwijstern USE wordt verwezen naar de voorkeurstern. Met de verwijstern **USE FOR (UF)** wordt de verhouding aangeduid tussen voorkeurstern en verwijstern. Bijvoorbeeld:

AO

UF Administratieve organisatie

Hiërarchische relatie

De hiërarchische relatie groepeert voorkeurstermen in een hiërarchische boomstructuur.

De standaarden spreken van **broader terms (bt)** en **narrower terms (nt)**.

Er zijn twee typen hiërarchische relaties: de generieke (genus >> species) en de partiële relaties (geheel >> deel).

Genus >> species		Geheel >> deel	
Genus (geslacht): fietsen		Geheel:	Armen
Species (soort):	racefietsen	Deel:	bovenarmen
	kinderfietsen		handen
	herenfietsen		onderarmen

Associatieve relaties

De associatieve relaties (**related terms: RT**) zijn relaties tussen termen die een inhoudelijk verwantschap met elkaar hebben en niet als een van de voorgaande relatie worden aangemerkt. Deze associatieve relaties zijn vooral interessant als ze leiden tot nuttige verwijzingen voor de gebruiker.

Scope Note (SN)

Naast deze drie relaties kan aan een term ook een zogenaamde scope note toegevoegd worden. Het kan gaan om een definitie van het gebruik van een term binnen de context van de thesaurus.

Een thesaurus is een goed instrument om de verschillen in taalgebruik tussen de bewerkingsfase (de ontsluiting of indexeerder) en de zoekfase (de informatiezoekende naar informatie) te overbruggen. Een aantal redenen voor het gebruik van een thesaurus zijn:

- ⋮ de natuurlijke taal van auteurs, indexeerders en gebruikers is vertaald naar een gecontroleerde woordenschat;
- ⋮ het verzekert de consistentie in het toewijzen van indextermen;
- ⋮ het geeft de semantische relaties tussen termen weer en
- ⋮ het is een goede zoekleiding.

Naast dit algemene gebruik in information retrieval, stellen Aitchison et al. (2000, p. 1) dat een thesaurus ook kan helpen bij het verwerven van een algemeen begrip van een onderwerpsdomein, bij het opstellen van semantische kenniskaarten en bij het definiëren van termen.

Internationale standaardisering

De term *'thesaurus'* heeft een eenduidige betekenis gekregen in de documentaire informatievoorziening. Er zijn nationale en internationale standaarden ontwikkeld die de belangrijkste aspecten van één- en meertalige thesauri regelen. Voor de ééntalige thesauri zijn de belangrijkste de Amerikaanse ANSI/NISO Z39.19-1993 en de ISO-norm 2788-1986, voor de meertalige gaat het vooral om de ISO 5964-1985.

Voorbeeldsite van een overheidsthesaurus

Government of Canada - Core Subject Thesaurus -

http://en.thesaurus.gc.ca/these/thes_e.html of http://en.thesaurus.gc.ca/bib_e.html

4.9.8 De OCW / ADION-thesaurus

De ADION-thesaurus is in de jaren '70 opgezet als basis voor de onderwerpsontsluiting voor databestanden van het toenmalige Ministerie van Onderwijs en Wetenschappen. Als basis diende de EUDISED-thesaurus (European Documentation and Information System for Education), uitgegeven door het Europees Centrum voor Onderwijsdocumentatie van de Raad van Europa. In 1977 werd deze veeltalige thesaurus vertaald in het Nederlands, in samenwerking met Nederlandstalig België. In 1978 werd Nederlandse nationale terminologie toegevoegd en ontstond de eerste, nog niet gedrukte, editie van de ADION-thesaurus. De tweede versie werd in 1983 gepubliceerd. In 1987 is de thesaurus in beperkte mate herzien en aangepast aan de eisen van die tijd. Zowel de catalogus van de bibliotheek als het documentatiebestand werden en worden ontsloten met behulp van deze thesaurus. De naam van de thesaurus was afkomstig van het ADION-documentatiebestand. De afkorting staat voor Automatisch Documentatie en Informatiesysteem voor Onderwijsliteratuur in Nederland.

In de periode 1994 - 1997 is de thesaurus geheel herzien en uitgebreid met het uit de VWS-thesaurus specifieke cultuur- en mediaterminologie en het incorporeren daarvan in de OCW-thesaurus. Vanaf november 1996 is de cultuur- en mediagerelateerde literatuur met een concept OCW-thesaurus ontsloten.

Doel en doelgroep van de ADION-thesaurus

De thesaurus is samengesteld ten behoeve van de onderwerpsontsluiting van beleidsrelevante literatuur op het gebied van onderwijs, wetenschap, cultuur en media. De thesaurus wordt gebruikt bij het indexeren van en zoeken naar deze literatuur in de catalogus van de bibliotheek van het Ministerie en van informatie in het departementale documentatiebestand. De primaire doelgroepen zijn medewerkers van de ministeriële bibliotheek en de beleidsambtenaren van OCW.

Zowel mensen met een grondige kennis van het onderwerp (maar met weinig ervaring met zoeken m.b.v. thesaurustermen) als mensen met een grote kennis van thesauri en zoeken (en wellicht minder kennis op specifieke vakgebieden) werken binnen de organisatie met dit instrument. In de uitwerking van de opzet van de thesaurus is daarom gekozen voor een grote mate van uitleg van zowel de werking van de thesaurus als van de inhoudelijke aspecten van de descriptorren.

Het gebruik van de thesaurus

Voor een goed gebruik wordt de OCW-thesaurus op diverse wijzen uitgebracht: een alfabetisch deel, een gefacetteerd deel, een kwic-index en een deel met normlijsten.

4.9.9 Rubrieken

Bij het ontwerp van een informatiesysteem is het van belang een indruk te krijgen van de omvang van de indexbestanden en het werk dat het kost om deze regelmatig te actualiseren.

Zowel het nauwelijks ontsluiten van documenten, met als gevolg een oceaan aan gegevens (veel treffers), als alle documenten uitputtend ontsluiten, is duur. Hier moet een balans in worden gevonden. Daarvoor is binnen OCW een zogenaamde rubriekenlijst ontstaan. De rubriekenlijst is een set van ongeveer 100 rubrieken (trefwoorden) die het mogelijk maakt om de vele (deel)onderwerpen binnen OCW van een rubricering te voorzien. De rubriekenlijst is voorzien van een nadere explicitering van wat onder de betreffende rubriek wordt verstaan.

4.9.10 Metadata-repository

In de onderzoeksperiode geleidelijk aan een beeld ontstaan over de behoefte aan een centrale bron voor metadata (metadata-repository). De term metadata-repository wordt in deze context gebruikt als een centrale database die als bron dient voor de ontsluiting van de departementale informatie.

Elementen van die database zijn:

- ⋮ een departementaal vocabularium (trefwoordsysteem of bij voorkeur een thesaurus);
- ⋮ het departementale ordeningsplan;
- ⋮ de handelingen van OCW en de daaraan gekoppelde bewaar- of vernietigingsindicaties;
- ⋮ de departementale thema's c.q. primaire activiteiten van OCW (zie § 4.7) en
- ⋮ de daaraan gelieerde werkprocessen.

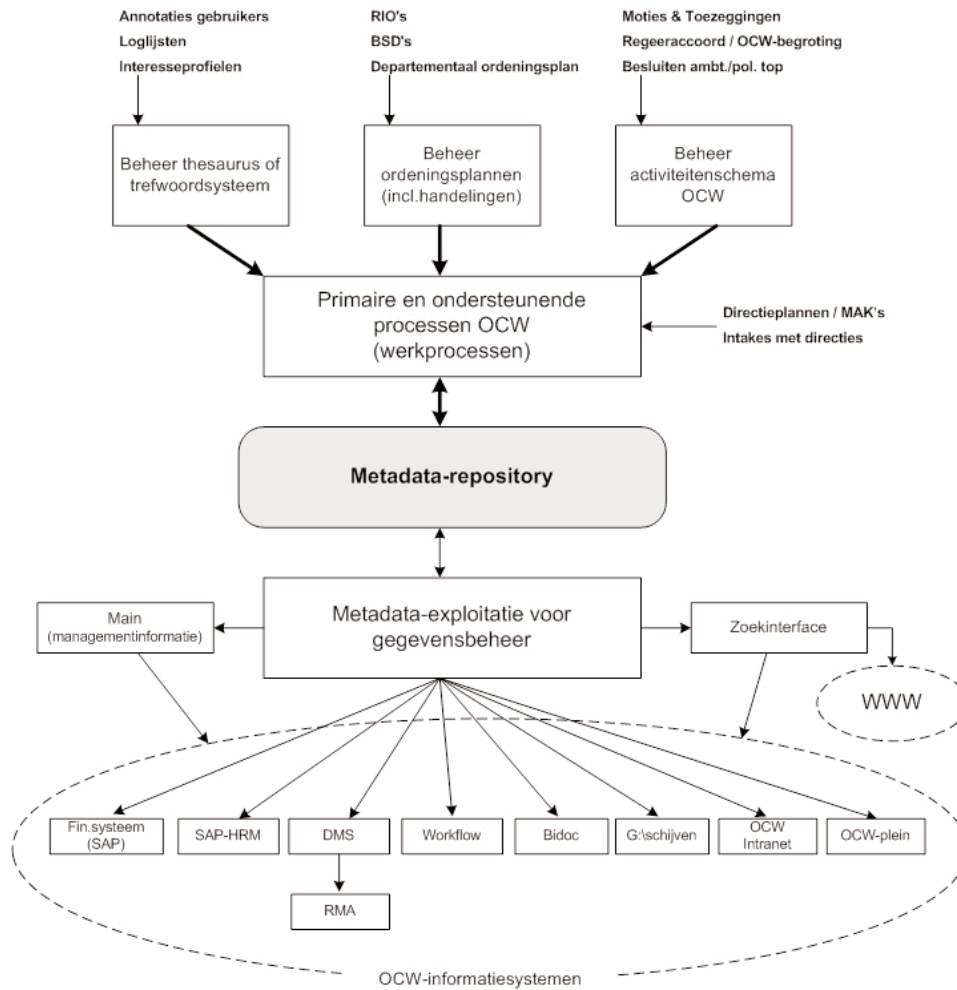
De grafische weergave van de metadata-repository database wordt in figuur 2 weergegeven. Het hart van de database zijn de werkprocessen van het bestuursdepartement van OCW. De ontsluiting van deze werkprocessen vindt centraal en vooraf plaats met thesaurustermen. Relaties worden gelegd met het departementaal ordeningsplan en de overkoepelende thema's (activiteiten-index). De activiteitenindex is één van de wegen naar buiten en kan als OCW-navigatiestructuur (taxonomie) gepresenteerd worden op het OCW-plein.

De thesaurus kan daarnaast in de verschillende representatievormen als zoekleiding fungeren voor zowel de interne OCW-medewerker als voor de informatiezoekende burger. Als de metadata-repository bijvoorbeeld via de werkprocessen gelinkt wordt aan de Active Directory, wordt het mogelijk om kennisprofielen te genereren. Uiteindelijk kunnen de Dublin Core metadata, bij de registratie van een document, bijna geheel geautomatiseerd worden toegekend. Uiteraard moet het toekomstig beleid er op gericht zijn dat alle relevante informatiesystemen (documentair, financieel, workflow, et cetera) op termijn gekoppeld worden aan deze centrale database.

Het voorgaande geeft een globale indruk van de functie van de metadata-repository als kennis-systeem voor de interne en externe informatievoorziening. Mede door een eenduidige ontsluiting van departementale (documentaire) informatie wordt de samenhang in informatievoorziening beter geborgd.

Omdat dit onderwerp verder buiten de scope van het onderzoek valt, is naar de levensvatbaarheid van een metadata-repository verder geen onderzoek verricht.

Figuur 2 metadata-repository



5 Conclusies en aanbevelingen

5.1 Interne en externe informatievoorziening

Beschikbaarheid en beschikbaarstelling zijn kernbegrippen in de wereld van de documentaire informatievoorziening. In toenemende mate worden er digitale (eind)producten geproduceerd. De ontwikkelingen bij de rijksoverheid rond het digitaal werken stimuleren dit. Deze documenten zijn van belang voor het afleggen van verantwoording, de reconstructie van het gevoerde beleid en op termijn voor het cultureel erfgoed. Voor wat betreft de eerste twee belangen wordt de maatschappelijke druk steeds groter. Van de overheid wordt verwacht dat zij op een efficiënte en transparante wijze haar taken voor de maatschappij uitvoert. Dit staat hoog op de politieke agenda. Zowel de OCW-medewerker als een burger moeten makkelijk en snel kunnen beschikken over alle gewenste informatie. Uiteraard is de informatie betrouwbaar en voorzien van actuele statuskenmerken. Ook de 24 uren economie speelt ook een steeds grotere rol. Het tijd- en plaats-onafhankelijk werken lijkt tot het gangbare leven te gaan behoren. Om dit te realiseren is een goed georganiseerde informatievoorziening nodig.

Hierop inspeland heeft OCW besloten een verdere uitbouw van het OCW-plein en een betere ontsluiting van het OCW-Intranet te realiseren. Per 1 september 2003 wordt het nieuwe webcontentmanagement-systeem geïmplementeerd.

De advisering en ondersteuning van directies bij het inrichten en ontsluiten van de G:\-schijf (gemeenschappelijke dataserver) is een actueel onderwerp van gesprek. Ook is er aandacht voor instrumentontwikkeling voor kennisdeling van impliciete kennis en voor ontsluiting van alle beschikbare relevante bronnen met één zoekfaciliteit.

Deze ontwikkelingen maken het noodzakelijk dat de toegankelijkheid en beschikbaarheid van de documentaire informatie anders georganiseerd moet worden en past bij de moderne ontsluitingsmechanismen. De rol van metadata hierbij is evident.

Het is voor de consistentie in de organisatie van de informatievoorziening belangrijk om aan te sluiten bij reeds ontwikkelde standaarden voor ontsluiting en het beschikbaar stellen van documentaire informatie.

De internationale Dublin Core Metadata standaard¹² kan daar als ontsluitingstaal een belangrijke rol bij vervullen. Mede ook omdat webtechnologie op dit moment een juiste keuze is voor het beschikbaar stellen van documentaire informatie, zowel voor de interne als externe informatievoorziening.

¹² ISO 15836:2003(E), Information and documentation - The Dublin Core metadata element set

Aanbevelingen

- Vindt niet het wiel opnieuw uit, maar sluit aan bij de internationale ontwikkelingen en standaarden.
- Werk met een internationaal geaccepteerde standaard (de Dublin Core Metadata) voor metadata ten aanzien van de technische eisen. Op deze manier kan gebruik worden gemaakt van 'proven technology'.
Deze standaarden hebben hun nut al bewezen. Informatie-ontsluiting veronderstelt metadata gebruik. Internationaal hebben verschillende overheidsorganisaties, zoals VS, Australië, Groot-Brittanië, dit al verplicht gesteld.
- Intensiveer de interdepartementale samenwerking over de ontsluiting van overheidsinformatie. Werk toe naar een Dublin Core metadatastandaard voor de rijksoverheid.
- Zorg voor een gemeenschappelijke overheidsbrede visie op de ontsluiting van overheidsinformatie en zorg voor een gemeenschappelijk punt binnen de overheid van waaruit aansturing plaatsvindt.

5.1.1 Gemeenschappelijk dataserver (G:\schijf)

Voor de informatievoorziening is de functie 'Zoeken en vinden' de functie waar het allemaal om draait. Het is van belang om in de gestaag groeiende omvang van digitale documenten sneller en preciezer de relevante informatie c.q. documenten te vinden. Afgezien van de tijdsbesparing in het zoekproces worden er beslissingen genomen op grond van de beschikbare informatie. Een groot deel van de documentaire informatie wordt opgeslagen op de gemeenschappelijke dataservers: de G:\-schijven.

De nieuw ingevoerde ordeningsstructuur op de G:\-schijf is inmiddels op de lagere niveaus (derde niveau en verder) fors uitgebreid en een mate van wildgroei wordt daarbij geconstateerd. De opslag van de documenten is onvoldoende eenduidig en er zijn interpretatieverschillen over de plek van opslag in de structuur. Ook worden dezelfde documenten op verschillende plekken op de G-schijf opgeslagen. Reacties van de respondenten en de nazorggesprekken (digitalisering fase 1) versterken dit beeld.

Naast opslag op een logische plek in de ordeningsstructuur worden elementen zoals onderwerp, datum en in sommige gevallen een versie-aanduiding in de bestandsbenamingen gebruikt om de toegankelijkheid van de documenten te vergroten. Dit is persoonsgebonden. De gebruikers zijn namelijk meer tevreden over de terugvindbaarheid van **de eigen** documenten dan van de terugvindbaarheid van documenten **van anderen**. Vooral de bestandbenamingen worden onvoldoende bevonden voor het bepalen van de relevantie.

Bij de opslag en plaatsing van documenten speelt tijd een belangrijke rol. Men heeft of gunt zich niet de tijd om een document met een goede naam op de juiste plek in de ordeningsstructuur op te slaan. Ook de opslag van functionele documenten op de persoonlijke H:\-schijven beïnvloedt het vinden van relevante informatie negatief.

De gemiddelde toename van het volume (aantal documenten) op de G:\-schijven is volgens de indexeringsstatistiek van Verity 3% per week.

De toegankelijkheid en beschikbaarheid van de documentaire informatie wordt negatief beïnvloed door het ontbreken of niet uitvoeren van (de)centrale beheerprocedures en -richtlijnen.

Er zijn geen (centrale) criteria voor overdracht van informatie c.q. kennis bij bijvoorbeeld functiewisseling, uitdiensttreding en schoningsacties. Doorgaans wordt op basis van eigen interpretatie bepaald of documenten bewaard blijven of worden verwijderd. Hierdoor gaat kennis verloren of dreigt verloren te gaan.

Aanbevelingen

- Stel criteria op voor de overdracht van informatie c.q. kennis bij functiewisseling en /of uitdiensttreding en stel hier een departementaal format beschikbaar.
- De functionele documenten met een status van definitief-concept dienen op de G:\-schijf van de directie te worden opgeslagen. De vertrouwelijke documenten moeten met behulp van een autorisatiemodel van algemene openbaarheid worden uitgezonderd. Als gevolg hiervan dient de zoekmachine het autorisatiemodel te ondersteunen. Stel hier een beleidskader voor vast.

5.1.2 Organisatie van de informatievoorziening

Er zijn weinig bezwaren tegen een centrale opslag, toegankelijkheid en beschikbaarstelling van de functionele documenten. Wel wordt er een duidelijk voorbehoud gemaakt over de vertrouwelijkheid van documenten. De vertrouwelijkheid in deze context gaat verder dan alleen personeelsvertrouwelijkheid. Hierbij is te denken aan privacyvertrouwelijk, politiek vertrouwelijk, bedrijfs- en beleidsvertrouwelijk. Bij de implementatie van zoekinstrumenten dient hiermee rekening te worden gehouden, alsmede bij de invoering van het metadatabeleid.

Het als auteur verantwoordelijk zijn voor de toegankelijkheid en beschikbaar stellen van functionele informatie wordt niet als een probleem gezien. Ook niet als zij daarvoor metadata **moeten** vastleggen. Men acht dit zelfs noodzakelijk. Daar worden wel voorwaarden aan gesteld zoals uniformiteit en gebruikersvriendelijke invoer- en zoekfunctionaliteit.

Tegen kennisdeling is weinig bezwaar als het gaat om het delen met directe collega's. Het bezwaar groeit echter als het gaat om kennisdeling met andere collega's (directie-overstijgend). De status van documenten speelt hierbij een belangrijke rol, maar het bezwaar wordt minder als de status van documenten een meer definitief karakter hebben.

Aanbevelingen

- Zorg voor een gemeenschappelijk departementale visie op de ontsluiting van documentaire informatie en stel departementaal beleid voor metadata vast op basis van de Office document-eigenschappen, HTML-metadata (simple tags) en de Dublin Core metadata.
- Maak ter uitvoering van het beleid zoveel mogelijk gebruik van 'proven technology' en het beschikbare instrumentarium.
- Draag er zorg voor dat de toekenning van metadata zoveel mogelijk geautomatiseerd kan plaatsvinden.
- Zorg voor een gebruikersvriendelijke invoer- en zoekfunctionaliteit.

5.1.3 Instrumentarium

De nieuwe sjabloontechologie (Smart Documents) biedt voldoende mogelijkheden voor het geautomatiseerd vastleggen van metadata bij de bron zoals titel, onderwerp, auteur, trefwoorden en rubrieken. De sjabloontechologie maakt per 1 september 2003 onderdeel uit van de standaard-werkplek.

De G:\-schijf en het Intranet blijken de belangrijkste bronnen van informatie te zijn. Maar ook het Internet en het OCW-plein veelvuldig geraadpleegd.

Binnen de huidige kantoorautomatisering (werkplekstandaard) worden separate zoekfunctionaliteiten aangeboden voor de verschillende informatiedomeinen. Dit wordt door de gebruikers als inefficiënt en en onduidelijke ervaren.

Aanbevelingen

- Vastgesteld moet worden of de zoekfunctionaliteiten, die nu naast elkaar worden aangeboden, geoptimaliseerd en geïntegreerd kunnen worden om vollediger en betere zoekresultaten te verkrijgen.
- Pas de nieuwe sjabloontechologie aan, zodanig dat de vastlegging van de basis metadata-elementen titel, onderwerp, auteur en organisatie-eenheid en rubriek (als verplicht veld) volledig geautomatiseerd kan plaatsvinden. Implementeer eveneens een gebruikersvriendelijk metadataformulier om de gebruiker in staat te stellen om metadata toe te voegen en/of te muteren.

- Biedt de gebruiker de mogelijkheid om zelf één of meerdere vrije trefwoorden toe te kennen binnen de sjabloontechnologie.
- Pas het zoek- en resultatscherm zoveel mogelijk aan aan de geïnventariseerde gebruikerswensen en implementeer de nieuwe zoekfunctionaliteit bij de directies.

5.1.4 *Trefwoordsysteem versus thesaurus*

Een thesaurus is bij uitstek geschikt voor het gebruik in een geautomatiseerde documentaire omgeving. Belangrijke voordelen van de thesaurus ten opzichte van een trefwoordsysteem zijn de verschillende presentatievormen.

Een hiërarchische representatie van de thesaurus biedt in een geautomatiseerde omgeving de generische zoekmogelijkheden en een goede navigatiestructuur (taxonomie). De thesaurus is hierdoor in staat om als 'beperkt' kennissysteem te functioneren. Gecombineerd met andere gevalideerde en gecontroleerde tabellen (zie grafische weergave in § 4.9.9.) neemt de thesaurus als kennissysteem in kracht en functionaliteit toe. De facetindeling is een niet minder krachtige representatie van de thesaurus.

Een ander belangrijk voordeel is dat de termen (descriptor) als concept in een betekenisvolle context zijn geplaatst. Door het gebruik van de verwijzingsystematiek (hiërarchisch en relationeel) krijgt de descriptor een betekenisvolle waarde. En als laatste voordeel ten opzichte van de trefwoordenlijst zijn de multilinguïstische mogelijkheden. In het licht van de Europese eenwording is dit een niet te onderschatten voordeel.

Aanbevelingen

- Doe nader onderzoek naar de toegevoegde waarde van een thesaurus als ontsluitings- en zoekinstrument (zoekleiding) voor de totale interne en externe informatievoorziening van OCW. Neem in de kosten en batenanalyse de rol van de thesaurus als kennissysteem en als onderdeel van de metadata-repository mee.
- Bij een positief onderzoeksresultaat: implementeer de thesaurus en faseer de trefwoordenlijst als ontsluitingsinstrument voor de documentaire informatievoorziening uit.

5.2 **Metadata-repository**

Er worden in toenemende mate informatiesystemen verworven om een bepaald bedrijfs- of werkproces te ondersteunen en faciliteren. Van belang is te beseffen dat de informatie en kennis die in de systemen zijn opgeslagen veelal op een geheel eigen wijze vastgelegd, ontsloten en beheerd worden. De samenhang ontbreekt.

Datamining en datawarehouses zijn dan de oplossing om geïntegreerde en samenhangende informatie uit de systemen te trekken.

Een metadata-repository kan, net als bij een Active Directory, een belangrijke centrale rol vervullen bij de structurering en ontsluiting van de interne en externe informatievoorziening.

Dit concept is in de onderzoeksperiode verder niet onderzocht of uitgewerkt. Nader onderzoek moet uitwijzen of dit een levensvatbaar concept is. Zo ja, dan wordt de effectiviteit van de activiteiten op het gebied van de informatievoorziening waarschijnlijk ongekend groot.

Aanbeveling

- Doe nader onderzoek naar de toegevoegde waarde van een metadata-repository als ontsluitingsinstrumentarium en als zoekleiding voor de interne en externe informatievoorziening.

5.3 Implementatiestrategie

Gezien de groeiende hoeveelheid digitale documenten is het organisatorisch wenselijk om het toekennen van metadata bij de maker (de bron) te beleggen. Het (achteraf) toekennen van metadata door bijvoorbeeld DIV-afdelingen zou een te groot beslag op deze capaciteit leggen en de kwaliteit van de terugvindbaarheid negatief beïnvloeden. Het kiezen voor toekennen aan de bron vereist echter wel voorzieningen en instructies waarmee eindgebruikers uit de voeten kunnen. Het doen beseffen van het belang van het toekennen van goede metadata voor de terugvindbaarheid dient hierbij voorop te staan.

Het laten vastleggen van één of meerdere trefwoorden uit de gecontroleerde trefwoordenlijst zal weerstand opleveren. Verrijking (een verdere ontsluiting) van de documentaire informatie kan plaatsvinden voor de archiefwaardige documenten door de decentrale DIV-medewerkers en voor de te publiceren documenten op het Intranet / OCW-plein door de lokale redacteurs.

Het wordt belangrijk gevonden dat de nieuwe functionaliteiten stap voor stap geïmplementeerd worden. Als de implementatie van nieuwe instrumenten het veranderingsvermogen te boven gaat of niet direct tot gewenste resultaten leidt, ontstaat er een drempel om met de nieuwe functionaliteiten aan de slag te gaan. De veranderingsbereidheid neemt hierdoor af.

Alle geboden functionaliteiten (organisatorisch, technisch en instrumenteel) moeten in hoge mate gebruikersvriendelijk zijn. Dit is een uitdrukkelijk acceptatiecriterium.

Aanbevelingen

- Implementeer het metadata-beleid gefaseerd. Bepaal in de implementatiestrategie welke producten het snelst tot zichtbare resultaten leiden.
- Betrek de gebruikers bij het ontwikkelen, implementeren en accepteren van de nieuwe functionaliteiten.

5.3.1 Communicatie

Het traject 'Auteurs-gegenereerde metadata' is een positieve volgende stap in het Digitaliseringstraject. Het is echter voor de OCW-medewerkers niet altijd even helder wat de samenhang is tussen de verschillende deeltrajecten. Een goede communicatiestrategie is één van de belangrijkste factoren voor een succesvolle implementatie. De veranderingsbereidheid wordt vergroot door de gebruikers er in een vroeg stadium bij te betrekken. Het zijn de gebruikers die de verandering moeten dragen en uitvoeren.

Door op het juiste moment, op passende wijze, tussen alle betrokkenen een heldere communicatie tot stand te brengen, wordt de kans op het welslagen van de innovatie groter.

In dit kader worden demonstratie- en instructiebijeenkomsten, waarin de gebruikers leren om de schermen voor metadata en de zoekfunctionaliteit goed te gebruiken, als zeer belangrijk gezien. Ook een adequate ondersteuning in de vorm van een goed bereikbaar helpdesk is een goed communicatiemiddel. Integratie in bestaande helpdesk(ken) garandeert de continuïteit beter.

Aanbevelingen

- Breng op het juiste moment, op passende wijze, tussen alle betrokkenen een heldere communicatie tot stand. Neem het belang van metadata op in de communicatiestrategie van het programma 'Digitaal werken bij OCW 2003-2006'
- Realiseer op de gebruikers toegesneden instructies, handleidingen en opleidingen die relevant zijn bij de invoering van metadata. Biedt deze zowel digitaal (Intranet) als fysiek aan.
- Dit dient vooraf te gaan aan de daadwerkelijke implementatie.
- Realiseer presentatie- en demonstratiebijeenkomsten voor de verschillende doelgroepen.
- Besteedt in de nieuwsbrief van het werkprogramma nadrukkelijk aandacht aan het belang en toegevoegde waarde van metadata.

6 De bron ontsloten met Dublin Core metadata

Metadata-element	Dublin Core label	Omschrijving
Titel	DC:Title	Haalbaarheidsonderzoek Auteurs-gegenereerde metadata
Auteur of maker	DC:Creator.PersonalName DC:Creator.CorporateName	Rombout, Wil Ministerie Onderwijs, Cultuur en Wetenschap
Onderwerp / trefwoorden	DC:Subject	advies; automatisering; beleid; contentmanagement; div; document; documentair beleid; documentaire informatievoorziening; gegevensbeheer; informatievoorziening; kennismanagement; metadata; onderzoek; ontsluiting; ontsluitingstechniek; thesaurus; trefwoorden; vocabularium
Omschrijving	DC:Description	Dit eindrapport beschrijft de te behalen doelen en de implementatiestrategie voor project Auteurs-gegenereerde metadata. Dit project is onderdeel van het departementsbrede werkprogramma Digitaal werken bij OCW. Een belangrijk element van het werkprogramma 'Digitaal werken' is dat medewerkers hun taken beter kunnen uitvoeren en dat de informatievoorziening hen daar efficiënt in ondersteunt. Zij moeten daartoe documentaire informatie op een goede en toegankelijke wijze kunnen opslaan en vooral ook terug kunnen vinden. Als uiteindelijk resultaat betekent dit, dat alle informatie met één 'user-interface' plaats- en tijdonafhankelijk benaderbaar moet zijn – ongeacht de plek van opslag – via op elkaar afgestemde en marktconforme 'Storage & retrieval' technieken. Om dat te kunnen bereiken is een eenduidige ontsluiting van de documentaire informatie noodzakelijk.
Uitgever	DC:Publisher	Ministerie Onderwijs, Cultuur en Wetenschap
Andere medewerkers	DC:Contributor.PersonalName	Seitzinger, Ruud
Datum	DC: Date	2003-08-01
Bestandstype	DC: Type	Tekst
Format	DC: Format	application/msword
Bestandsidentificatie	DC: Identifier	
Bron	DC: Source	
Taal	DC: Language	NI
Relatie	DC: Relation	
Dekking	DC: Coverage	
Copyright	DC: Rights	Ministerie van Onderwijs, Cultuur en Wetenschap

6.1 Voorbeeld uitwerking metadata (html en Dublin Core metadatatags)

<!--General metadata for Search Engines -->

```
<title>Haalbaarheidsonderzoek Auteurs-gegenereerde metadata</title>
<META HTTP-EQUIV="content-type" CONTENT="text/html; charset=ISO-8859-1">
<META NAME="title" CONTENT=" Haalbaarheidsonderzoek Auteurs-gegenereerde metadata">
<META NAME="author" CONTENT="Rombout, Wil">
<META NAME="author" CONTENT="Ministerie van Onderwijs, Cultuur en Wetenschap">
<META NAME="copyright" LANG="nl" CONTENT="Ministerie van Onderwijs, Cultuur en
Wetenschap">
<META NAME="description" LANG="nl" CONTENT="Dit eindrapport beschrijft de te behalen doelen
en de implementatiestrategie voor project Auteurs-gegenereerde metadata. Dit project is onder-
deel van het departementsbrede werkprogramma Digitaal werken bij OCW. Een belangrijk element
van het werkprogramma 'Digitaal werken' is dat medewerkers hun taken beter kunnen uitvoeren en
dat de informatievoorziening hen daar efficiënt in ondersteunt. Zij moeten daartoe documentaire
informatie op een goede en toegankelijke wijze kunnen opslaan en vooral ook terug kunnen vinden.
Als uiteindelijk resultaat betekent dit, dat alle informatie met één 'user-interface' plaats- en tijd-
onafhankelijk benaderbaar moet zijn – ongeacht de plek van opslag – via op elkaar afgestemde en
marktconforme 'Storage & retrieval' technieken. Om dat te kunnen bereiken is een eenduidige
ontsluiting van de documentaire informatie noodzakelijk.">
<META NAME="keywords" LANG="nl" CONTENT="advies; automatisering; beleidsplan; content-
management; div; document; documentair beleid; documentaire informatievoorziening; gegevens-
beheer; informatievoorziening; kennismanagement; metadata; onderzoek; ontsluiting; ontsluitings-
techniek; thesaurus; trefwoorden; vocabularium">
```

<!-- Dublin Core Metadata -->

```
<!-- These metadata have been made with the DONOR metadatagenerator. -->
<!-- http://www.konbib.nl/donor/mg-start.html -->
```

<!-- relevant links for used schemes -->

```
<LINK REL="IMT" HREF="ftp://ftp.isi.edu/in-notes/rfc2046.txt">
<LINK REL="ISO8601" HREF="http://www.w3.org/TR/NOTE-datetime">
<LINK REL="ISO639-1" HREF="ftp://ftp.isi.edu/in-notes/rfc1766.txt">
<LINK REL="URL" HREF="ftp://ftp.isi.edu/in-notes/rfc1808.txt">
```

<META NAME="DC.Title" CONTENT="Haalbaarheidsonderzoek Auteurs-gegenereerde metadata">
<META NAME="DC.Creator.PersonalName" CONTENT="Rombout, W.A.">
<META NAME="DC.Creator.CorporateName" CONTENT="Ministerie Onderwijs, Cultuur en Wetenschap">
<META NAME="DC.Subject" CONTENT="advies; automatisering; beleidsplan; contentmanagement; div; document; documentair beleid; documentaire informatievoorziening; gegevensbeheer; informatievoorziening; kennismanagement; metadata; onderzoek; ontsluiting; ontsluitingstechniek; thesaurus; trefwoorden; vocabularium">
<META NAME="DC.Description" CONTENT="Dit eindrapport beschrijft de te behalen doelen en de implementatiestrategie voor project Auteurs-gegenereerde metadata. Dit project is onderdeel van het departementsbrede werkprogramma Digitaal werken bij OCW. Een belangrijk element van het werkprogramma 'Digitaal werken'; is dat medewerkers hun taken beter kunnen uitvoeren en dat de informatievoorziening hen daar efficiënt in ondersteunt. Zij moeten daartoe documentaire informatie op een goede en toegankelijke wijze kunnen opslaan en vooral ook terug kunnen vinden. Als uiteindelijk resultaat betekent dit, dat alle informatie met één 'user-interface' plaats- en tijd-onafhankelijk benaderbaar moet zijn - ongeacht de plek van opslag - via op elkaar afgestemde en marktconforme 'Storage & retrieval' technieken. Om dat te kunnen bereiken is een eenduidige ontsluiting van de documentaire informatie noodzakelijk. ">
<META NAME="DC.Publisher" CONTENT="Ministerie van Onderwijs, Cultuur en Wetenschap">
<META NAME="DC.Contributor.PersonalName" CONTENT="Seitzinger, R.">
<META NAME="DC.Date" SCHEME="ISO8601" CONTENT="2003-08-01">
<META NAME="DC.Type" CONTENT="text">
<META NAME="DC.Format" SCHEME="IMT" CONTENT="application/msword">
<META NAME="DC.Language" SCHEME="ISO639-1" CONTENT="nl">

7 Bronvermeldingen

7.1 Literatuur en publicaties

- ⋮ *AIV Informatiekunde voor het HBO*. Derksen en Crins (2000). Academic Service, Schoonhoven
- ⋮ *Advieskunde voor praktijkstages, organisatieverandering als leerproces*. Kempen en Keizer (2000). Wolters Noordhoff BV, Groningen
- ⋮ *Basisboek Methoden en technieken: Handleiding voor het opzetten en uitvoeren van onderzoek*. Baarda en de Goede (2001). Stenfert Kroese: Groningen, 3e herziene druk.
- ⋮ Beheersdocument G:\schijf, FacB: september 2001
- ⋮ *Begrip in document management*. Ekelenburg, P. van, Waard J. de (1999). F&G publishing.
- ⋮ *Bibliotheek en documentatie: Handboek ten dienste van de opleidingen*. Schneiders, dr. P., et al (1977). Van Loghum Slaterus bv: Deventer, 3e herziene en vermeerderde druk 1984.
- ⋮ Brief Secretaris-generaal over de taakstelling. OCW: 13 oktober 2002
- ⋮ *Burger en overheid in de informatiesamenleving: De noodzaak van institutionele innovatie*. Eenmalige Adviescommissie ICT en Overheid: Den Haag, september 2001.
- ⋮ *De probleemstelling van een onderzoek*. Verschuren, P.J.M. (1986). Hentenaar: Wijk bij Duurstede, 2e verbeterde druk 1988.
- ⋮ *Documentaire informatiesystemen*. Meer, K. van der (1998). NBLC Uitgeverij: Den Haag
- ⋮ *Eindrapport Elektronisch Openbaarheid: Naar een transparantere overheid door proces-innovatie*. Het Expertise centrum, 's-Gravenhage, 8-05-2002
- ⋮ *In dienst van de democratie*. Het rapport van de Commissie Toekomst Overheidscommunicatie. SDU Uitgevers: Den Haag, augustus 2001.
- ⋮ *Inhoudelijk ontsluiting: begrippenkader en begripsvorming*. Haagse Hogeschool: Den Haag
- ⋮ *Informatie-analyse: een inleiding*. Haagse Hogeschool: Den Haag
- ⋮ *Notitie 'Ontsluiting van elektronische documenten d.m.v. Dublin Core metadata'*. Roald Maris, 18-09-2002
- ⋮ *Onderzoeksplan 'Auteurs-gegenereerde metadata' (Onderzoeksplan_metadata.doc)* Wil Rombout, januari 2003
- ⋮ *Ontsluiting van elektronische documenten d.m.v. Dublin Core metadata*. Roald Maris, 18 september 2002.
- ⋮ *Projectevaluatie departementsbrede implementatie digitalisering documentopslag*. HEC, 4 april 2002.
- ⋮ Template herstructurering G:\-schijf, FacB: september 2001.
- ⋮ *Thesauri: Inleiding en begrippen*. Haagse Hogeschool: Den Haag

- ⌘ Verslag Intakegesprek project Metadata met Ed de Bruin, Informatiemanager OCW, 13-11-2002
- ⌘ Verslag terugkoppelbijeenkomst oriëntatiefase 'Auteurs-gegenereerde metadata', 13-12-2002
- ⌘ Verslag evaluatiebijeenkomst Pilotdirecties Voorlichting en RZO, 24-04-2003
- ⌘ *Werkprogramma Digitaal Werken bij OCW 2003-2006 (werkprogr. projectenportfolio 2003 vs2.doc)*
- ⌘ *Woordsystemen: Theorie en praktijk van thesauri en trefwoordsystemen.* Magrijn, M. et al. (1997). NBLC Uitgeverij: Den Haag

7.2 Internetbronnen

- ⌘ **'De kunst van het vinden'**
Vooronderzoek naar mogelijkheden tot het verbeteren van de vindbaarheid van (rijks)overheidinformatie. Stichting Itafit, April 2002. <http://www.zenc.nl/publicaties/>
- ⌘ **'Author-generated Dublin Core Metadata for Web Resources'**
A baseline study in an Organisation. Jane Greenberg et al. University of North Carolina at Chapel Hill
[Journal of Digital information, volume 2 issue 2, Themes: Information discovery 2001-11-06](http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Greenberg/)
<http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Greenberg/>
- ⌘ **'Overcoming the Obstacles for META DATA Management'**
Mike Butt. Application Development Trends, june 2000.
<http://www.adtmag.com/article.asp?id=2641>
- ⌘ **'UK metadataframework, may 2001'**
e-Government Metadata Framework - Frameworks - Office of the e-Envoy
[http://www.e-envoy.gov.uk/oeo/oeo.nsf/sections/framework-egmf/\\$file/egmf.htm](http://www.e-envoy.gov.uk/oeo/oeo.nsf/sections/framework-egmf/$file/egmf.htm)
- ⌘ **'Version 2 of the e-Government Metadata Standard'**
UK Govtalk
<http://www.govtalk.gov.uk/>
- ⌘ **'ISO 15836:20003(E) Information and documentation — The Dublin Core metadata element set'**
Dublin Core Metadata Initiative
<http://dublincore.org/en> <http://www.niso.org/international/SC4/sc4docs.html>
- ⌘ **'Metadata'**
Verkennen – metadata en zo
<http://www.mmi.unimaas.nl/vincent/metadata.htm>

- ⌘ **'Regeringsverklaring Kabinet Balkenende: Duidelijkheid en daadkracht, 26 juli 2002'**
Regering.nl – Balkenende II'
<http://www.regering.nl/regeringsbeleid/bronnen/index.jsp>
- ⌘ **Raamwerk voor metadata in de SURFnet Video Portal (DRAFT, Pva+PF+lvO, v0.21, 30/110/00)**
http://www.surfnet.nl/innovatie/surfworks/svp/doc/Metadata_Raamwerk.pdf
- ⌘ **'Archieflexicon voor Nederland en Vlaanderen, concepttekst, januari 2000'**
Teuling, A.J.M. den (2001). Koninklijke Vereniging van Archivarissen in Nederland
www.kvan.nl/nieuws
- ⌘ **'Een veilige omgeving voor digitale archiefbescheiden: een standaard voor record management applications'**
Zwolse groep (1999).
www.gemeentearchief.amsterdam.nl/concerndiensten/hulpmiddelen/digitale_omgeving/-veilige.pdf
- ⌘ **'Div-woordenboek van digital display'**
Digital display
<http://www.ddisplay.nl/>

Dit is een publicatie van het Ministerie van Onderwijs, Cultuur en Wetenschap

Auteur	Wil Rombout
Vormgeving	Wil Hermans, Den Haag
Foto	Remco Vellekoop, Den Haag
Druk	Speed Print, Zoetermeer

ISBN	90 5910 331 9
------	---------------

Uitgave	augustus 2003
Zie ook	www.minocw.nl/organisatie/bronontsloten

OCW 10.871/3.750