

# Multi-Media vanuit bibliotheekperspectief

## Citation for published version (APA):

Franssen, J. (1997). *Multi-Media vanuit bibliotheekperspectief*. Universiteitsbibliotheek Maastricht.

## Document status and date:

Published: 01/01/1997

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

**MULTI-MEDIA**

**vanuit**

**BIBLIOTHEEKPERSPECTIEF**

## **Inhoudsopgave**

0 Verantwoording

### **Deel I De noodzaak tot anticipatie**

1 Inleiding

2 Inventarisatie van het aanwezige beeld-, geluids- en video-materiaal

2.1 Gesprekspartners

2.2 Collectie-grootte

2.2.1 Foto-archief Unigraphic

2.2.2 Vakgroepen en medewerkers

2.2.3 UB-collectie

2.3 Collectie-soorten

2.4 Collectie-gebruik

2.5 Conclusies

3 De noodzaak tot anticipatie

3.1 Het groeiende aanbod elektronische documenten

3.2 De aard van multi-media documenten

3.3 Ontwikkeling van computerondersteund onderwijs (COO)

3.4 Beschikbaarheid van aanwezig beeldmateriaal

3.5 De vraag naar 'on-demand'-diensten

3.5.1 Noodzakelijke apparatuur

3.5.2 De concurrentie-positie

3.5.3 Gegarandeerde beschikbaarheid

### **Deel II Functionele systeemdelen**

1 Multi-media database management systemen; een eerste kenschetsing

2 Document capture

2.1 Digitalisering van documenten

2.2 Document-analyse

2.3 Zelf digitaliseren en uitbesteden

2.3.1 Centrale plek binnen de instelling

2.3.2 Uitbesteden

2.4 Copyright en patientgebonden materiaal

3 Document Opslag

3.1 Opslag-media

3.2 Opslag grootte

3.3 Documentformaten

3.3.1 Het belang van formaten

3.3.2 Tekstdocument-formaten

3.3.2.1 Word, WP en PostScript

3.3.2.2 SGML en HTML

3.3.2.3 TIFF - Tagged Image File Format

3.3.2.4 PDF - Portable Document Format

3.3.2.5 Mogelijke voorkeuren

3.3.3 Beelddocument-formaten

3.3.3.1 Veel gebruikte formaten

- 3.3.3.2 JPEG
- 3.3.3.3 FIF - Fractal Image Format
- 3.3.4 Videodocument-formaten (MPEG)
- 3.3.5 Geluidsdocument-formaten (AIF, AU,SND,MID,WAV,RA)

#### 4 Document transport

- 4.1 Bandbreedte
- 4.2 Protocol
- 4.3 ATM en Fast-Ethernet

#### 5 Document retrieval

- 5.1 Ontsluiting elektronische documenten
- 5.2 Bibliografische ontsluiting
  - 5.2.1 Bibliografische ontsluiting tekst-georiënteerde documenten
    - 5.2.1.1 Pica richtlijnen voor het catalogiseren van online resources
    - 5.2.1.2 Bibliografisch moeilijk identificeerbaar materiaal
  - 5.2.2 Bibliografische ontsluiting fotomateriaal
  - 5.2.3 Bibliografische ontsluiting video-materiaal
- 5.3 Onderwerpsontsluiting
  - 5.3.1 Onderwerpsontsluiting elektronische tekstdocumenten
  - 5.3.2 Onderwerpsontsluiting beelddocumenten
    - 5.3.2.1 Aparte status van beeldmateriaal
    - 5.3.2.2 Thesaurus gebaseerde indexering
    - 5.3.2.3 Content-based-image retrieval
    - 5.3.2.4 Relevante beeld-eigenschappen
    - 5.3.2.5 Trefwoord-indexering versus Content-based-image retrieval
  - 5.3.3 Onderwerpsontsluiting videodocumenten

#### 6 Document Display, Manipulation and Copy

- 6.1 Interface
- 6.2 Multi-media werkstations
- 6.3 Aanverwante apparatuur

### **Deel III Organisatorische consequenties van multi-media**

#### 1 Algemene organisatorische consequenties voor de UB

- 1.1 Sterkere markt-oriëntatie
- 1.2 Samenwerking met rekencentra
- 1.3 Expertise ontwikkeling.

#### 2 Consequenties op acquisitie-niveau

- 2.1 Kwaliteitsbeoordeling en verantwoordelijkheid van aanschaf
- 2.2 Attendering
- 2.3 Samenwerking met inhoudsdeskundigen
- 2.4 Diversificatie van de markt

#### 3 Consequenties voor catalogisering

- 3.1 Het bibliografische werk
- 3.2 Het inhoudelijke werk

#### 4 Consequenties voor ISD (uitleen)

## 5 Consequenties voor het systeem-beheer en -onderhoud

- 5.1 Monitoring
- 5.2 Back-up's
- 5.3 Formaat-, applicatie- en conversie-ondersteuning
- 5.4 Database-ontwerp, -onderhoud en -beheer

## 6 Consequenties voor het bibliotheekonderwijs

## 7 Mogelijke consequenties voor het onderwijs-leerproces

## 8 Mogelijke consequenties voor het onderzoek

## 9 Samenvatting

## **IV Aanbevelingen voor het vervolgtraject**

### 1 Fase I

### 2 Fase II

- 2.1 Maasnet-2
- 2.2 Centrale organisatie en coördinatie

## Literatuurlijst

## 0. Verantwoording

Dit rapport is geschreven in het kader van een UB-project waarin gestreefd werd naar de opstelling van een gedetailleerd projectuitvoeringsplan voor een daadwerkelijke implementatie van een multimedia-database-management-systeem, waarmee tekst-, geluid-, grafisch-, beeld- en video-materiaal op een lokaal netwerk beschikbaar gesteld kan worden.

Tijdens het onderzoek bleek al snel dat kant en klare systemen nog niet bestaan en er nogal wat probleemgebieden zijn, waar geen eenduidige oplossingen voor te geven zijn. Commercieel leverbare systemen zijn zeer domeingericht ontworpen en bieden doorgaans te weinig mogelijkheden voor een adequate bibliografische en inhoudelijke ontsluiting van documenten zoals een universiteitsbibliotheek zich dit zou wensen. Daarnaast bleek dat een inhoudelijke ontsluiting van met name beeldmateriaal, consequenties heeft, die zeer ingrijpend zijn. Omdat automatische beeld-ontsluitings-technieken zich nog in een experimenteel stadium bevinden, zullen bibliotheken, alvorens ze tot een inhoudelijke ontsluiting van beeldmateriaal willen overgaan, zeer goed over de consequenties moeten hebben nagedacht.

Als resultaat van deze bevindingen is het oorspronkelijke projectdoel bijgesteld. Dit rapport bevat dan ook niet het projectuitvoeringsplan, maar alleen de inventarisatie van beschikbaar te stellen materiaal en een beschrijving van functies en probleemgebieden die bij de constructie van een multi-media database-management-systeem een rol spelen. Omdat er over nogal wat punten door bibliotheken beslissingen genomen moeten worden, is er ook aandacht voor de organisatorische consequenties van elektronische beschikbaarstelling. Alhoewel de behandeling globaal blijft, poogt dit rapport toch enig houvast te bieden aan personen die de vernieuwing van de infrastructuur en de functionaliteit ervan mee willen sturen. Omdat grote investeringen beslist enige notie van de technieken vergen, al was het maar voor de herkenning in overlegsituaties met technische deskundigen, wordt er op sommige punten in dit rapport ook op de technische aspecten ingegaan. Omdat we ons in het onderzoek beperkt hebben tot lokale beschikbaarstelling via het universiteitsnetwerk, wijden we niet uit over mogelijke technieken en standaarden voor documentleveranties tot aan de thuiswerkplek.<sup>1</sup>

De inventarisatie van de multi-media wensen aan de instelling is overigens niet uitputtend. Omdat ook hier al snel bleek dat het te beschikbaar stellen materiaal vrij divers is en een generieke database de enige optie zou zijn, is er van verdere interviews afgezien.

Het zal duidelijk zijn dat er aan de kostenoverzichten, zoals ze in het oorspronkelijke projectvoorstel voorzien waren, in dit rapport niet toegekomen kan worden. Over veel factoren bestaat nog onduidelijkheid. Bovendien hangen deze factoren ook nog af van de wijze waarop het database-management-systeem vormgegeven wordt. Kwantificering van kosten in dit stadium is eenvoudig onmogelijk. Pas wanneer alle organisatorische consequenties en technische vereisten bekend zouden zijn, zou een inschatting van de kosten gemaakt kunnen worden. Een eerste aanzet hiertoe vormt de inhoud van dit rapport.

J.T. Franssen  
1996 en 1997.

---

<sup>1</sup>. Hiervoor zij verwezen naar Niemegeers (1995), Tuck (1996) en de resultaten van het UB-project 'A03 Documentlevering/IBL vanuit het landelijk Open Bibliotheeknetwerk (OBN) tot aan de werkplek'.



## Deel I

### 1 Inleiding

Met de ontwikkeling van snelle computer-processoren en het beschikbaar komen van nieuwe netwerk-technologieën staan multi-media-toepassingen sterk in de belangstelling, ook in de onderwijswereld.

De interesse in deze ontwikkelingen vanuit onderwijs-kringen, ligt op een lijn met het belang dat aan de informatie-technologie in zijn algemeen toegekend wordt. Dat belang is gelegen in de vele toepassingsmogelijkheden die met behulp van één interactief medium, de computer, gerealiseerd kunnen worden. De kracht van de huidige informatie-technologie bestaat erin dat er, in principe vanuit de gehele wereld, op één werkstation voorzien kan worden in wat Barker & Yeates (1985, p. 25) nog "interesting possibilities" noemden, namelijk:

- (1) *sound effects and analysis,*
- (2) *static and dynamic imagery through computer graphics,*
- (3) *text handling facilities,*
- (4) *control of external devices and of learning progress,*
- (5) *a variety of data capture techniques,*
- (6) *facilities for data archival, retrieval and dissemination,*
- (7) *a means of achieving highly individualized instruction,*
- (8) *a highly interactive learning environment, and*
- (9) *facilities for pattern matching, computation and decision making.*

Inmiddels zijn een aantal van deze faciliteiten in het hoger onderwijs gerealiseerd. Computers worden door praktisch alle studenten op een bepaald moment in hun studie gebruikt, veelal als instrument voor het produceren van tekst, tabellen en grafieken, en als krachtige rekenmachine. Beheers- en administratieve taken ten behoeve van studenten zijn op veel instellingen geautomatiseerd zodat relevante gegevens, zoals bijvoorbeeld studievoortgangs-gegevens, elektronisch aan studenten ter beschikking gesteld kunnen worden. Universiteitsbibliotheken hebben een on-line catalogus die ook via Internet bereikbaar is en experimenteren met elektronische documentlevering tot aan de werkplek.

Een probleemgebied vormen toepassingen die naast tekstdata ook met andere data, zoals geluid, beeld en video, van doen hebben. Vanuit de onderwijskant kan gedacht worden aan computerondersteunde onderwijsprogramma's (COO's) en aan andere leermiddelen waar de integratie en het gebruik van andersoortige data een wezenlijke rol zouden kunnen spelen. Vanuit het wetenschappelijke onderzoek kan men denken aan de bestudering van (medisch) beeld- en geluidsmateriaal en vanuit de taken van een universiteitsbibliotheek aan de beschikbaarstelling ervan. Louter door de grootte van deze databestanden leveren ze problemen op voor de gebruikelijke netwerken aan onderwijsinstellingen en stellen ze zware eisen aan de processor-snelheden en geheugens van pc's. Gebruikersvriendelijke en betaalbare software die ook grafische afbeeldingen, geluid, foto's en video aankan, is pas sinds kort op grotere schaal beschikbaar en kent dan ook nog geen hoge penetratiegraad in het hoger onderwijs, (zie ook Dieleman & Schenau 1995, p.11).

Met de snelle ontwikkeling in de mogelijkheden van computers en netwerken om meerdere soorten data te verwerken, dringt ook de vraag zich op hoe deze veelsoortige informatie het best opgeslagen en beschikbaar gesteld kan worden. Op deze vraag zal in deel II van dit rapport worden ingegaan, nadat in hoofdstuk 2 het huidige gebruik van beeldmateriaal, en in hoofdstuk 3 de noodzaak tot anticipatie op de technologische ontwikkelingen ter sprake is gekomen. Deel III van het rapport bevat een beschrijving van de organisatorische consequenties voor de UB. Het rapport wordt afgesloten met een aantal aanbevelingen voor een vervolgproject (Deel IV).



## **2 Inventarisatie van het aanwezige beeld-, geluids- en video-materiaal**

### **2.1 Gesprekspartners**

In eerste instantie werd er in het projectvoorstel voorzien in een uitputtende inventarisatie van het aan de instelling aanwezige beeld-, geluids- en video-materiaal. Gaandeweg is deze opzet verlaten. Niet alleen omdat afspraken zeer moeilijk te plannen waren maar vooral omdat al snel duidelijk werd dat, gegeven de aantallen en eisen, de benodigde opslagcapaciteit enorm zou moeten zijn en een algemeen database-ontwerp de enige optie zou zijn. Het ontwerp van het multi-media-database-management-systeem moest zo algemeen zijn dat het niet alleen aan zeer divers materiaal plaats zou bieden maar ook mogelijkheden moest bevatten voor een toegesneden bibliografische en inhoudelijke ontsluiting van het beeldmateriaal op vakgroepniveau. Toen dit duidelijk was, is er van verdere interviews afgezien. Gesproken is dan ook alleen met vakgroepen en diensten waarvan op voorhand duidelijk was dat er veel met beeldmateriaal gewerkt werd of zou worden. Gesproken is met: Visuele Cultuur, Klinische Genetica, Radiodiagnostiek, Skills-Lab, Elektronenmicroscopie, Unigraphic en met de beheerders van een aantal studielandschappen.

### **2.2 Collectie-grootte:**

#### **2.2.1 Foto-archief Unigraphic**

Een apart staande collectie beeldmateriaal wordt gevormd door het fotoarchief van Unigraphic. Deze collectie is hier opgenomen omdat door de gebrekkige archivering van dit snel groeiende en zeer diverse archief, de nog jonge historie van de eigen instelling snel verloren kan gaan. De voor de instelling interessante collectie, voornamelijk personen en gebouwen, bestaat uit een 1500-tal foto-negatieven en dia-positieven.

#### **2.2.2 Vakgroepen en medewerkers:**

Op vakgroepniveau was opmerkelijk dat er slechts enkelen concreet konden aangeven hoeveel beeldmateriaal er voor onderzoek en onderwijs aanwezig is en welk deel van de verzameling op het netwerk beschikbaar gesteld zou moeten worden. Alleen Elektronenmicroscopie en Klinische Genetica konden concrete getallen noemen. Zo beschikt klinische genetica b.v. over een verzameling van 10.000 dia's, 1000 z/w foto's, 25 video's en een tiental audiocassetes. Op de vraag, welk deel publiek beschikbaar gesteld zou moeten worden, kon geen antwoord gegeven worden omdat het gebruik erg privacy-gevoelig ligt. Elektronenmicroscopie daarentegen, met een verzameling van rond de 91.000 foto-negatieven met zeer hoge resoluties, gaf te kennen dat ongeveer 10% van deze collectie voor eindgebruikers raadpleegbaar zou moeten zijn. Dit zou ook moeten gelden voor de jaarlijkse aanwas van 5400 negatieven waarvan ook 10 % digitaal beschikbaar zou moeten komen. Het Skills-lab, dat zowel dia's, foto's, cd-rom's en video's gebruikt, werd alleen over de video's concreet. Hoewel zij slechts beschikken over een 20-tal onderwijsgebonden video's, produceren ze jaarlijks wel een 400-tal, digitaal opgenomen, video's met patient-simulaties die ter evaluatie door docenten en studenten, naar de bibliotheek gebracht worden. Met name voor de docenten zou elektronische raadpleging op de werkplek een uitkomst bieden. Deze moeten nu, veelal vanuit andere instellingen, voor de evaluatie speciaal naar de bibliotheek komen waar ze de video's in de av-groepskamers kunnen bekijken.

Hoewel de andere vakgroepen minder concreet konden aangeven om welke aantallen het gaat, is het toch duidelijk dat de (potentiële) collecties enorm zijn. Hier moet niet alleen Radiognostiek genoemd worden, die grote aantallen digitaal beeldmateriaal genereert, maar ook de onderscheiden collecties die de individuele medewerkers in bezit hebben en, zij het incidenteel, voor onderwijsdoeleinden gebruiken. Het was markant dat in de gesprekken met name privé-collecties van collega's genoemd werden.

### **2.2.3 UB-collectie:**

Een niet onaanzienlijke collectie beeldmateriaal, die specifiek met het ook op het onderwijs beheerd en aangeschaft wordt, vormt de collectie AV-media in de studielandschappen. Wanneer er alleen al naar de video's gekeken wordt dan spreken we over een 180-tal video's voor FdR, 1264 stuks voor FdG, FdGW en Psychologie en over een 250-tal video's voor FdCW en FdEWB. Bovendien worden er voor geneeskunde ook nog eens 280 band-dia-series beheerd die elk uit een tachtigtal dia's bestaan. Gewezen zou hier ook kunnen worden naar de collectie Dittrich, die uit een groot aantal kunsthistorische werken bestaat, met een schat aan beeldmateriaal voor FdCW. Een ander deel van de UB-collectie die voor digitalisering in aanmerking zou kunnen komen, wordt gevormd door onderdelen van de zogenaamde Jezuïeten-collectie. Vanuit cultuurwetenschappelijke hoek (FdCW) wordt er bijvoorbeeld over de digitalisatie van een aantal tijdschriften gedacht die vanwege hun fysieke staat moeilijk meer op traditionele wijze, verantwoord te raadplegen zijn.

### **2.3 Collectie-soorten**

Een onderscheid dat in de gesprekken al snel naar voren kwam, was een onderscheid tussen onderwijsgebonden materiaal en onderzoeksmateriaal. Voor het onderwijsgebonden materiaal was men van mening dat dit in principe voor de studenten en onderwijzers digitaal beschikbaar moest zijn. Voor het onderzoeksmateriaal was van mening dat dit niet direct publiek toegankelijk hoefde te zijn maar men gaf wel te kennen dat een digitale kopie voor eigen gebruik, wel handig zou zijn. In de medische sfeer blijken met name privacy- en beheersproblemen een rol te spelen wanneer naar publieke beschikbaarstelling gevraagd wordt. Zo gaf Klinische Genetica en Radiodiagnostiek te kennen dat de patientgebondenheid van het materiaal een bijna onoverkomelijk probleem vormt als het gebruik buiten de reguliere patientenzorg zou vallen. Hoewel ze het er allemaal over eens zijn dat digitalisering van het beeldmateriaal grote voordelen voor het wetenschappelijk onderzoek zou bieden, wordt de toestemming voor een andersoortig gebruik zelden gevraagd. Patienten zijn vaak zelf niet beslissingsbevoegd en moet men voor de toestemming familieleden benaderen. Dit gebeurt slechts zelden en alleen in uitzonderlijke gevallen. Men wees er wel op dat er voor onderwijsdoeleinden steeds meer commercieel (medisch-)beeldmateriaal, met name cd-rom, op de markt komt dat gemakkelijk via het netwerk beschikbaar gesteld zou kunnen en moeten worden. Nu gebeurt dit voornamelijk middels stand-alone configuraties die, vanwege de licentie-overeenkomsten, alleen in aparte kamers gebruikt kunnen worden. Deze auteursrechtelijke problematiek werd ook door Visuele Cultuur naar voren gebracht die voor het onderwijs zo'n 30 tot 40 uur video beschikbaar zou willen stellen.

Het onderscheid tussen een onderwijs- en een onderzoeks-collectie kwam ook tot uitdrukking in de mogelijke bewerkingen die men voor beeldmateriaal noodzakelijk acht. Aan de eindgebruikers-zijde, c.q. de onderwijzer, denkt men voornamelijk aan louter een kwalitatief goede presentatie van afbeeldingen en video. Alleen in de medische sfeer wordt er nog aan tomografische software gedacht waarmee afbeeldingen geroteerd kunnen worden. Aan de onderzoekskant daarentegen wenst men meer mogelijkheden en zouden afbeeldingen en video ingelezen moeten kunnen worden voor verdere bewerkingen, berekeningen en analyses. Ook de inhoudelijke en bibliografische ontsluiting zou voor het onderzoeksmateriaal dieper moeten gaan dan voor het onderwijsgebonden materiaal. Met name Visuele Cultuur en Elektronenmicroscopie, hadden uitgesproken ideeën over de ontsluiting van beeldmateriaal. Zonder lang nadenken werden een aantal zeer specifieke zoek-criteria en analysetechnieken genoemd die voor een effectief gebruik van een onderzoekscollectie noodzakelijk zouden zijn en veel verder reiken dan de mogelijkheden die de OPC momenteel biedt.

## 2.4 Collectie-gebruik

Het huidige gebruik van de aanwezige beeld- en video-collecties is moeilijk onder één noemer te brengen. Het grootste deel van het aan de instelling beschikbare beeldmateriaal wordt gegenereerd in het kader van de reguliere patiëntenzorg en slechts een klein deel hiervan wordt voor onderwijs en wetenschappelijk onderzoek gebruikt. Ook voor de prive-collecties geldt dat er slechts incidenteel afbeeldingen voor het onderwijs gebruikt worden. Van geheel andere aard is het fotoarchief van Unigraphic waar voornamelijk voor voorlichtingsdoeleinden een beroep op wordt gedaan. Weer van een andere aard zijn de AV-media en cd-rom's die wel specifiek voor het onderwijs aangeschaft en beheerd worden. Maar ook hier geldt dat feitelijke gebruikscijfers moeilijk te geven zijn. Veel van dit materiaal staat namelijk in een open-opstelling waardoor het feitelijk gebruik moeilijk te bepalen valt.

Bekend is, zij het op basis van beperkte gegevens over een korte periode (zie Dahlmans 1996 en Werkgroep AV-media 1996), dat bij rechten veelvuldig gebruik gemaakt wordt van docentvervangende video's (college's) en andere videobanden die een duidelijke relatie met het lopende onderwijs hebben. Bij FdCW en FdEWB is bekend dat er van de 250 titels slechts een aantal vaker bekeken worden. Hoewel het videogebruik bij deze laatste faculteiten over de jaren heen bescheiden is toegenomen, zijn er toch flinke schommelingen in de beijk-frequentie van individuele titels. Het vermoeden bestaat dat de vermelding van de video-titels in de blokboeken op dit gebruik van invloed is. Dezelfde indruk bestaat er ook bij het video-gebruik bij geneeskunde, gezondheidswetenschappen en psychologie waar van de 1264 video-titels een 700-tal in blokboeken zijn opgenomen.

Een door meerdere mensen genoemde waardevolle collectie beeldmateriaal vormen de band-dia-series. Hoewel het gebruik, met uitzondering van oogheelkunde, over de jaren heen afneemt, geeft men toch te kennen dat de series voor digitalisering in aanmerking zouden moeten komen. Het bevat niet alleen uniek beeldmateriaal maar is bovendien vrij kwetsbaar zodat het snel verloren kan gaan. Verwacht wordt dat bij een elektronische beschikbaarstelling het gebruik met sprongen omhoog zal gaan en men verlost zal zijn van de beheersproblemen die het gebruik nu kenmerken. Bovendien blijken de series een goede bron voor de ontwikkeling van computer-ondersteunde onderwijs-programma's te zijn. Kwalitatief goed beeldmateriaal is voor de ontwikkeling van deze programma's een noodzaak. Een tweetal series zijn bijvoorbeeld al met het oog op de ontwikkeling van een COO-programma over het oog door de vakgroep O&O zelf gedigitaliseerd.

## 2.5 Conclusies

Ondanks de inventarisatie niet uitputtend is geweest, kunnen er toch een aantal conclusies getrokken worden. Duidelijk is dat:

- de potentiële collecties die voor digitalisering in aanmerking komen enorm zijn
- de wens tot digitalisering voorkomt uit de archiveringsproblematiek van het analoge materiaal, een betere beschikbaarheid voor grotere groepen en de mogelijkheid tot bewerking van digitaal materiaal
- een toename in het gebruik van beeldmateriaal verwacht wordt, voor zowel het onderwijs als onderzoek, als het eenmaal digitaal beschikbaar is
- een toename in het commercieel beschikbare materiaal verwacht en bespeurd wordt
- de aard van het materiaal zeer divers is en uiteenloopt van een te scannen tekst met plaatjes tot aan video's toe
- door het ontbreken van feitelijke gebruikscijfers, er zeer kritisch naar het belang van het te digitaliseren materiaal gekeken moet worden
- digitalisering een aantal beheersproblemen voor de bibliotheek zou oplossen (zoals bij de band-dia-series en een aantal tijdschriften uit de Jezuïeten-collectie)
- aan de onderwijskant voornamelijk gedacht wordt aan illustratief beeldgebruik, docentvervangende informatieoverdracht en onderwijs-ondersteuning met behulp van video en, bij elektronische beschikbaarheid van beeldmateriaal, aan betere mogelijkheden tot COO-ontwikkeling

- binnen het huidige gebruik door studenten sterk de nadruk ligt op video
- er voornamelijk problemen worden verwacht rond de privacy van patiënten en auteursrechten op commercieel materiaal

Voor het ontwerp van het multi-media-management-systeem betekent het bovenstaande voornamelijk het volgende:

- Gezien de potentiële collecties zal het systeem over een grote opslagcapaciteit moeten beschikken of in ieder geval gemakkelijk uit te breiden zijn
- Het systeem zal een multi-user functionaliteit moeten bezitten
- Omdat documenten in diverse formaten aangeleverd zullen worden, zullen er voor sommige documenten conversies in eigen beheer uitgevoerd moeten worden
- Gezien het diverse gebruik zullen er, voor de toegang tot de database, uitgebreide autorisatie-mogelijkheden aanwezig moeten zijn
- Het systeem zal mogelijkheden moeten bevatten tot een diepe inhoudelijke en bibliografische ontsluiting van zowel beeld- als video-materiaal, en
- op meerdere plaatsen mogelijkheden tot bewerking van het materiaal moeten bevatten.

Wanneer naar het huidige gebruik van audio-visueel materiaal gekeken wordt dan zou al snel de vraag kunnen rijzen of de investeringen die met een multi-media-database-management-systeem gemoeid zijn, wel opwegen tegen de baten. Zoals Vorst & Looijen (1995) opmerken kennen investeringen in IT-infrastructuren bij gebrek aan directe baten een rechtvaardigingsprobleem.

Niettemin zijn er een aantal goede argumenten te geven waarom bibliotheken op de beschikbaarstelling van multi-media dienen te anticiperen en te investeren.

### 3. De noodzaak tot anticipatie

#### 3.1 Het groeiende aanbod elektronische documenten

De noodzaak om, als bibliotheek, te anticiperen op de snelle ontwikkelingen in de mogelijkheden tot digitale informatie-overdracht en -opslag, ligt voornamelijk in de veranderingen in het informatie-aanbod en de onomkeerbaarheid ervan.<sup>2</sup> Er is niet alleen een toename in het aanbod van elektronische documenten te bemerken, maar ook een verschuiving van tekst-georiënteerde naar multi-media documenten vanwege de nieuwe mogelijkheden waarmee beeld en tekst geïntegreerd kunnen worden, (Brown, 1996 en Roes, 1996).

In de bibliotheekwereld wordt over de manier waarop op de toename van het elektronische aanbod gereageerd moet worden, nogal verschillend gedacht.

In het rapport van De Koninklijke Bibliotheek: "Collecties op achterstand...." (1996), wordt betoogd dat het aanbod van elektronische publicaties maar een fractie van het gedrukte aanbod zal blijven, ook bij een onstuimige groei. De winst van de informatie-technologie moet volgens de samenstellers van rapport vooral gezocht worden in de toegevoegde waarde en niet in de vervanging van het gedrukte aanbod. Zij gaan er schijnbaar vanuit dat naast het elektronische document het gedrukte exemplaar zal blijven bestaan en dat dit door een bibliotheek beschikbaar gesteld zou moeten worden. Hoelang deze werkwijze door bibliotheken gehandhaafd kan worden blijft echter de vraag. Zeker gezien het nog steeds groeiende aanbod van wetenschappelijke literatuur, de nog steeds stijgende prijzen en feit dat, nu al, sommige documenten alleen elektronisch aangeboden worden. Deze documenten zouden, in de optiek van het rapport, door de bibliotheek dus zelf geprint en gebonden moeten worden. Een werkwijze die bijna haaks staat op het streven van veel bibliotheken om elektronische documentlevering tot op de werkplek te willen realiseren en waarmee vooruitgelopen wordt op de 'on-demand'-diensten die bibliotheken in het elektronische tijdperk zullen moeten gaan leveren (zie par 3.5).

Andere auteurs zijn daarentegen van mening dat de informatie-technologie het enige middel is dat ons ter beschikking staat om het steeds groeiende aanbod aan wetenschappelijke literatuur het hoofd te bieden, (zie b.v. Rawlins 1992 en 1993, Clark 1995, Odlyzko 1995). In meerdere artikelen wordt getracht, zelfs aan de hand van rekenvoorbeelden, uitgevers ervan te overtuigen dat elektronisch publiceren, zowel financieel als organisatorisch, de beste oplossing voor alle partijen is. Hoewel er over het te volgen economische model nog discussie bestaat ( King & Griffiths 1995, Hayes 1996, Brown 1996) en nog geen eenduidigheid bestaat over de manier waarop de wetenschappelijke communicatie in elektronische vorm het beste georganiseerd kan worden (Lancaster 1995b, Okerson & O'Donnell 1995, Tenopir 1995, Bide 1996, Giles 1996, Peek 1996, Silverman 1996), zijn alle auteurs van mening dat met elektronische publicaties terdege rekening gehouden moet worden en actie-lijnen uitgezet dienen te worden.

---

<sup>2</sup>. Voor algemene en onderwijskundige argumenten waarom het onderwijs zou moeten anticiperen op informatie-technologische ontwikkelingen zij verwezen naar Franssen (1994) en Wolf (1995).

Hoe snel er op deze verwachte toename door bibliotheken gereageerd moet worden, is moeilijk te beantwoorden. Niet alleen de bibliotheken en uitgevers spelen hier een rol, maar ook de eindgebruikers. Deze zullen bereid, en in staat moeten zijn met elektronisch aangeleverde documenten om te gaan. Dit betekent dat het document platform-onafhankelijk aangeleverd moet worden maar ook dat, aan de gebruikerskant, de juiste soft- en hard-ware voorhanden moet zijn waarmee het document te lezen, te printen en zondig te bewerken is. Een situatie die voor interne gebruikers binnen de instelling sneller te regelen is dan voor externe gebruikers en toekomstige thuiswerkstations. Zo lang aan de gebruikerskant niet aan deze voorwaarden voldaan is kan een bibliotheek alleen intern, organisatorisch en financieel voordeel behalen met het elektronisch verkrijgen en opslaan van documenten, (Task Force on Archiving of Digital Information, 1996).<sup>3</sup>

### **3.2 De aard van multi-media documenten**

Een tweede argument voor extra investeringen in informatie-technologieën, en met name in beeldinformatie management-systemen, ligt in de aard van multi-media documenten die moeilijk meer op de traditionele manieren beschikbaar gesteld kunnen worden. Zou men de informatie uit zulke documenten op traditionele wijze willen aanbieden dan zou dit kunnen betekenen dat men als bibliotheek, naast een stuk tekst, tevens foto's, geluidcassettes, videobanden en bijbehorende afspeelapparatuur beschikbaar zou moeten stellen. De beheersproblematiek is dan niet van de lucht en verklaart voor een deel het sporadische gebruik van de traditionele audio-visuele materialen in het onderwijs.

Bij deze beschikbaarstelling speelt niet alleen een praktisch maar ook een onderwijskundig argument. Juist het gebruik van AV-materialen zou de leerstof aanschouwelijker kunnen maken en de kwaliteit van het onderwijs kunnen bevorderen. Of zoals het door Graziadei (1995) formuleert wordt:

*Through the provision of high-quality sound, images, and video, multiple senses in the learner are engaged in the learning process. And when course materials are more engaging, they tend to be more "memorable".*

### **3.3 Ontwikkeling van computerondersteund onderwijs (COO)**

Een vierde reden voor investeringen in beeld-informatie management-systemen vormt de ontwikkeling van computerondersteunde onderwijsprogramma's. Een van de knelpunten waarom dit soort onderwijs niet goed van de grond is kunnen komen, was een gebrek aan mogelijkheden om beeldmateriaal in computerondersteund onderwijs te integreren, (Pilot 1994). Nu wel goede hardware, software en netwerktechnologie beschikbaar zijn, zou men de financieel/organisatorische en onderwijskundige voordelen van computerondersteund onderwijs kunnen gaan gebruiken<sup>4</sup>.

### **3.4 Beschikbaarheid van aanwezig beeldmateriaal**

Een vijfde reden voor de aanschaf van een beeldinformatie management-systeem vormt reeds aanwezig beeldmateriaal. Vaak is dit alleen in "hard-copy"-versies beschikbaar en slecht toegankelijk voor onderzoek en onderwijs. Elektronische beschikbaarheid, gekoppeld aan een goede bibliografische en inhoudelijke ontsluiting, zou hier een uitkomst kunnen bieden. Dit geldt ook voor bijvoorbeeld foto-archieven die niet direct betrekking hebben op onderwijs of onderzoek, maar bijvoorbeeld op de

---

<sup>3</sup>. Een rekenvoorbeeld, het Yale model, dat uitgaat van een vergelijking tussen enerzijds, digitale opslag en geprinte uitvoer, met anderzijds, papieren opslag en gecopieerde uitvoer, is te vinden in het rapport "Preserving Digital Information" van de Task Force on Archiving of Digital Information (1996). Uit dit rapport waar met zeer restrictieve aannames gewerkt wordt, blijken twee dingen. Ten eerste, wanneer in een gedistribueerde omgeving gewerkt wordt waar alleen die documenten via licenties gedownload worden waar veel vraag naar is, blijkt al na zeven jaren elektronische opslag financieel de voordeligste te worden. Ten tweede, dat het financiële voordeel niet in het technologische aspect van digitale opslag ligt maar gerealiseerd wordt door aanpassingen op een organisatorisch vlak zoals bijvoorbeeld in de wijze van acquisitie en elektronische opslag.

<sup>4</sup>. Voor een beschrijving van deze voordelen zij verwezen naar Franssen (1994) en Wolf (1995).

instellingsgeschiedenis. Van dit soort archieven is de inhoud en wijze van archivering vaak alleen bekend aan de beherende persoon.

Algemeen gesproken biedt een geautomatiseerde opslag van beeldmateriaal de volgende voordelen, (Beek, 1996):

- *het behoud van de collectie: het fysiek hanteren van de originelen wordt beperkt tot en de collectie kan in een aparte ruimte, met ideale bewaaromstandigheden, geconserveerd worden.*
- *de zoeksnelheid wordt vergroot en bedraagt - afhankelijk van het opslagmedium, de apparatuur, kwaliteit en compressie van de afbeeldingen - meestal enkele tot tientallen seconden.*
- *verkeerd terugzetten is niet meer mogelijk*
- *via "thumbnails" is browsen mogelijk*
- *de distributie van documenten wordt vergemakkelijkt*

### **3.5 De vraag naar 'on-demand'-diensten**

#### **3.5.1 Noodzakelijke apparatuur**

Een laatste reden tot anticipatie op, en aanschaf van, een multimedia-database-management-systeem vormen de consequenties en de reikwijdte van de vraag naar 'on-demand'-diensten.

Een van de belangrijke punten in de literatuur over elektronisch publiceren is dat, naast het feit dat bibliotheken op de diensten afgerekend zullen gaan worden (Leskien 1996), de 'on-demand'-diensten zeer ver kunnen gaan en er voor moeten zorgen dat de gebruiker krijgt wat hij vraagt, op een wijze die hem het meeste schikt. Dit noodzaakt bibliotheken niet alleen tot een grondige reorganisatie van de bestaande werkwijzen, maar ook tot de aanschaf en beheer van allerlei extra (rand-)apparatuur en software waarmee naast extern, ook (bibliografisch en praktisch) moeilijk toegankelijk eigen materiaal beschikbaar gesteld zou kunnen worden, (Melin 1996, Corral 1995, Vries & Minnigh 1995, White 1995, Richard 1995, Sever & Harel 1995 en Sever 1995).

#### **3.5.2 De concurrentie-positie**

Een andere consequentie van de vraag naar 'on-demand'-diensten is dat bibliotheken tot samenwerking (resource sharing) gedwongen worden met instellingen die hun concurrenten zouden kunnen zijn, (Leskien 1996, Dannelly 1995). Dit kan leiden tot paradoxale situaties waar de verantwoordelijkheden niet altijd even duidelijk zijn maar waar wel door iemand het initiatief genomen zal (moeten) worden. Dannelly (1995) wijst er terecht op dat voor 'resource-sharing' er ook een te delen collectie moet zijn, maar dat dit niet betekend dat er al op voorhand duidelijk moet zijn wie, wat en waar elektronisch beschikbaar zal gaan stellen en tegen welke kosten. Een afwachtende houding zal ten koste gaan van de noodzakelijke opbouw van expertise en de concurrentiepositie alleen maar kunnen verzwakken.

#### **3.5.3 Gegarandeerde beschikbaarheid**

Een derde consequentie van de 'on-demand'-diensten en reden tot anticipatie is dat bibliotheken ook nieuwe archiverings-rollen toebedeeld zullen krijgen. Ook deze, in zekere zin, paradoxale ontwikkeling wordt in meerdere publicaties onder de aandacht gebracht. Hierin wordt gesteld dat met een toename in het elektronische aanbod, bibliotheken verantwoordelijkheden zullen krijgen die voorheen bij de uitgevers lagen en gewaarborgd werden door het papieren medium waarmee de informatie in boeken en tijdschriften verspreid werd. De verantwoordelijkheid waar het hier om gaat betreft de zogenaamde 'integriteit' van het document. Deze kan bij een elektronisch document veel sneller verloren gaan dan bij een papieren exemplaar waarvan de inhoud en vorm na publicatie zowat eeuwig vastligt. Bij elektronische documenten ligt dit wezenlijk anders. Deze kunnen niet alleen vrij snel van vorm en inhoud veranderen, maar ook de vindplaats en de noodzakelijk software voor de raadpleging kunnen binnen enkele seconden wijzigen. Wil men een voor het wetenschappelijk bedrijf bruikbaar en citeerbaar document overhouden, dan zal iemand de authenticiteit en beschikbaarheid moeten garanderen (zie b.v. Brichford & Mahler 1995, Guenter 1996, Luksch & Schultheiss 1996, Lynch 1996 en Task Force on Digital Archiving 1996).

Men zou, zoals Rawlins (1992 en 1993), kunnen overwegen om deze verantwoordelijk bij de uitgevers te

leggen. Deze zouden voor de unieke beschikbaarheid verantwoordelijk blijven en met een redelijke download-prijs verzekerd zijn van voldoende winsten. Maar of de uitgevers deze verantwoordelijkheid ook nemen en kunnen nemen, valt te bezien.

In de eerste plaats is het onduidelijk of ze nog steeds evenveel wetenschappelijk interessant werk aangeboden zullen krijgen. Met een toename van, door collega-wetenschappers, elektronisch gerefereerde tijdschriften zal ook meer werk op eigen initiatief verspreid gaan worden (Harnad 1996, Hurtado 1996, Roes 1996).

Ten tweede is het bestand dat de uitgevers beschikbaar zullen moeten houden vrij groot en onduidelijk welk deel hiervan regelmatig gedownload zal gaan worden. Als profit-organisatie zullen ze snel geneigd zijn om minder frequent geraadpleegde werken van de server te halen. In principe gebeurt hetzelfde nu al bij de papieren 'backsets', waar voor een copy vaak naar het Amerikaanse UMI verwezen wordt waar tijdschriften op microfiche bewaard worden. Kortom, de archief-functie van non-profit organisaties, zoals bibliotheken, zal met een toename in het elektronisch aanbod eerder toe- dan afnemen.

Een derde en laatste probleem bij de beschikbaarheid van documenten wordt bij de uitgevers gevormd door beperkte bandbreedte van het Internet. Voordat de beperkte bandbreedte van het Internet een probleem werd gingen er stemmen op dat de rol van bibliotheken bij de beschikbaarstelling van materialen, geminimaliseerd zou worden. Hooguit zou er nog van een bemiddelende of verwijzende rol gesproken kunnen worden. Bij een alsmat toenemend elektronisch aanbod, waar gebruikers via het Internet zelf gemakkelijk aan zouden kunnen komen, kan een bibliotheek alleen nog maar een verwijzende functie vervullen. Ze hoeft alleen nog maar duidelijk te maken waar het een en ander en tegen welke prijs, te verkrijgen is. Alleen uniek materiaal, dat nergens anders verkrijgbaar zou zijn, zou dan in aanmerking komen voor digitalisatie en elektronische beschikbaarstelling door de eigen instelling. Een veel gehoorde kreet in deze was en is nog vaak: "Van collectie-management naar informatie-management". Een probleem is echter dat deze kreet alleen maar geldt voor materiaal dat inderdaad via externe bronnen door het netwerk aangeleverd kan worden. Wanneer deze externe bronnen moeilijk te bereiken zijn, de bestanden relatief groot zijn en het Internet zijn huidige bandbreedte blijft behouden, dan zal van de veel gevraagde "on-demand"-diensten, van externe leveranciers weinig overblijven. Een gevolg hiervan is dat ook de collectie-management-taken van bibliotheken weer sterker op de voorgrond zullen gaan treden (Leskien 1996, Melin 1996) en zeker wanneer deze de beschikbaarheid en snelle levering over een lokaal netwerk kunnen garanderen.

Het zal duidelijk zijn dat wanneer bibliotheken op deze, in wezen traditionele, taken aangesproken zullen gaan worden en 'on-demand'-diensten zullen moeten gaan leveren, er niet alleen voldoende expertise aanwezig moet zijn maar ook de noodzakelijk apparatuur waaronder een multi-media-database-management systeem.



## Deel II Functionele systeemdelen

### 1 Multi-media-database-management-systemen; een eerste kenschetsing

Wanneer in de literatuur over multi-media database management-systemen (MMDBMS) gesproken wordt dan spitst het zich al snel toe op beeld-informatie managementsystemen omdat deze, qua techniek en ontwerp, de meeste problemen opleveren. In de literatuur worden deze systemen vaak afgezet tegen traditionele database-management systemen (DBMS) die al vrij lang in gebruik zijn. Omdat deze systemen primair ontworpen zijn voor alfa-numerieke informatie, zoals teksten, is de opslag, het beheer en de beschikbaarstelling van dit soort informatie tot het uiterste geoptimaliseerd. Dit i.t.t. niet alfa-numerieke informatie zoals beelden, die louter door hun grootte en (potentiële) informatie-inhoud in traditionele database management-systemen voor problemen kunnen zorgen. Beeldmateriaal wordt hierin meestal opgeslagen als een groot binair object, ook wel BLOB genoemd, en de ontsluiting geregeld via trefwoord-indexen. Voor sommige toepassingen mag dit voldoende zijn maar, voor bijvoorbeeld video-materiaal en voor symbolisch rijk beeldmateriaal, is deze benadering minder bevredigend. Een inhoudelijke diepte-ontsluiting met behulp van trefwoorden is bij dit soort beeldmateriaal enorm tijdrovend en de wijze van opslag laat ook geen benadering van delen van een video of foto toe.

Een voor beeldmateriaal geoptimaliseerd systeem zou naast de bekende database-functies en hogere processorsnelheden, moeten beschikken over zowel automatische segmentatie-technieken voor video alsook automatische beeldinformatie-extractie-technieken. Segmentatie-technieken voor bijvoorbeeld de herkenning van video-shots en scenes, en extractie-technieken voor de automatische generering van beeld metadata (kleur, vorm, grootte, textuur etc.) die voor een additionele inhoudelijke ontsluiting kunnen zorgen. Wanneer een systeem ook deze eigenschappen bezit dan spreekt men over een multi-media database management-systeem (MMDBMS).

Kenmerkend voor de huidige beeld- en video- database-systemen is dat ze ontworpen zijn voor specifieke toepassingsgebieden (Raghavan, V. & Gudivada, V., 1994, Furht & Smoliar & Zhang 1995 en Ahanger, G. & Little, T., 1996). Zo zijn er systemen voor "video-on-demand", "teleconferencing" en diverse andere soorten beeld-archieven maar nog geen generieke beeld- en video-database-management-systemen. Systemen die de technieken voor de behandeling van alle soorten materiaal (tekst, grafiek, beeld, geluid en video) geïntegreerd hebben vindt men helemaal niet in de literatuur terug. Zo'n systeem zal zelf ontworpen en/of samengesteld moeten worden.

De delen en functies waaruit een multi-media managementsysteem zou kunnen/moeten bestaan, wordt in het navolgende beschreven aan de hand van wat men, in navolging van de "Image Chain", de "Document Chain" zou kunnen noemen. Onderdelen van de "Document Chain" die achtereenvolgens ter sprake zullen komen, zijn:

- Document Capture (vastleggen)
- Document Storage (opslag)
- Document Transport (netwerk transport)
- Document Retrieval (ontsluiting)
- Document Display (weergave)
- Document Manipulation (bewerking)
- Document Copy (hard copy)

## 2 Document capture

### 2.1 Digitalisering van documenten

Documenten kunnen kant en klaar in een bepaald elektronisch formaat aangeleverd worden of documenten zullen zelf gedigitaliseerd moeten worden. Bij kant en klare aanlevering zou alleen nog een formaat-conversie moeten plaatsvinden wanneer het formaat niet met het gewenste opslag-, transmissie- of print-formaat overeen zou komen. Wanneer er zelf gedigitaliseerd moet worden dan zijn er meer stappen te nemen, afhankelijk van het gewenste eindresultaat. Soms volstaat een "digitale foto" van een origineel terwijl er in andere toepassingen ook nog sprake is van de (semi-)automatische omzetting van de inhoud (letters, cijfers en pagina-opmaak) van een document, in computercodes. Dit gebeurt met behulp van zogenaamde Optical Character Recognition (OCR)-software die een machineleesbaar tekstdocument als einddoel heeft.

Een tweetal veel gebruikte methodes voor digitalisatie zijn scannen en het opnemen met een videocamera. Scanning heeft, volgens van der Beek (1996), de achterstand op de videocamera wat betreft de beeldkwaliteit inmiddels ingehaald. Volgens hem zijn er zeer mooie resultaten mee te bereiken en is het meest gangbare type een flatbed-scanner, al dan niet uitgerust met een documentdoorvoersysteem.<sup>5</sup> Het gebruik van een (video-)camera heeft het voordeel dat het beeld direct (in zijn geheel) op het beeldscherm van de computer te zien is, waarbij software-matig instellingen als scherpte en belichting kunnen worden bijgesteld. Digitale camera's worden nog niet zoveel gebruikt vanwege de hoge prijzen en de hardware eisen. Gebruikelijker zijn analoge video-camera's met een speciale conversiekaart (A/D converter) die de analoge signalen omzet in digitale gegevens en ook gebruikt kan worden voor de digitalisering van video-banden.

### 2.2 Document-analyse

Een belangrijke fase bij het digitaliseren van documenten is wat Horik (1996) de 'document-analyse' noemt en waarin naar voren gebracht wordt welke eisen aan de hardware, software en aan het bedienend personeel gesteld moeten worden.

Deze fase is belangrijk omdat het document met een bepaalde vorm en omvang, niet alleen bepaalt welke type scanner of (video-)camera gebruikt kan worden, maar met zijn inhoud en gewenst gebruik, ook de optimale resolutie. Een verkeerde inschatting van de relevante details, het feitelijk gebruik en/of andere inhoudelijke aspecten van het te digitaliseren document, beïnvloedt niet alleen de waardering van het gehele systeem maar leidt er ook toe dat het origineel in een later stadium opnieuw gedigitaliseerd moet worden. Dit is een zeer arbeidsintensief proces dat ten alle tijden voorkomen moet worden (Horik 1996). Zo zijn bijvoorbeeld bepaalde medische afbeeldingen alleen bij hoge beeld-resoluties goed bruikbaar en veel tekstdocumenten alleen wanneer ze machineleesbaar gemaakt zijn. Voor kunsthistorische afbeeldingen zijn bijvoorbeeld weer kleur en zichtbaarheid van bepaalde details belangrijker.

Voor het bedienend personeel betekent dit dat ze over een stuk vakinhoudelijke kennis moeten beschikken of in ieder geval een inhoudsdeskundige moeten kunnen raadplegen over de relevantie van bepaalde details. Bij gebruik van OCR-software geldt bovendien de eis dat het bedienend personeel de tekst moet kunnen lezen, ook al is er sprake van perfect gedrukte cyrillische of gothische letters.

De belangrijkste factoren waarover in het conversieproces een beslissing genomen moet worden, zijn:

---

<sup>5</sup>. Voor zwaar en gebonden materiaal kan een overhead scanner gebruikt worden en voor dia's of microfiches respectievelijk dia- en microfiche-scanners.

- de instellingen van de hard- en soft-ware waarbij de beeld-resolutie, het aantal kleuren en het beeldcontrast de belangrijkste opties zijn<sup>6</sup>,
- voorbewerkingen, zoals eventuele formaat-omzettingen en tekst-aanpassingen, zoals het rechtzetten van schuin gedrukte tekst en vergroting van het beeldcontrast bij OCR,
- nabewerkingen op de OCR-uitvoer zoals het corrigeren van: vergeten tekens (rejection errors), van gemiste tekens (insertion errors), of van verkeerd gelezen tekens (substitution errors).

## **2.3 Zelf digitalisering en uitbesteden**

### **2.3.1 Centrale plek binnen de instelling**

Een thema waar in de literatuur iedere keer weer op gewezen wordt, is dat het hele conversieproces zeer arbeidsintensief is en er nogal wat expertise en organisatie noodzakelijk is om het proces tot een goed einde te brengen (Beek 1996, Horik 1996, SURFdoc 1995). Met name de tijd die nodig is voor het digitaliseren van het materiaal en corrigeren daarvan blijkt nogal eens onderschat te worden. Uit de projecten die Horik (1996) noemt, blijkt ook dat niet alleen een brede kennis van hardware en software vereist is, maar ook kennis van de infrastructuur van de organisatie waarin het conversiesysteem geplaatst wordt. Digitalisering van documenten op een centrale plek binnen de eigen instelling ligt dan ook voor hand. Expertise, onderhoud van soft- en hard-ware, organisatie en databasebeheer kunnen zo op één plek gehouden worden. Problemen kunnen echter ontstaan wanneer de communicatie tussen diverse belanghebbenden, zoals aanbieders en bewerkers van het materiaal, niet goed geregeld wordt. Dit blijkt uit de SURFdoc projecten waar het beeld-transport naar een scanlaboratorium via een netwerk geregeld was. De organisatie van een centrale dienst vergt een grondige voorbereiding met duidelijke afspraken en communicatielijnen met alle belanghebbenden.

### **2.3.2 Uitbesteden**

Uitbesteding van digitalisatie komt in aanmerking wanneer:

- de noodzakelijke inhoudelijke expertise niet in eigen huis aanwezig zou zijn, (denk aan oude drukken en specifieke vakgebieden), of
- de technische expertise niet aanwezig zou zijn omdat er speciale apparatuur gebruikt moet worden, en wanneer
- het materiaal betreft dat zich gemakkelijk leent voor een 'bulk'-behandeling zoals bijvoorbeeld dia-series.

Met name bij deze laatste optie is gebleken dat 'bulk'-scanning door een ander partij (bijv. Kodak) veelal goedkoper is en het resultaat met routinematige kleurcorrecties aangepast kan worden. Het zal van de document-analyse afhangen of digitalisatie in eigen beheer of digitalisatie door een andere partij in aanmerking komt.

---

<sup>6</sup>. Algemeen gesproken moet de scanresolutie zo hoog mogelijk gekozen worden. Voor transport is het niet langer noodzakelijk om lage resoluties te kiezen omdat het verzenden schaalbaar gemaakt kan worden. Bovendien zijn de compressietechnieken nog volop in ontwikkeling en is geheugen niet meer zo duur als een paar jaar geleden. Daarnaast zullen aan de presentatie van afbeeldingen steeds hogere eisen gesteld worden en is het raadzaam om daar nu al op in te spelen. Door nu al te kiezen voor een hoge resolutie en een groot aantal kleuren, kan voorkomen worden dat in later stadium, wanneer alsnog moet worden gekozen voor een betere kwaliteit, alles opnieuw moet worden ingescand en gecorrigeerd. Het zal duidelijk zijn dat in de praktijk grenzen aan de mogelijke resolutie gesteld zijn omdat er zowel financiële, praktische alsook systeem- en netwerktechnische beperking zullen zijn.

## 2.4 Copyright en patientgebonden materiaal

Een steeds terugkerend probleem bij digitalisatie van materiaal is het copyright dat op verschillende soorten materiaal kan rusten. Ook bij de SURFdoc-projecten was dit een steeds terugkomend probleem, (Heijne & Rexwinkel, 1995). Dit feit dwingt bibliotheken nu alleen dát materiaal elektronisch beschikbaar te stellen waarvan het copyright bij de universiteit berust of gekocht kan worden van de rechthebbende. Een ander punt vormt de patient-gebondenheid van veel bio-medisch beeld- en geluids-materiaal dat niet zonder meer aan derden en/of via een netwerk beschikbaar gesteld mag worden. Ook hier geldt dat er een stuk organisatie geregeld moet worden die:

- uitzoekt of op bepaald materiaal rechten rusten en/of tot patienten herleid kan worden
- regels opstelt hoe te handelen bij het aanbod van patientgebonden of copyright gebonden materiaal
- zorgt voor de naleving en controle op de regels
- zondig zorg draagt voor de aankoop van rechten
- eventueel de toestemming verzorgd tot gebruik van patientgebonden materiaal, en
- voor commercieel aangeboden materiaal gebruikslicenties afsluit met aanbieders en afnemers

## 3. Document Opslag

### 3.1 Opslag-media

Er is een groot aantal, al dan niet gestandaardiseerde, media beschikbaar voor de opslag van gedigitaliseerd materiaal. De voornaamste digitale opslagmedia zijn de CD-ROM, WORM (Write Once Read Many-disks), Photo-CD, MOD (Magnetic Optical Disk) en harddisk.

Een belangrijk aspect is of het medium te beschrijven, te herschrijven, of alleen te lezen is. Het zal van de aard van de toepassing afhangen welke vorm de voorkeur verdient.

Een ander aspect vormt de standaardisatie. CD-ROM, MOD en Photo-CD zijn goed gestandaardiseerd. Dit betekent dat de structuur, opbouw etc. van het medium is vastgelegd. Hierdoor kan uitwisseling van gegevens eenvoudig gerealiseerd worden. WORM is minder goed gestandaardiseerd en zal uitgelezen moeten door hetzelfde apparaat dat de schrijf beschreven heeft. Dit hoeft echter geen probleem te zijn als het een gesloten systeem betreft, waarbij alle informatie binnen het systeem gebruikt zal gaan worden.

Een derde variabele die bij de keuze van een opslagmedium speelt is de data-beschikbaarheid; de snelheid waarmee de gegevens uitgelezen kunnen worden. Langzame cd-rom's en hard-disks kunnen de databeschikbaarheid reduceren, ook al is de server snel genoeg. Voor een acceptabele video-kwaliteit moet bijvoorbeeld minimaal 1.5 Mbit/sec. geleverd kunnen worden. Is de server te langzaam (b.v. bij teveel simultane aanvragen) of de harddisk, dan gaat dit bij video het eerst ten koste van het aantal beelden per seconde dat weergegeven wordt en in het ergste geval ten koste van de geluidswaergave waar de gebruiker het meest gevoelig voor is.

Een vierde variabele vormt de frequentie van het gebruik. Worden bepaalde documenten zelden gebruikt dan kan men overwegen ze van een harddisk naar een ander, zgn. tertiair, opslagmedium over te zetten dat langzamer is. De gebruiker die dit document wil inzien krijgt dan in de meeste gevallen een melding dat het document opgehaald wordt en een indicatie van de tijd die dit vergt.

In de literatuur worden meerdere servers voor on-line-opslag genoemd. Namen die men het meest tegenkomt, zijn:

- de n-CUBE-Based Oracle Media Server

- de Silicon Graphics Media Server, en
- de Hewlett-Packard Media Server.

De verschillen tussen deze merken zijn echter niet altijd even duidelijk en hangen voornamelijk samen met de bijhorende software tot de inrichting van de servers. Belangrijker is echter dat de snelheid waarmee de data van de harde schijven gelezen kunnen worden en dat de snelheid waarmee de data op het netwerk gezet kunnen worden, groter moeten zijn dan de 'real-time rates' van applicaties, zodat meerdere simultane transmissies kunnen plaatsvinden van dezelfde informatie.

### 3.2 Opslag grootte

Om een indruk te krijgen hoe groot de geheugen capaciteit moet zijn, volgt hieronder een overzicht met een aantal gemiddelde file-groottes en bit-rates voor bewegend beeld.

	OBJECT TYPE	SIZE AND BANDWIDTH
TEXT	- ASCII - EBCDIC	<b>2 KB</b> per page
IMAGE	- Bit-mapped graphics - Still photos - Faxes	- Simple (grayscale) <b>77 KB</b> per image (320x240x8bits)  - Detailed (color) <b>3 MB</b> per image (1100x900x24bits)
AUDIO	Non-coded stream of digitized audio or voice	Voice/Phone 8KHz/8 bits (mono) <b>6-44 KB/s</b>  Audio CD 44.1 KHz/16 bits (stereo) <b>176 KB/s</b> (44.1 KHz x 2 ch x 16 bits)
ANIMATION	Synched image and audio stream at 15-19 frames/s	16 bit color, 16 frames/sec  <b>6.5 MB/s</b> (320 x 640 x 16 bits x 16 frames/sec)
VIDEO	TV analog or digital image with synched streams at 24-30 frames/s	24 bit color, 30 frames/sec  <b>27.6 MB/s</b> (640 x 480 x 24 bits x 30 frames/sec)

uit: Furht e.a (1994)

Uit dit overzicht blijkt al snel dat de geheugencapaciteit behoorlijk moet zijn. Wanneer alleen al naar de banddia-series (280 stuks met elk 80 dia's) gekeken wordt dat spreken we bij een gemiddelde van 3 MB per beeld al over 67200 MB opslagcapaciteit. Zonder compressie zal hier niet gemakkelijk mee om te gaan zijn. Niet alleen vanwege de loutere bestandsgrootte maar ook vanwege de snelheid waarmee zo'n ongecomprimeerd beeld uitgelezen zou kunnen worden.

Omdat de compressie-ratio's met de documentformaten gegeven zijn, volgt in het navolgende een bespreking van veel gebruikte documentformaten.

### **3.3 Documentformaten**

#### **3.3.1 Het belang van formaten**

Documentformaten zijn om een aantal redenen van belang. Ze bepalen niet alleen de bestandsgrootte en het mogelijk gebruik van relatief langzame opslag-media, maar ook:

- de transmissie snelheid via een netwerk
- de eventuele migratie snelheid van het oudere formaat naar een nieuw
- de keuze van hard- en soft-ware en het aantal noodzakelijke plug-in's en dus uiteindelijk,
- de waardering van de gebruiker van het gehele systeem.

Het zal duidelijk zijn dat een keuze van het opslag formaat ook financiële consequenties zal hebben. Zeker in een netwerk- omgeving waar voor sommige produkten licentiekosten betaald zullen moeten worden en de geheugencapaciteit aanzienlijk zal moeten zijn, (zie ook het rapport van de Task Force on Digital Archiving). Achtereenvolgens zullen de volgende veel gebruikte formaten besproken worden:

- Tekstdocument-formaten ( Word, WP, Postscript, HTML, SGML en de bestandsformaten TIFF en PDF)
- Beelddocument formaten ( GIF, PCD, JPEG, BMP, PIC, TARGA, PNG, en FIF)
- Videodocument formaten (MPEG)

#### **3.3.2 Tekstdocument-formaten**

##### **3.3.2.1 Word, WP en PostScript**

In SURFdoc, een onderzoeksproject van de Stichting SURF naar opslag, inzage en verwerking van elektronische documenten gedurende de jaren 1992 tot en met 1994, werd om praktische redenen bij voorbaat besloten dat teksten opgeslagen zouden moeten worden in ASCII, in tekstverwerkingsformaten zoals WP en Word of PostScript. Dit waren de populaire formaten waarbij weinig technische problemen leken te bestaan. Maar de verschillende formaten bleken alle hun beperkingen te hebben. Zo voorzag ASCII niet in integratie van tekst en afbeeldingen en gaven WP en MS-Word conversieproblemen. Postscript bood weliswaar integratie van tekst en beeld, maar ontbeerde een aantal viewers om die documenten te raadplegen, (Dempsey & Heijne 1995, p.180-181). De software voor het bekijken ('viewen') van een PostScript bestand werkt niet voor alle platforms even gemakkelijk en goed. Het bestand kan wel geprint maar niet op beeldscherm bekeken worden. Zelfs met speciale plug-in's voor Windows, zijn PS-bestanden vaak niet goed te raadplegen op het beeldscherm.

##### **3.3.2.2 SGML en HTML**

In een vervolgproject, Premium genaamd, heeft men alternatieve documentformaten onderzocht die gedurende het SURFdoc-project voorhanden kwamen: SGML (Standardized General Markup Language) en HTML (Hypertext Markup Language).

De ISO-standaard SGML bestaat sinds 1986 en is gericht op het benoemen van inhoudelijke elementen in een tekst door middel van markeringen die 'tags' genoemd worden. SGML is niet gebonden aan bepaalde hardware of software, maar is algemeen inzetbaar. Het is gebaseerd op ASCII, zodat teksten eenvoudig over netwerken kunnen worden verstuurd, en definieert de structuur van een document, niet de uiterlijke verschijningsvorm. Het is een standaarddefinitie van de wijze waarop de tags (coderingen) voor de structurele elementen van een tekst, zoals titel, auteur en noten, moeten worden geconstrueerd. De gecodeerde tekst kan dan in principe worden gebruikt door elk softwarepakket dat de tags kan ontcijferen, en zo op verschillende manieren worden verwerkt: voor het afdrukken van een opgemaakt document, voor schermweergave, voor zoek- en opvraagacties, voor het samenstellen van een bibliografie, enzovoort. Het grote voordeel van SGML is dat wanneer codering van de structuur van een document plaatsvindt met een volledige toepassing van SGML, een maximale flexibiliteit voor toekomstig gebruik van het document in andere applicaties garandeerd wordt. Wanneer in 1996 de ISO 10646 (Unicode) wordt toegepast voor

het weergeven en transporteren van niet-westerse geschriften is de portabiliteit maximaal.

Voor een toepassing binnen een bepaalde omgeving kan uit de totale set van beschikbare tags binnen SGML een keuze gemaakt worden en een soort subset gedefinieerd worden. Dit wordt een DTD (Document Type definition) genoemd. HTML is zo'n DTD, dat in het World Wide Web gebruikt wordt en binnen de academische wereld vrij populair is. HTML dankt zijn populariteit aan:

- het feit dat het aantrekkelijke en relatief eenvoudige mogelijkheden biedt voor het integreren van tekst en plaatjes aangevuld met mogelijkheden om geluid en beeld weer te geven, ingebed in een hypertext omgeving
- de producten via Public Domain beschikbaar zijn
- er veel soorten documenten verwerkt kunnen worden omdat HTML slechts een klein deel van de structuur van documenten vastlegt, en
- er veel aandacht voor de presentatie van een document is.

Een nadeel is dat HTML niet gestandaardiseerd is en er internationaal (nog) geen overeenstemming bestaat over eenduidige toepassing van de DTD, wat zichtbaar is in de verschillende producten voor Web-servers en -clients. Voor een verdergaande standaardisatie wordt er gewerkt aan een versie van HTML die meer naar SGML neigt en HTML+ of HTML 3.0 genoemd wordt. Maar ondanks dat blijft HTML een produkt met een beperkter SGML-structuur dan volledige SGML-documenten en is er internationaal geen overeenstemming over eenduidige toepassing van HTML, (SURFnet BV 1996).

Een ander, niet onbelangrijk, nadeel van zowel SGML en HTML is dat de uitvoer van ontwikkelde software voor de conversie van tekstverwerkerbestanden naar zowel HTML als SGML, handmatig moet worden bijgestuurd en gecorrigeerd. Een activiteit die bijna niet opweegt tegen het werk dat iemand zou hebben, wanneer hij het oorspronkelijk document meteen met de noodzakelijke stuurtekens voor HTML of SGML zou opmaken. Een blijvend probleem zal dit zijn voor bestanden die de noodzakelijke stuurtekens voor een correcte automatische conversie naar SGML of HTML missen. Bij dit soort bestanden, en dat zijn er velen, zal altijd handmatige convertering moeten plaatsvinden.

Een derde nadeel van HTML en SGML, met name voor de levering van documenten aan eindgebruikers, is dat deze formaten niet voldoende garanties bieden dat de integriteit en opmaak van het document, zoals door de producent/uitgever aangeleverd, behouden blijven. Dit heeft met name uitgevers doen omzien naar formaten die na levering niet meer gewijzigd kunnen worden. Ze bieden gescande pagina's aan in een formaat zoals TIFF en PDF.

### **3.3.2.3 TIFF - Tagged Image File Format**

Het TIFF-formaat, dat voor opslag en uitwisseling van beeldmateriaal in kleur en grijstinten ontwikkeld is, en dat het bruikbaar maakt voor meer soorten data dan alleen tekst-images, is op vrijwel ieder platform te gebruiken en wordt voor het interbibliotheekverkeer veel toegepast in samenhang met Group 4-faxcompressie. In feite is het een de facto standaard, zowel nationaal als internationaal, (zie Heijne & Rexwinkel, 1995; Tuck, 1996 en Dempsey e.a. 1996). Het enige nadeel van dit formaat dat in het onderzoek voor dit rapport naar voren gekomen is, naast de onveranderlijkheid van de tekst, de bestandsgrootte. TIFF bestanden zijn bijgevolg langzaam te printen. Een compacter bestandsformaat (vaak 1 op 10) is PDF.

### **3.3.2.4 PDF -Portable Document Format**

Het produkt waarin PDF (de platform onafhankelijke opvolger van Postscript) is opgenomen heet Acrobat en biedt modules voor opslag en annotatie van een document, voor het uitwisselen van een document en voor het bekijken van een document op het scherm. Opslag kan gedaan worden bijvoorbeeld door conversie van een tekstverwerkingsdocument in PostScript (en dit kan vanuit vrijwel alle tekstverwerkers) en vervolgens in PDF. In een op deze wijze opgenomen document kan alleen nog 'geannoteerd' worden, d.w.z. dat de tekst niet veranderd kan worden maar er wel aantekeningen bij de tekst gemaakt kunnen worden. Waar de aantekening thuishoort kan visueel worden gemaakt in het document. PDF biedt geen mogelijkheid om ook de documentstructuur vast te leggen maar kan, volgens SURFnet BV (1996), wel uitstekend worden gebruikt als outputmedium voor documenten die bijvoorbeeld met HTML of SGML zijn opgeslagen.

Naast het feit dat een tekst niet meer veranderd kan worden in het PDF-formaat, hetgeen aan de

gebruikerszijde een bezwaar zou kunnen zijn, is er ook nog een ander nadeel. Steeds weer wordt erop gewezen dat PDF geen officiële standaard is maar een leveranciersgebonden produkt. In situaties waarin documenten voor een onbepaalde tijd ter beschikking moeten blijven zou een leveranciersgebonden produkt een risico zijn. Hoe zwaar aan dit laatste argument getild moet worden, is onduidelijk. Van de Kodak standaard PCD, voor de Photo-CD, is lange tijd hetzelfde beweerd maar het is nu wel het meest gebruikte opslagmedium voor beeldmateriaal met uitstekende referenties. Hetzelfde geldt voor Windows als besturingssysteem. Ook dit is een leveranciersgebonden produkt, maar niemand zal vanuit dit gegeven besluiten Windows niet te gebruiken, men volgt gewoon de markt.

### **3.3.2.5 Mogelijke voorkeuren**

Welk van de besproken tekstdocument-formaten de voorkeur verdient, is moeilijk te beantwoorden. Duidelijk is alleen dat men de markt volgt en formaten kiest waarvoor ook betaalbare en gebruiksvriendelijke produkten zijn.<sup>7</sup> Theoretisch gaat de voorkeur uit naar leveranciers- en tijdsafhankelijke produkten maar de praktijk dwingt vaak tot andere keuzes.

Uitgevers van wetenschappelijke tijdschriften zullen de toon aangeven wat betreft de externe toevoer van elektronische documenten. Hoewel Elsevier Science en Kluwer hebben aangekondigd bepaalde produkten in SGML te gaan aanbieden, (Dempsey & Heijne, 1995), worden momenteel de meeste documenten in PDF, PS of TIFF aangeleverd. Met name het PDF-formaat zou belangrijk kunnen worden omdat de leverancier van Acrobat (en dus PDF) met Netscape een overeenkomst aangegaan is om Acrobat als documentformaat te integreren. Dit zou kunnen betekenen dat PDF een belangrijker positie zal gaan innemen dan eerst werd aangenomen.

---

<sup>7</sup>. Het ODA-formaat wordt in de literatuur over het algemeen omschreven als superieur aan SGML, maar door het gebrek aan produkten wordt ODA wereldwijd nog nauwelijks toegepast.



Wat betreft de interne produktie en daarmee samenhangende standaards, zal veel afhangen van de mogelijkheden die geboden zullen worden door de standaard tekstverwerkers om de output te genereren in formaten als SGML en HTML. Veel zal ook afhangen van copyright overwegingen van de auteurs. Zelfs wanneer een bibliotheek de voorkeur zou geven aan SGML, iets wat in het SURFdoc- en Premium-project geadviseerd wordt, dan betekent dat nog niet dat de auteur dit zou willen. Het interne aanbod zal dan ook afhangen van de mate waarin een bibliotheek kan garanderen dat bij een beschikbaarstelling buiten het lokale netwerk een onveranderlijk formaat aangeboden wordt. D.w.z. dat er een automatische convertering moet plaatsvinden van bijvoorbeeld SGML naar PDF of een ander onveranderlijk formaat. Ook wanneer copyright-overwegingen geen rol spelen, maar de beschikbaarstelling via het Internet, dan is een keuze niet gemakkelijk. Het gebruiksgemak en verspreiding van HTML zal afgewogen moeten worden tegen de portabiliteit en het meerwerk van een volledige codering in SGML. Daarnaast spelen ook nog de resultaten een rol van het onderzoek naar de mogelijkheid om de interface, van in SGML opgeslagen documenten, te baseren op HTML.<sup>8</sup> Ook de bruikbaarheid van de onlangs op de markt gebrachte eerste Public Domain browser voor SGML-documenten moet in de overwegingen meegenomen worden.

De keuze voor een bepaald tekstdocument-formaat voor bibliotheken is dus niet gemakkelijk en er spelen vele factoren een rol. Het ziet ernaar uit dat ook de bibliotheken marktvolger zullen moeten blijven en vele formaten moeten blijven ondersteunen.<sup>9</sup> Een uitspraak over de rol van bibliotheken, als intermediar of als opslagplaats, zal hier niet veel aan veranderen. Een eensluidend advies zou pas gegeven kunnen worden wanneer er internationaal tot een standaardisering besloten wordt. Gezien de veelheid aan initiatieven op dit gebied en het vrije karakter van Internet, zal dit lang op zich laten wachten. Het enige waar momenteel naar gekeken zou kunnen worden is gebruik, gebruiksgemak en kosten, en vanuit deze optiek is het HTML, PDF en TIFF die de toon aangeven.

### **3.3.3 Beelddocument-formaten**

#### **3.3.3.1 Veel gebruikte formaten**

Veel gebruikte beeldformaten zijn GIF, PCD, JPEG, BMP, PIC, TARGA, PNG, en FIF. Belangrijke overwegingen bij een beeldformaat keuze, vormen:

- de mogelijke convertering van afbeeldingen vanuit het Kodak Photo-CD, het zgn. PCD-formaat dat op dit moment een de facto standaard is voor de opslag, c.q. back-up van fotomateriaal
- de mogelijkheid to inzoomen
- de compressie ratio die met het formaat bereikt kan worden, en
- een platform onafhankelijke gebruiksmogelijkheid

Omdat in dit rapport ook een Netscape interface in de overweging meegenomen moet worden (toekomstige thuiswerkstations en/of een uniforme interface voor de beelddatabase), vormt het mogelijk gebruik binnen Netscape ook een criterium. Formaten die dan te gebruiken zijn, zijn JPEG, GIF, BMP, PNG en FIF. Wanneer naar de conversietijd van een PCD-bestand en naar de beeldkwaliteit van een formaat gekeken wordt dan blijven JPEG en FIF over. Hierbij pleit de beeldkwaliteit in het voordeel voor FIF en de conversietijd in het voordeel van JPEG, (Flerackers e.a., 1996, p.17)

#### **3.3.3.2 JPEG**

JPEG is een standaard die ontwikkeld is door de Joint Photographic Experts Group. Het is een populaire compressiemethode voor kleurenimages, waarbij nauwelijks verlies van kwaliteit optreedt en dat erg

---

<sup>8</sup>. Deze strategie wordt gevolgd aan het "Electronic Text Center" van de University of Virginia waar alle on-line teksten in SGML opgemaakt worden volgens de TEI richtlijnen, (zie Seaman 1996).

<sup>9</sup>. In het WebDOC-project van Pica is besloten voor opslag in de formaten HTML, PDF, PS, TIFF en Tex dat zich goed leent voor scheikundige en wiskundige symbolen.

effectief is bij lage compressie-ratio's van zo rond de 25:1, (Furht 1995, p.172).

Een van de belangrijkste nadelen van JPEG-gecomprimeerde beelden is dat ze resolutie afhankelijk zijn. D.w.z. dat wanneer het decomprimeerde beeld in hogere resoluties afgebeeld wordt dan het origineel, het weergegeven beeld blokkerig wordt vanwege pixel-replicatie. Inzoomen op een JPEG-gecomprimeerd beeld gaat dan ook ten koste van de beeldkwaliteit. Bovendien bieden grafische kaarten en printers steeds hogere resoluties en dit zou kunnen betekenen dat JPEG-beelden steeds opnieuw gescand en gecomprimeerd moeten worden als men de voordelen van de nieuwe technologieën wil blijven bieden. Dit geldt overigens voor alle eerder genoemde zogenaamde 'raster-based-graphics' zoals, BMP GIF en zijn navolger PNG. Beeld-formaten die deze problematiek niet kennen zijn zogenaamde 'fractal-based-graphics' die inzoomen mogelijk maken omdat de mathematische eigenschappen van beeldelementen bij resolutieverandering behouden blijven. Het bekendste fractale formaat is FIF.

### 3.3.3.3 FIF - Fractal Image Format

Dit formaat is gebaseerd op fractale compressie waarbij een afbeelding in stukken gebroken wordt en waarmee, via bepaalde berekeningen, de gehele afbeelding opnieuw berekend en weergegeven kan worden. Fractale compressie en decompressie is een asymmetrisch proces. D.w.z. dat compressie veel meer tijd in beslag neemt dan decompressie. Wanneer aan de compressie-tijden geen stricte grenzen gesteld worden dan kunnen zeer hoge compressie-ratio's bereikt worden zonder dat dit een toename in decompressie-tijd tot gevolg heeft of een afname van beeld-kwaliteit.

Bij een vergelijking tussen JPEG compressie en fractale compressie bleek het volgende (Furht 1995, p.176):

*"In the case of an 800 x 800 color image, which requires 1.44 MB of disk space, the JPEG software compression and decompression took 41 seconds each on a 386/33 MHz system.... The fractal image compression was done in 8 minutes, but the decompression took only 7 seconds. .... "....on several 320 x 200 grayscale images, fractal compresssion ratio was in average 15, while JPEG compression ratios varied between 3 and 20, depending on selected image quality. .... In summary, in contrast to JPEG, fractal compression offers fast decompression, resolution independence, and high compression ratios. It can be usefull for asymmetric multimedia applications where quick access to high-quality images is important".*

Dezelfde conclusie wordt ook getrokken in het rapport van Flerackers e.a.(1996) waar naast de inzoom mogelijkheid ook het feit genoemd wordt dat voor FIF geen extra hardware nodig is. Flerackers e.a. (1996, p.17) concluderen:

*"Fractal image compression is most appropriate, if small file sizes with superior image quality are important".*

### 3.3.4 Videodocument-formaten

In de literatuur wordt voor videodistributie over netwerken in feite maar over één standaard gesproken, en dat is MPEG.<sup>10</sup>

MPEG werkt met interframe compressie. Dit i.t.t. intraframe compressie dat elk individueel video-frame comprimeert en opslaat. Interframe compressie is gebaseerd op het idee dat, ondanks er handelingen plaatsvinden, de achtergronden in de meeste videostukken relatief stabiel blijven. M.a.w. er is een stuk redundantie. Compressie vindt plaats door het creëren van een sequentie van zogenaamde I-, P- en B-frames. Het aantal I-frames binnen een videosegment, of ook wel referentie-frames genoemd, kunnen zelf worden bepaald met de instelling van de apparatuur. Gebruikelijk is dat ze zo om 9 tot 15 frames geplaatst worden. Deze frames bevatten een compleet videobeeld en worden het minst gecomprimeerd. Gebruikt worden deze I-frames voor, videorecorder-functies zoals 'fastforward' en 'fastback' en als mogelijk adres voor een database. Omdat de tussen de I-frames gelegen beelden slechts weinig zullen verschillen van de I-frames kan tijdens de comprimering volstaan worden met de opslag van alleen de verschillen tussen deze beelden. Deze verschillen worden als P- of als B-frame opgeslagen waarbij bij de berekening van P-frames alleen gebruik gemaakt wordt van voorafgaande beelden en bij B-frames van

<sup>10</sup>. Voor alternatieve en experimentele video-(compressie-) formaten zij verwezen naar Furht (1995), hoofdstuk 8.

voorafgaande én navolgende beelden. Hiermee kunnen compressie-ratio's gehaald worden van meer dan 200:1.

Een herhaalbare sequentie die zich in tal van applicaties als effectief bewezen heeft is: IBBPBBPBB. Hiermee wordt een relatief hoge compressie-ratio bereikt terwijl 'random acces', d.w.z. de adresseerbaarheid van videosegmenten, om de 9 beelden (I-frames) mogelijk is, (Furht 1995, p.146). Bij 25 frames per seconde zou dit betekenen dat een MPEG-gecomprimeerde video om de 360 milliseconde (9 gedeeld door 25) adresseerbaar is. Wanneer een nog hogere 'random acces' gewenst zou zijn dan kan men de sequentie geheel als I-frames coderen. MPEG wordt dan identiek aan M-JPEG (Moving-JPEG). Hoge compressie-ratio's kunnen echter alleen verkregen worden wanneer er in de sequentie veel B-frames gebruikt worden. Voor de compressie van audio-data bezit MPEG een ingebouwd algoritme met compressie-ratio's van 5:1 tot 10:1. De sterke kanten van MPEG-compressie vormen:

- random access,
- fast forward/reverse searches,
- reverse playback
- audio-visual synchronization
- robustness to errors
- editability
- format flexibility, and
- cost trade-off.

Onderscheid moet er gemaakt worden tussen MPEG-1 en MPEG-2. MPEG-1 is bedoelt voor de compressie van videobeelden met een resolutie van 352 x 240, bij 30 beelden per seconde. De resulterende datasnelheid na comprimering varieert tussen de 1.2 en 3 Mbps. Een nadeel van MPEG-1 is de beeldweergave niet schaalbaar is. Een MPEG-1 stream moet altijd gedecodeerd worden tot de volledige resolutie, (Girod, 1994).

MPEG-2 is bedoelt voor de comprimering van videobeelden met hogere resoluties zoals, CCIR 601 (720x486), EDTV (960x486) en HDTV (1920x1080). Bij 30 beelden per seconde variëren de datasnelheden respectievelijk tussen de 5 en 10, 7 en 15 en 20 en 40 Mbps. Bij 60 beelden per seconde, wat MPEG-2 aankan, verdubbelen de datasnelheden.

Het zal duidelijk zijn dat het gebruik van MPEG-2 alleen mogelijk is, wanneer bandbreedte geen probleem vormt en een hoge kwaliteit voorop staat. Het gebruik van MPEG-2 moet ook overwogen worden wanneer een video naar multi-media-werkstations gezonden wordt met verschillende AV-specificaties. Met MPEG-2 kunnen namelijk videosignalen verdeeld worden in bit-streams die de video representeren in verschillende resoluties, beeldkwaliteiten of beelden per seconde. Uit deze onderscheiden data, die tegelijkertijd verzonden worden, pakt het individuele werkstation alleen dat signaal dat ook verwerkt kan worden, (MPEG Primer, 1996).

### **3.3.5 Geluids-documenten**

Geluidsbestanden worden momenteel in veel formaten aangeleverd. De meest bekende zijn: AIF, AU, SND, MID, WAV, RA. Alle formaten worden door Flerackers e.a (1996) kwalitatief als goed beoordeeld. De voorkeur gaat bij hun uit naar het zogenaamde .WAV-formaat van Microsoft ondanks het formaat resulteert in relatief grote geluidsbestanden. De bekendheid met het formaat, het feit dat het zowel bij PC- en Mac-platforms gebruikt kan worden en ook door Netscape 3.0 ondersteund wordt, geven de doorslag.

## 4 Document transport via netwerken

### 4.1 Bandbreedte

Een belangrijke grootheid die bij het transport van multi-media over netwerken een rol speelt, is de bandbreedte of bandwijdte, c.q. het aantal bits dat per seconde door de bekabeling van het netwerk verzonden kan worden. Aan de bandbreedte zijn grenzen gesteld en deze worden bepaald door technische factoren zoals de soort bekabeling, gebruikte schakeltechniek, apparatuur etc.. Ethernet, een veel gebruikte netwerktechnologie met coax-bekabeling, kent bijvoorbeeld een bovengrens van 10 Mb/s. Fast Ethernet en VG-Anylan kennen een bovengrens van 100 Mb/s maar vragen hiervoor wel aangepaste schakelapparatuur en een zgn. UTP (categorie 5)-bekabeling.

### 4.2 Protocol

De tweede belangrijke grootheid, het netwerk-protocol, staat voor de manier waarop de data op het netwerk gedistribueerd worden. Kenmerkend voor de juist genoemde netwerken, is dat ze ontworpen zijn ter maximalisatie van netwerkgebruik voor relatief grote groepen gebruikers waarbij iedere gebruiker hetzelfde dienstenniveau toebedeeld krijgt. Bij deze, zgn. 'fairness-based' protocollen, waar meerdere computers op eenzelfde segment van de netwerkbekabeling zijn aangesloten, daalt de performance wanneer het aantal gebruikers toeneemt. Dit soort protocollen zijn adequaat als het aantal gebruikers van het lokale netwerksegment laag is en variable vertragingen in de datatransmissie voor de gebruikte applicaties geen probleem vormen zoals bij tekst en afbeeldingen. Anders is het gesteld bij video waarbij, voor bijvoorbeeld de lipsynchronisatie, een strakke synchronisatie tussen de beeld- en geluids-data vereist is en voor een ononderbroken weergave een minimale datatransmissie-snelheid gegarandeerd moeten kunnen worden die bij de meeste formaten rond 1.5 Mb/s ligt. Bij een Ethernet-netwerk, met een bovengrens van 10 Mb/s, blijkt dan ook dat de performance al bij 5 simultane video-gebruikers achteruit gaat en het beeld vertraagd weergegeven wordt. Bij 100 Mb/s-netwerken zou dit aantal rond de 50 gebruikers liggen, vooropgesteld dat er geen ander netwerkgebruik is.

### 4.3 ATM en Fast-Ethernet

Opmerkelijk is dat, hoewel netwerken met een bovengrens van 100 Mb/s mogelijkheden bieden voor de distributie van beeldmateriaal, sommige auteurs Fast Ethernet of AV-AnyLan als interim-oplossingen zien. Met deze technieken zouden alleen snelle lokale upgrades gerealiseerd kunnen worden maar toekomstvast zouden ze niet zijn. Niemegeers (1995) is zelfs van mening dat technische maatregelen ter verlenging van de levensduur van de huidige netwerkstructuren al gauw kapitaalvernietiging betekenen en men op den duur niet om vervanging van de bestaande netwerkvoorzieningen heen kan. De beleidsvraag betreft, volgens hem, alleen het moment van opwaardering, waarbij verhuizing en verbouwingen de overgang zouden vergemakkelijken. Bovendien zouden de kosten van de noodzakelijke overdimensionering, ten gevolge van een onvermijdelijke en toenemende vraag naar multi-mediale netwerkdiensten, bij nieuwe aanleg beperkt zijn. Ter ondersteuning van multi-media zou meteen gekozen moeten worden voor protocollen zoals ATM (Asynchronous Transfer Mode) waarmee aan gebruikers bandbreedtes toegewezen kunnen worden en/of aan speciale data, zoals video, hogere prioriteiten toegekend kunnen worden. Het voordeel hiervan is gelegen in de gegarandeerde bandbreedte waarmee variabele vertragingen voorkomen kunnen worden. Bovendien kent ATM geen bovengrens voor netwerksegmenten en kunnen er, in theorie, datatransmissie-snelheden van 2.5 Gbit/sec. gehaald worden. Ook Furht (1995, p.47) concludeert na een vergelijking van een aantal netwerk-technologieën, dat:

*"....,it is clear that ATM is the superior switching technology for multimedia communications. ATM is capable of supporting multimedia traffic, it provides various types of connectivity, it guarantees performance, its bandwidth is very large (up to several Gbit/s), its end-to-end delay is relatively low, and it guarantees media synchronization".*

Schenau (1995a) is minder voortvarend en verwacht dat ATM in eerste instantie ingezet zal worden als subnetwerktechnologie voor b.v. de backbone van een netwerk en dat netwerksegmenten voorzien zullen worden van Fast-Ethernet-technologie. Hoewel Schenau de redenen voor deze verwachting niet duidelijk maakt, kunnen er toch een aantal genoemd worden.

Ten eerste vragen beide technologieën om UTP-bekabeling, maar is de ATM-apparatuur in vergelijking met Fast-Ethernet relatief duur. Ten tweede ondersteunt de ATM-apparatuur die momenteel te koop is slechts datasnelheden van 25 tot 155 Mb/s, wat niet veel hoger is dan de 100Mb/s van Fast-Ethernet. Hogere snelheden, momenteel tot 622 Mb/s, zijn ook bij ATM pas mogelijk met gebruik van dure optische bekabeling. En ten laatste, kunnen met Fast-Ethernet toch, zij het niet theoretisch maar wel praktisch, minimale datatransmissie-snelheden gerealiseerd worden door verdere netwerk-segmentering en een intelligente bekabeling van werkstations. Bovendien wordt er op congressen, voor Fast-Ethernet met UTP-bekabeling, over datatransmissie-snelheden gesproken die vele malen hoger liggen dan het huidige maximum van 100 Mb/s. Kortom: ook voor de middellange termijn biedt Fast-Ethernet voor de zwaartse toepassingen, zoals digitale video, goede distributie-mogelijkheden.

Vereiste datasnelheden voor verschillende media:

Information Type	Bit Rate	Quality and remarks
DATA	Wide range of bit rates	Continuous, burst and packet-oriented data
TEXT	Several Kbps	Higher bit rates for downloading of large volumes
GRAPHICS	Relatively low bit rates Higher bit rates - 100 Mbps and higher	Depending on transfer time required Exchange of complex 3D computer model
IMAGE	64 Kbps Various Up to 30 Mbps	Group-4 telefax Corresponds to JPEG std. High-quality professional images
VIDEO	64-128Kbps 384 Kbps-2 Mbps 1.5 Mbps 5-10 Mbps 34/35 Mbps 50 Mbps or less 100 Mbps or more	Video telephony (H.261) Videoconferencing (H.261) MPEG-1 TV quality (MPEG-2) TV distribution HDTV quality Studio-to-studio HDTV video downloading
AUDIO	n x 64 Kbps	3.1 KHz, or 7,5 KHz, or hi-fi baseband signals

uit: Furht e.a (1994), p.34.

## 5 Document retrieval

### 5.1 Ontsluiting elektronische documenten

Voor de ontsluiting van elektronische documenten worden doorgaans 'records' in een database aangemaakt die een verwijzing bevatten naar het feitelijke document dat op een server staat.<sup>11</sup> Voor tekstdocumenten bevat het record meestal een aantal velden voor bibliografische informatie, een aantal trefwoordvelden voor de inhoudelijke ontsluiting, en vaak een vrij-tekst- of zgn. annotatie-veld waar men een beschrijvend stuk tekst in kwijt kan. In beeld-databases wordt bij het document-record tevens een plaats ingeruimd voor het zogenaamde 'thumbnail-image' dat, als resultaat van een zoekactie, het eerst getoond wordt voordat men het eigenlijke document op de server kan opvragen. Een vergelijkbare aanpak wordt gevolgd bij video-databases, zij het dat het 'thumbnail-image' een zgn. 'key-frame-image' genoemd wordt dat handmatig, dan wel elektronisch, gekozen wordt ter representatie van de inhoud van een filmshot.

Voor zowel de bibliografische als de inhoudelijke beschrijving van elektronische documenten ligt het als UB voor de hand dat men probeert aan te sluiten bij methodieken die ook elders gebruikt worden en waardoor een zekere uitwisseling van gegevens mogelijk wordt, vergelijkbaar met de Gemeenschappelijke Catalogus.

### 5.2 Bibliografische ontsluiting

#### 5.2.1 Bibliografische ontsluiting tekst-georiënteerde documenten

##### 5.2.1.1 Pica richtlijnen voor het catalogiseren van online resources

Als Pica-bibliotheek, ligt het bovendien voor de hand dat de "Pica richtlijnen voor het catalogiseren van online resources" gevolgd worden, (Pica, 1995).<sup>12</sup> Er bestaat niet alleen al een ruime ervaring in de toepassing van de omvattendere Pica-richtlijnen voor de beschrijving van audio-visueel materiaal, maar het gebruik biedt wellicht ook de mogelijkheid tot koppeling van de lokale database met de gemeenschappelijke Pica-database.

Een online resource wordt in de richtlijnen gedefinieerd als een elektronisch document,

- dat bibliografisch identificeerbaar is,
- dat in machineleesbare vorm is opgeslagen op een elektronisch opslagmedium en
- dat online beschikbaar is

Gezien de definitie zou het bij de richtlijnen behorende beschrijvingsformaat, in principe, voor alle soorten gedigitaliseerd materiaal gebruikt kunnen worden. Hiertoe zijn in het formaat voor een aantal specifieke kenmerken van elektronische documenten extra plaatsen ingeruimd waarin men gegevens kwijt, zoals:

- editie-, c.q. update gegevens,
- plaatst en naam van de host,

---

<sup>11</sup>. Een tweetal veel gebruikte databases die men in de literatuur tegenkomt zijn de 'Oracle 7 RDBMS' en de 'Delphi 2.0 Client Server Suite'.

<sup>12</sup>. Pogingen om buiten de OPC om, met behulp van netwerkmethoden te indexerem zijn weinig bevredigend geweest. Een produkt als WAIS blijkt b.v. lastig om mee te werken en slecht gedocumenteerd. Dergelijke onvolkomenheden van netwerkmethoden pleiten voor het blijven nastreven van een integratie van de OPC met het aanbod van informatie op het netwerk. Voor een aantal nadelen van het werken met WAIS binnen bibliotheekomgeving zij verwezen naar Heijne (1995, p. 20).

- specieke materiaal aanduidingen,
- illustratie vermeldingen zoals geluid of bewegend beeld,
- toegangsgegevens van het bestand zoals de URL, filegrootte, compressieformaten, IP-adressen, software-vereisten voor het viewen, etc., en
- eventuele verwijzingen naar papieren versies van het document

Ondanks met deze extra kenmercodes voorzien wordt in het dynamische karakter van elektronische documenten, die qua vindplaats en inhoud snel kunnen veranderen, kent het huidige beschrijvingsformaat toch nog een aantal beperkingen die alleen weggenomen kunnen worden wanneer het formaat lokaal, danwel gemeenschappelijk, aangepast wordt.

### 5.2.1.2 Bibliografisch moeilijk identificeerbaar materiaal

Problematisch wordt het gebruik van het beschrijvingsformaat voor materiaal waarbij relevante gegevens voor een bibliografische beschrijving ontbreken of onbekend zijn. Zo ontbreekt bij materiaal, afkomstig van Internet of uit fotoarchieven, vaak een auteurs- of producent-naam en andere gegevens die voor de beschrijving in de pica-database verplicht zijn. Alleen het concrete object is voorhanden en objectieve metadata ter beschrijving ontbreken. Een voor de hand liggende (defensieve) reactie zou kunnen zijn dat er dusdanige kwaliteitseisen aan elektronische documenten gesteld gaan worden dat ze alleen gecatalogiseerd worden als zij de degelijkheid van geschreven tekst aannemen. In het SURFdoc-project wordt terecht de vraag gesteld of dit de juiste weg is en op termijn vol te houden is, (SURFdoc 1995, p. 51). De beperkingen van een catalogiseersysteem behoren geen beletsel te zijn bij de verdergaande introductie van elektronische documenten.

Een belangrijke discussie in deze, vormt de toekenning van zogenaamde URN's (Uniform Resource Names) en URC's (Uniform Resource Characteristics) bij documenten op het World Wide Web. Vanwege de veranderlijkheid van de URL's (Uniform Resource Locators), die aangepast moeten worden wanneer documenten naar andere machines migreren, wordt door de Internet Engineering Task Force (IETF) overwogen zogenaamde URN's toe te kennen. Met deze naamgeving zou een document tijds- en plaats-onafhankelijk identificeerbaar zijn, vergelijkbaar met het ISBN en het ISSN. Over de wijze waarop de naamgeving zou moeten plaatsvinden, en over het mechanisme dat de vertaling van URN's naar geldige URL's moet regelen, wordt nagedacht, (Task Force on Digital Archiving, 1996).

Hoewel de URN's zich momenteel nog in een conceptueel stadium bevinden zal, in afwachting van de ontwikkelingen, het beschrijvingsformaat moeten voorzien in een veld voor de retrospectieve toekenning van zo'n URN. Dit geldt ook voor de zogenaamde URC's die de 'Task Force' noodzakelijk acht en waarin men algemenere informatie over het document kwijt kan zoals eventuele download-kosten en b.v. namen van rechthebbenden.<sup>13</sup> Met de toekenning van URN's en URC's verzekerd men zich in ieder geval van extra mogelijke koppelingen tussen de database, het document en het Internet, en wat wellicht het belangrijkste is; een eenduidige identificatie van bibliografisch moeilijk identificeerbaar materiaal.

In hoeverre men in het licht van genoemde ontwikkelingen bereid is het pica-formaat en de richtlijnen op gemeenschappelijk niveau aan te passen, is moeilijk in schatten. Het zou gemakkelijk zijn wanneer er in de gemeenschappelijke database beschreven kon worden maar waarschijnlijk is dit niet. Bibliografisch moeilijk identificeerbaar materiaal zal voornamelijk in een lokale database beschreven moeten worden. Landelijke beschikbaarheid zal afhangen van, formaat-ondersteuning, auteursrechten en netwerkbependingen, en van de wijze van toekenning van URN's.

---

<sup>13</sup>. Hoewel het Pica-formaat met kenmerkcode 4201 reeds in deze mogelijkheid voorziet, is het onwaarschijnlijk dat de URC's qua opbouw en mogelijkheden met de pica-code overeenkomen.

### 5.2.2 Bibliografische ontsluiting foto- en beeld-materiaal

Naast het bovengenoemde probleem bij de beschrijving van online resources, spelen bij fotografisch materiaal, ook nog andere catalogiseer-problemen. Omdat het Pica-beschrijvings-formaat in eerste instantie ontwikkeld is voor de beschrijving van gedigitaliseerd tekstmateriaal, waarbij afbeeldingen meestal als illustratie dienen, is er weinig aandacht geweest voor bibliografische gegevens die men voor fotomateriaal belangrijk acht, (Hogenboom & Voort, 1994). Zo ontbreken in het Pica-formaat bijvoorbeeld specifieke kenmercodes voor de aard van een foto-origineel, de datum van opname, gebruikte techniek en een zogenaamd 'concept'-veld waar men kan aangeven welk thema of abstract begrip de maker heeft willen uitdrukken of met welk doel de foto is gemaakt of tot welk genre de originele foto gerekend kan worden. Deze gegevens zouden nu, bij het bestaande Pica-formaat, bij een annotatieveld ingevuld moeten worden. Zou men met deze gegevens een zoekactie willen uitvoeren dan zouden de gebruikte annotatie-velden het eerst geïndexeerd moeten worden voordat men de gebruikte gegevenssoorten als zoek sleutels zou kunnen aanbieden. Voor een daadwerkelijk gebruik zal het Pica-formaat ook hiervoor aangepast dan wel uitgebreid moeten worden.

Of en in hoeverre men op gemeenschappelijk niveau bereid is het formaat aan te passen, zal mede afhangen van het aantal gegevens die voor de bibliografische beschrijving van fotomateriaal noodzakelijk acht. Hoewel er internationaal nog lang geen eenheid bestaat in het formele beschrijven van afbeeldingen zij er wel ontwikkelingen in de goede richting, (Starre, 1996). Een zo'n goede richting is de minimum-standaard voor de uitwisseling van fotogegevens, die in Nederland geformuleerd is door het 'Samenwerkingsproject Informatisering Fotografiebestanden', (Hogenboom & Voort, 1994). Omdat het voorgestelde uitwisselings-formaat slechts een twaalfstal gegevens-soorten bevat, waarvan er acht bibliografisch van aard zijn, is het goed voorstelbaar dat deze gegevens binnen het Pica-formaat ingepast kunnen worden. Hiertoe zouden de belangengroepen in nader overleg kunnen treden.<sup>14</sup> Bij enkel lokale beschrijving zou men kunnen handelen naar eigen goeddunken en voor fotografisch materiaal eigen lokale kenmercodes kunnen definiëren.

### 5.2.3 Bibliografische ontsluiting video-materiaal

Video's vormen voor het Pica-beschrijvingsformaat in eerste instantie geen probleem. Voor de formele beschrijving verschillen video's niet wezenlijk van monografieën. Wanneer men echter weer kijkt naar de bibliografische kenmerken die men specifiek voor film belangrijk acht, zoals het genre, de 'cast' en eventuele oscar nominaties (Turner 1994, Rowe e.a. 1994), dan blijkt dat het Pica-formaat hier geen specifieke kenmercodes voor heeft. Deze gegevens zouden in het bestaande formaat weer in annotatie-velden ingevoerd moeten worden.

Een probleem van principiële aard betreft delen van video. In het huidige formaat zouden deze delen als apart bestand op de server opgeslagen moeten worden en de database-records uitgebreid moeten worden met relevante kenmercodes voor video-shots, scenes en segmenten zoals bv.:

- begin- en eind-tijden in uren, minuten, seconden of frames
- soort overgang waarmee de shot begint
- soort overgang waarmee de shot eindigt
- het type shot (close-up, medium, full)
- het deel van de dag dat verbeeld wordt in de shot of scene (ochtend, avond etc.)
- het object dat de scene definieert (acteur, plaats of lokatie)
- personen die in een scene voorkomen
- objecten die in een scene voorkomen
- geografische lokatie waar de scene speelt

Het zal duidelijk zijn dat wanneer men zou besluiten tot de ontsluiting van video-fragmenten met gebruikelijk database technieken de server erg snel vol zal raken. Men is in deze aangewezen op nieuwe technieken zoals videoservers met bijhorende programmatuur zodat de video slechts eenmaal digitaal opgeslagen hoeft te worden en, voor de identificatie van de segmenten, toch tot op frame-niveau door de database adresseerbaar blijft.

---

<sup>14</sup>. Hetzelfde voor de bibliografische beschrijving van geluidsdocumenten die in het kader van dit rapport geen object van onderzoek zijn geweest.



Hoewel het voorstelbaar is dat het beschrijvingsformaat plaatsbiedt aan 'frame'-intervallen is het ondenkbaar dat Pica een beschrijving van het geheel en zijn delen binnen de database zal toestaan. Ook bij artikelen binnen een bundel gebeurt dit niet en het zal niet waarschijnlijk zijn dat voor video een uitzondering gemaakt wordt. Lokale bibliografische beschrijving volgens een aangepast pica-formaat dringt zich ook hier op.

### **5.3 Onderwerpsontsluiting**

#### **5.3.1 Onderwerpsontsluiting elektronische documenten**

Een apart probleem bij de ontsluiting van elektronische documenten betreft de onderwerpsontsluiting. Zoals gezegd, biedt het pica-formaat een beperkt aantal mogelijkheden voor de onderwerpsontsluiting. Hiervoor zijn in het beschrijvingsformaat een aantal gemeenschappelijke en lokale kenmerkcodes gereserveerd waarin men met vakgebieds-aanduidingen en trefwoorden de inhoud zo nauwkeurig mogelijk probeert aan te geven. Voor teksten is deze manier van ontsluiten goed te verdedigen. Vaak wordt in de tekst zelf al aangegeven wat het onderwerp is en worden de meeste teksten met een bepaald doel geschreven. Verwacht mag worden, zij het met bepaalde voorbehouden, dat de potentiële gebruikers en de onderwerpsontsluiters een tekst vanuit dit doel benaderen. Ondanks de meeste databases beelden op eenzelfde manier ontsluiten als tekst, (Gudivada & Raghavan, 1994 en Furht e.a., 1995), ligt dit bij afbeeldingen toch moeilijker.

#### **5.3.2 Onderwerpsontsluiting beelddocumenten**

##### **5.3.2.1 Aparte status van beeld-materiaal**

I.t.t. boeken en andere documenten, worden beelden zelden voorzien van titels, voorwoorden, introducties of samenvattingen. De indexerder zal op zijn eigen interpretatie van de eigenlijke inhoud moeten vertrouwen, hetgeen bijzonder moeilijk is wanneer het gaat om beelden van symbolische en allegorische aard. Daarnaast zijn beelden principieel van interdisciplinaire aard. Ze bevatten altijd een variëteit aan eigenschappen die elk van potentieel belang kunnen zijn vanuit verschillende onderzoeksperspectieven. Beelden worden dan ook vaak gebruikt voor doeleinden die ver verwijderd zijn van de doelen die de maker voor ogen had. Een volledige representatie van de inhoud d.m.v. trefwoorden is bijna onmogelijk omdat de mogelijke interpretaties niet te overzien zijn. Dit moge blijken uit de onderzoeken die door Baxter & Anderson (1995) genoemd worden en waaruit blijkt dat niet alleen inter-indexeer consistentie bij beelden laag is maar ook de intra-indexeer consistentie. Bovendien levert een ongelimiteerde codering met trefwoorden een geheugen-probleem op bij de indexering.

##### **5.3.2.2 Thesaurus gebaseerde indexering**

Ter beperking van het aantal trefwoorden en verhoging van de consistentie zou men het gebruik van gecontroleerde trefwoord-thesauri kunnen overwegen. Maar een keuze is hier niet gemakkelijk. Bestaande thesauri zijn domeinspecifiek ontwikkeld en moeilijk aan te passen voor algemener gebruik. Bovendien veronderstellen ze vaak specifieke vakkennis en zijn ze moeilijk en tijdrovend in gebruik, (Baxter & Anderson 1995). Zo worden bijvoorbeeld belangrijke thesauri, zoals de AAT (Art and Architecture Thesaurus) en ICONCLASS, bekritiseerd omdat ze ofwel:

- tijdrovend in het gebruik zijn (AAT: 40 minuten/dia, ICONCLASS 30 minuten/beeld)
- alleen door goed gekwalificeerde kunsthistorici te gebruiken zijn en
- minder bevredigend werken voor gewone objecten zoals tafels, stoelen etc.

Ook de thesaurus van het op het eerste gezicht algemenere, Nederlandse uitwisselings-formaat voor fotocollecties, (Hogenboom & Voort 1994), kent een aantal van deze problemen.<sup>15</sup> Voor een goed gebruik

---

<sup>15</sup>. Voor de onderwerpsontsluiting van afbeeldingen is in dit formaat voor meerdere mogelijkheden gekozen. Het formaat onderscheidt in eerste instantie een tweetal niveau's. Het eerste niveau, een soort dakthesaurus, bestaat uit een achttal hoofdcatego-

wordt, naast cultuurhistorische kennis, ook nogal wat fotografische kennis verondersteld en zijn biomedische afbeeldingen pas na een grondige uitbreiding van de thesaurus in te delen.<sup>16</sup>

Deze domein-specifieke thesauri zouden wel goed te gebruiken zijn wanneer er onderscheiden collecties aangeboden zouden worden. Kunsthistorisch en algemener beeldmateriaal zou op onderwerp ontsloten kunnen worden met voor dit soort afbeeldingen ontworpen thesauri zoals AAT, ICONCLASS en de thesaurus van het foto-uitwisselingsformaat. Voor medisch materiaal zou men een beroep kunnen doen op de "Medical Subject Headings". Voor het databaserecord zou dit betekenen dat men collectie specifieke beeld-trefwoord-velden zou moeten definiëren en bij opname in de database moeten beslissen tot welke collectie(s) de te beschrijven afbeelding zou behoren. Een zekere willekeur bij dit laatste, kan niet uitgesloten worden gezien de interpretatie mogelijkheden bij afbeeldingen. Een ander probleem bij deze strategie is dat van de gebruiker verondersteld wordt dat hij kennis heeft van de manier waarop de onderscheiden thesauri de werkelijkheid indelen en welke overwegingen bij de constructie van de thesauri een rol gespeeld hebben.

Een andere strategie zou het gebruik van vrije trefwoorden kunnen zijn. Vanuit de gebruikerskant is deze strategie wel aan te bevelen omdat de gebruiker de database vanuit zijn interessegebied zal bevragen en niet geïnteresseerd zal zijn in alle mogelijke perspectieven die in de onderscheiden aangeboden trefwoorden tot uitdrukking kunnen komen. Het is echter zeer de vraag of de trefwoorden die voor kunsthistorische en medische beeldcollecties gebruikt worden in de bestaande, lokale of landelijk gebruikte, trefwoordlijsten in te passen zijn. Een integratie van deze trefwoorden zal in ieder geval tot gevolg hebben dat er bij beeldmateriaal geen onderscheid meer gemaakt kan worden tussen de 'concepten', c.q. ziektebeelden, die de afbeeldingen tot uitdrukking zouden kunnen brengen, en de benoemde objecten die toevallig zichtbaar zijn.

---

rieën (persoon, dier, plant, object, activiteit, gebeurtenis, ruimtelijke omgeving en non-figuratief) die elkaar wederzijds uitsluiten en waar in principe iedere afbeelding in onder te brengen valt. Het specifiekere niveau bestaat uit een aantal subcategorieën (kind, jongen, meisje etc. binnen de persoonscategorie) die door de instelling zelf uitgebreid kunnen worden zolang er relaties met de dakthesaurus gelegd worden. Omdat met deze indeling veel afbeeldingen moeilijk en/of onbevredigend ontsloten worden, kent het formaat ook nog een zogenaamd 'concept'- en een 'voorstelling specifiek'-trefwoordveld. Het 'concept'-veld dient voor aanduiding van het thema of abstract begrip dat de maker van de afbeelding heeft willen uitdrukken. Het 'voorstelling specifiek'-veld kan gebruikt worden voor trefwoorden of codes uit een eigen, reeds gehanteerde trefwoordenlijst of classificatie (codering). Hoewel het gebruik van deze thesaurus op eerste gezicht aantrekkelijk lijkt, levert het voor een diepere onderwerpsontsluiting dan het voor de uitwisseling verplichte voorstellings-trefwoord, geen besparing op.

<sup>16</sup>. Met name biomedische afbeeldingen die veelal gemaakt worden ter illustratie van een pathologische afwijking, ziekte of toestand, zullen met enkel de twee niveau's onbevredigend te plaatsen zijn omdat bij de constructie van de thesaurus de gewone zichtbare wereld als criterium gefunctioneerd heeft en het beeld zelf, de hoofdcategorie-keuze dient te bepalen.

### 5.3.2.3 Content-based-image retrieval

Een alternatief voor trefwoord-indexering van beeldmateriaal, of beter, een aanvulling op trefwoord-indexering, dat in de literatuur aangedragen wordt, heet 'content-based-image-retrieval' en is gebaseerd op beeldeigenschappen zoals bijvoorbeeld kleur en objectvormen. Deze manier van indexering is dermate afwijkend van de gangbare onderwerpsontsluiting en van de manier waarop gebruikers een database kunnen bevragen, dat Furht e.a. (1995) spreken van een ander paradigma, (p.226-227).

*"Queries based on image content require a paradigm which differs significantly from that of both traditional databases and text-based image understanding systems. First, such queries cannot be as logically rigorous as those expressed in text. Images are not atomic symbols, so symbol equality is not a particularly realistic predicate. Instead, searches tend to be based on similarity; and resemblance is more important than perfectly matching bit patterns. On the other hand all that is similar is not necessarily correct, so this paradigm tends to entail the retrieval of false positives which must then be manually discarded by the user. In other words the paradigm rejects the idea that queries may be expressed in terms of necessary and sufficient conditions which will determine exactly which images one wishes to retrieve. Instead, a query is more like an information filter, defined in terms of specific image properties, which reduces the user's search task by providing only a small number of candidates to be examined; and it is more important that candidates which are likely to be of interest are not excluded than it is that possible irrelevant candidates be included".*

In principe ondersteunt het paradigma dus drie visuele zoekmogelijkheden. Men zou ze als volgt kunnen omschrijven:

1. *Zoeken met beeldeigenschappen.* Gebruikers specificeren een eigenschap van een afbeelding zoals b.v. een kleurverdeling (of schetsen b.v. een object) en het systeem zoekt afbeeldingen met gelijkende eigenschappen. Sommige systemen bieden ook nog de mogelijkheid tot specificatie van de locatie waarin de eigenschap moet voorkomen.
2. *Zoeken m.b.h.v. beeldgelijkenis.* Gebruikers bieden het systeem een complete afbeelding als voorbeeld of template aan. Het systeem zoekt vervolgens afbeeldingen die visueel gelijkend zijn. Sommige systemen bieden de mogelijkheid om bepaalde gebieden van een voorbeeld te negeren zodat het gevonden aantal vergroot kan worden.
3. *Zoek-verfijning of systematisch bladeren.* Het resultaat van een van de zojuist genoemde manieren is een verzameling beelden. Een blader-vraag is er een die de vorige vraag verfijnt door ofwel een beeld uit de verzameling als voorbeeld te nemen of door een of andere verandering van de parameters van de originele zoekvraag.

### 5.3.2.4 Relevante beeld-eigenschappen

Het antwoord op de vraag welke beeld-eigenschappen voor bibliotheekgebruik relevant zijn, is afhankelijk van de vertrouwdheid van de gebruiker met het aanwezige beeldmateriaal. Wanneer een gebruiker goed op de hoogte is van de beeldcollectie kan men een zogenaamde model-benadering hanteren. Hierbij worden, voor de effectiviteit van het systeem, een hele reeks voorwaarden gedefinieerd waaraan de afbeeldingen voor opname in database moeten voldoen. Een voorbeeld is een gezichtsherkenning systeem waarbij foto's steeds vanuit dezelfde hoek, op gelijke afstand, met dezelfde belichting en vergrotingsfactor genomen worden. Beeld-eigenschappen zoals neuslengte, oogafstand zijn dan vrij gemakkelijk te definiëren en te gebruiken voor een zoekactie.

Anders is het bij een algemene benadering waarbij niet zulke stringente voorwaarden te definiëren zijn. Een bibliotheekcollectie met medisch, kunsthistorisch en algemeen beeldmateriaal is een voorbeeld. Van de gebruiker mag niet verwacht worden dat hij de gehele collectie kent en weet met welke specifieke

beeld-eigenschappen hij de, eventueel apart aangeboden collecties, kan bevragen. Voor zulk een algemener systeem komen dan ook meer algemene beeld-eigenschappen in aanmerking. Uit onderzoek blijkt dat met name,

*"...color, texture, shape of component objects, and relationships among edges which may be expressed in terms of line sketches",*

de meest effectieve zijn, (Furht e.a. 1995, p.229).<sup>17</sup>

### **5.3.2.5 Trefwoord-indexering versus Content-based-image retrieval**

Wanneer men de hoeveelheid literatuur bekijkt die over 'content-based-image-retrieval' verschenen is, zou men snel de indruk kunnen krijgen dat dit de oplossing voor de onderwerpsontsluiting van beeldmateriaal zou zijn. Hoewel het bij sommige toepassingen het enige middel zou kunnen zijn dat ter beschikking staat moet het toch niet als een vervanging van trefwoord-indexering gezien worden. Trefwoorden-indexering zal belangrijk blijven.

Een eerste reden is dat beeld-eigenschappen die bij "content-based-retrieval" gebruikt worden niet de semantische inhoud van een beeld kunnen vatten en er geen herkenning mee uitgedrukt kan worden. Een systeem is bijvoorbeeld wel in staat, met vooraf opgegeven eigenschappen of afbeeldingen, andere stadsgezichten te identificeren maar kan niet op basis van de verstrekte informatie bepalen of het Amsterdam of Rotterdam is. Met de toekenning van het trefwoord Amsterdam of Rotterdam is deze herkenning wel tot uitdrukking te brengen.

De tweede reden vormt de snelheid waarmee databases met trefwoorden doorzocht kunnen. Het zoeken met behulp van beeldeigenschappen is vele malen langzamer. Een laatste reden waarom, volgens Furth (1995), trefwoorden belangrijk zullen blijven, vormt het ingeburgerde gebruik van trefwoorden.

### **5.3.3 Onderwerpsontsluiting videodocumenten**

Het Pica beschrijvingsformaat biedt voor de onderwerpsontsluiting van video in principe de dezelfde mogelijkheden als voor tekstmateriaal. Voor de video als geheel, die men als een monografie zou kunnen opvatten, gelden in feite dezelfde regels en kent men trefwoorden toe die betrekking hebben op de inhoud van de video als geheel. Anders is het weer gesteld met delen van een video die men op onderwerp zou willen ontsluiten. In het Pica-formaat is, evenals voor bibliografische gegevens van een video- of film-clip, geen plaats voor de inhoudelijke informatie van delen ingeruimd. Hiervoor zou lokaal weer een en ander geregeld moeten worden.

---

<sup>17</sup>. Hoe deze eigenschappen op zowel scenes (beelden met 0 of meer objecten) als op objecten binnen scenes toe te passen, te berekenen en te vergeleken zijn, zij verwezen naar Furht e.a. (1995), hoofdstuk 11.

Omdat het praktisch en technisch ondoenlijk is om elk videoframe van trefwoorden te voorzien zijn de inspanningen van het onderzoek naar video-ontsluiting gericht op zogenaamde 'parsing'- of 'partitioning'- technieken en de automatische herkenning van zogenaamde 'key-frames' die de inhoud van een film- of video-shot het beste representeren. Met de 'parsing'-technieken wordt een soort structurele analyse van de video uitgevoerd en kunnen automatisch shot- en andere (effect-)overgangen opgespoord worden.<sup>18</sup> Vervolgens bestaat de mogelijkheid om binnen de gevonden shots automatisch 'key-frames' te laten zoeken.<sup>19</sup> Het zijn deze 'key-frames' die gecomprimeerd in de database opgeslagen worden, voorzien met:

- een verwijzing naar de video en de bijbehorende frame-intervallen,
- trefwoorden en/of tekstannotaties en/of
- beeldinformatiegegevens verkregen door 'content-based'-analyse.

Met deze technieken wordt de inhoudelijke ontsluiting van delen van een video in feite teruggebracht tot de ontsluiting van statische beelden met dezelfde mogelijkheden en problemen. Voor manieren waarop structurele informatie over de video weergegeven kan worden (de plaats van de shot in de gehele film of video) en mogelijke interacties met deze informatie, zij verwezen naar Furth e.a. (1995) en Hjelsvold e.a. (1995).

Een belangrijk punt dat in de literatuur bij video-databases naar voren wordt gebracht, maar ook voor beeld-databases zou kunnen gelden, is dat van de gangbare relationele database benadering afgeweken moet worden. De voorkeur zou, volgens Lee & Li & Xiong (1997) moeten uitgaan naar een aangepaste object-georiënteerde benadering die beter met de complexe video-data om zou kunnen gaan.

*"....., the entity relationship (ER) model for relational databases is not sufficient for describing video data, since it may involve complex objects as well as more relationships that can not be adequately accommodated by relational and ER modeling approaches. That is why object-oriented databases (OODBs) are being utilized as a better means due to their increased power of semantic modeling, in particular behavioral modeling and encapsulation, complex object support, type hierarchy and inheritance, and so on. Despite these powerful features, conventional OODB systems are still not sufficient for video data management, particularly with respect to such subjective activities as video editing and production, which form the basis for facilitating the desired video data manipulability".*

Een van de belangrijkste redenen waarom OODBs uitbreiding behoeven, is de mogelijkheid tot het 'dynamisch' toevoegen en verwijderen van recordvelden. Omdat betekenisvolle video-segmenten vaak pas achteraf geïdentificeerd worden, c.q. nadat de video in de database opgenomen is, moet de mogelijkheid bestaan dat video-segmenten met behulp van flexibele data structuren beschreven kunnen worden. Omdat het ondoenlijk is om bij de opbouw van database, met alle interesses rekening te houden en alle mogelijke eigenschappen vooraf te definiëren, moeten er nieuwe klassen gedefinieerd maar ook verwijderd kunnen worden. Conventionele OODMs zouden zich hier niet goed voor lenen. Voor een systeems-architectuur die hier wel rekening mee houdt zij verwezen naar Lee & Li & Xiong (1997) zelf.

---

<sup>18</sup>. Overzichten van te gebruiken technieken zijn te vinden in Ahanger & Little (1996) en Furth e.a (1995). Experimentele data met betrekking tot parsing-technieken gebaseerd op MPEG-sequenties (data-streams) wijzen erop dat zogenaamde hybride algoritmes, meerdere vergelijkingstechnieken gebaseerd op I- én B-frames van de MPEG stream, het beste resultaat opleveren. Alleen hybride algoritmes blijken alle shot-overgangen en speciale effecten zoals pan's etc te localiseren. (Furth 1995, p.333 en Shen & Delp 1996).

<sup>19</sup>. Zie: Hanjalic & Legendijk & Biemond (1996), Ardizzone & Cascia (1997) en Lee & Li & Xiong (1997).

## 6 Document Display, Manipulation and Copy

### 6.1 De interface

Een belangrijke factor, voordat het eigenlijke document opgevraagd en getoond kan worden, vormt de interface van de database. De beste oplossing zou hier zijn dat een gebruiker zich wendt tot een vertrouwd catalogusscherm (momenteel nog de OPC) waar hij met behulp van een of meerdere zoekleutel(s) de beschrijvingen zoekt, van voor hem relevante documenten. Wanneer het om een elektronisch document gaat, kan hij met een eenvoudige doorschakelingsmogelijkheid het document van de server naar het scherm of schijf halen.<sup>20</sup> Bij beeld- en video(clip)-documenten is er nog een tussenstap. Voordat de gebruiker de mogelijkheid tot downloaden, printen of viewen krijgt, worden 'thumbnail-images' getoond met aanvullende informatie ter controle van de relevantie. Hiermee kunnen behoorlijke netwerkbelastingen worden.<sup>21</sup>

Een andere situatie is, wanneer onderscheiden catalogi met elk hun specifieke zoekleutels en interfaces, aangeboden worden. Een voor de hand liggend onderscheid, in navolging van het onderscheid tussen de relationele en object-georiënteerde benadering, is hier een interface voor de tekst-database (Pica) en een meer visuele interface voor de beeld-, c.q. video-database die uitgebreidere browse-mogelijkheden moeten bieden dan een interface voor teksten. Omdat bij beeldmateriaal, en zeker bij het gebruik van content-based-retrieval dat gebaseerd is op gelijkenis, het zoekresultaat vele malen groter zal zijn dan bij tekst-retrieval, is een visuele (browse-)interface uitermate belangrijk; zowel voor het stellen van een zoekvraag als voor een verfijning van de resultaten van een zoekvraag. Furth (1995, p.31), formuleert dit punt als volgt:

*"Recovering information from such a database will, in general, involve a tightly coupled integration of retrieval and browsing. The purpose of retrieval is to compose and satisfy precisely formulated questions, but browsing is often necessary first before one begins to understand just how a question should be formulated. Similarly, very large databases may provide very large responses to queries; so browsing is likely to be just as important after a query has been processed as before it has been articulated".*

Een minimum-eis die dan ook aan de interface voor beelden en video gesteld wordt is, dat de resultaten van een zoekactie in de vorm van een verzameling dia-beelden getoond moet kunnen worden waarin gebladerd kan worden. Bovendien dient de interface selectiemogelijkheden te bevatten voor het viewen, downloaden, en printen van de documenten evenals meerdere mogelijkheden tot verfijning van de resultaten van een zoekactie met behulp van beeldeigenschappen en/of andere discriminerende informatie.<sup>22</sup>

Met welke programmatuur deze functionaliteit van de interface het beste vormgegeven kan worden is moeilijk te beantwoorden. Het antwoord op deze vraag is mede afhankelijk van het antwoord op vraag of de toegang tot de documenten dient verlopen via het OPC-scherm of dat een eigen interface

---

<sup>20</sup>. Een situatie die momenteel met het WEBcat en WEBdoc-project van PICA nagestreeft wordt.

<sup>21</sup>. Het tonen van 'thumbnail-images', c.q. lage resolutie afbeeldingen en/of lage resolutie key-frames van video, zou kunnen betekenen dat het gebruik van de Pica-database voor de beschrijving van beeld- en videodocumenten uitgesloten is. Zelf bij zeer sterk gecomprimeerde 'thumbnail-images' zal het netwerktransport problemen opleveren.

<sup>22</sup>. Bij een video-clip bestaat overigens de mogelijkheid tot presentatie van zogenaamde 'micons' waarbij een aantal frames als pagina's elkaar getoond worden en waarbij de mogelijkheid bestaat om verschillende frames achtereenvolgens uit de verzameling te tillen zodat ook een stuk structurele informatie over de frames binnen de sequentie verkregen kan worden. Hiervoor zij verwezen naar Furth (1995).

ontwikkeld moet worden. Vaak wordt er bijvoorbeeld gekozen voor een Internet-browser aan de client-kant vanwege de beschikbaarheid, bekendheid, gebruiksvriendelijkheid en de platform-onafhankelijkheid. Bovendien wordt hiermee een toegang tot de database geleverd die overeenkomt met een mogelijke thuissituatie. Een ander voordeel dat hier genoemd kan worden is het gemak waarmee viewers geïnstalleerd kunnen worden en de eis van een specifieke opslagformaat voor elektronische documenten minder dwingend wordt.

## 6.2 Multimedia werkstations

Een tweede belangrijke factor bij de weergave van elektronische documenten vormt de configuratie van het werkstation. Wil men multi-mediale documenten op een adequate manier centraal beschikbaar stellen dat zal de configuratie minimaal moeten bestaan uit:

- een video/audio subsysteem en voldoende viewers die zorgen voor de convertering en/of decompressie van de data voor het tonen en afspelen van het aangeboden,
- een hoogwaardig kleurenscherm met hoge resolutie
- een voldoende groot werkgeheugen en
- een voldoende snelle netwerkkaart

Een configuratie die momenteel veel genoemd wordt is:

- \* een Pentium 200 MHz
- \* een hoogwaardig SVGA en XGA (1024x768, 256 kleuren) kleurenbeeldscherm van 15 inch of groter
- \* 32 Mb intern RAM-geheugen
- \* 2 Gb (voldoende snel) schijfgeheugen
- \* een CD-ROM 4-lezer
- \* een gecombineerde video-audio kaart MPEG-2, en
- \* een netwerkkaart van 100 Mbps

## 6.3 Aanverwante apparatuur en software

Vanuit een gebruikers perspectief ligt de meerwaarde van digitale informatie niet zozeer in de beschikbaarheid ervan maar in de mogelijkheden tot be- en verwerking. De gebruiker zal het vrij weinig interesseren of een bibliotheek met digitale beschikbaarstelling zijn voordeel zou kunnen doen. Waar het de gebruiker om gaat is, dat hij de eenmaal gevonden informatie op een gemakkelijke manier, kan raadplegen, printen, downloaden, bewerken, verwerken en opslaan en daarbij niet gehinderd wordt door incompatibele software en hardware. Voor een bibliotheek betekend dit dat men, wellicht centraal, zal moeten voorzien in:

- soft- en hardware voor mogelijke bewerking van het digitale materiaal (video-montage-tafels, foto-stylers etc.), en apparatuur voor
- de uitvoer van het digitale materiaal (kleurenprinters, projectors, dia-printers, etc.)

Daarnaast zal er op centraal niveau natuurlijk apparatuur aanwezig moeten zijn die zorgt voor:

- de omzetting van analoog naar digitaal materiaal (converters, scanners, videorecorders, digitale camera's, OCR-apparatuur, etc):
- de opslag van het digitale materiaal (servers, laserdisks, etc.) en
- een voldoende snel netwerk

## **1. Algemene organisatorische consequenties voor de UB**

### **1.1 Sterkere markt-oriëntatie**

Zoals in deel I gesteld werd, dwingen de vraag naar 'on-demand'-diensten en de toename in het aanbod van multi-mediaal materiaal, bibliotheken niet alleen tot de aanschaf en beheer van allerlei extra (rand-)apparatuur en software maar ook tot een grondige reorganisatie van de bestaande werkwijzen die op een efficiënte acquisitie, catalogisering en uitlending van een eigen gedrukte collectie gericht zijn. Voordat er op specifiekere consequenties voor de afdelingen binnen de UB ingegaan wordt, kunnen ook een aantal algemene consequenties genoemd worden.

Een eerste algemene consequentie van de geschetste ontwikkelingen is dat, ondanks de beperkte bandbreedte van het Internet, de gebruiker toch een relatieve onafhankelijkheid ten aanzien van de opslagplaats van de collectie zal krijgen. Een gevolg hiervan is dat bibliotheken een sterkere marktwerking zullen ondervinden die hun dwingt zoveel mogelijk materiaal tegen een zo gunstig mogelijke prijs te leveren. Volgens een aantal auteurs zou men van een traditionele taakorganisatie met input-financiering moeten overgaan tot een marktorganisatie, met betaling van activiteiten per transactie. Bij deze vercommercialisering zou ook horen dat de bibliotheken hun bestaan als digitaal archief bekend maken en dienen te vermelden welk materiaal en diensten aangeboden zullen worden.

### **1.2 Samenwerking met rekencentra**

Een tweede consequentie van de ontwikkelingen is dat de taken van bibliotheken en rekencentra korter bij elkaar komen te liggen omdat beiden op de beschikbaarheid van documenten aangesproken zullen worden. De noodzakelijke samenwerking tussen deze twee diensten wordt volgens Allen (1995) door een aantal factoren bemoeilijkt.

De eerste factoren die hij noemt en samenvat onder de term "clashes of organizational culture", betreffen de dienstengerichtheid en gebruikers-gecentreerdheid van bibliotheken tegenover de productgerichtheid en opdrachtgerichtheid van rekencentra. Allen (1995) citeert een onderzoek waaruit zou blijken dat de dienstengerichtheid van bibliotheken zijn oorsprong vindt in de notie van vrije toegang tot informatie. De productgerichtheid van rekencentra zou daarentegen samengaan met een commerciële inslag die informatie beschouwt als een goed dat beschermt en verkocht moeten worden. Een groter verschil met een opvatting over de waarde van informatie als iets dat breed verspreid en gebruikt zou moeten worden, zou volgens hem niet te vinden zijn. Een laatste verschil in organisatie-cultuur zou gelegen zijn de gebruikerscentreeerdheid van bibliotheken versus de opdrachtgecentreerdheid van rekencentra. Bibliotheken zouden doorgaans met de belangen van alle universitaire gebruikers rekening houden terwijl andere diensten sneller geneigd zouden zijn om alleen aan specifieke wensen van opdrachtgevers (faculteiten en onderzoekers) tegemoet te komen.

De strategie waarmee deze moeilijkheden overwonnen kunnen worden, wordt door Allen (1995, p.652) als volgt samengevat:

*"...., although the organizational cultures of the two units are different, this barrier can be overcome by appropriate organizational structures. Permeable boundaries, special administrative forums to deal with crucial issues of schedules and priorities, and enhanced communications mechanisms can allow units on campus with different organizational cultures to work together productively".*

Een derde factor die samenwerking zou bemoeilijken, wordt gevormd door de 'personal incompatibilities'. Deze zouden te maken hebben met het feit dat mensen vanwege hun onderscheiden karakters tot verschillend soort werk aangetrokken worden. Onderzoek hieromtrent wijst echter uit dat er geen duidelijke stereotypen te geven zijn en dat wel degelijk samenwerking tussen automatiseerders en bibliotheekmedewerkers mogelijk is. Ongeveer 1 op 10 bibliotheekmedewerkers past in het profiel van de automatiseerder. Het is de taak van het management deze mensen bij elkaar te brengen.



Een laatste factor die door Allen (1995) genoemd wordt, vormen de verschillen in reactie op veranderingen. Op dit punt wordt met name bibliotheken vaak verweten dat ze vrij passief zijn als het gaat om de inpassing van nieuwe mogelijkheden die de informatietechnologie biedt. De aangewezen strategie is volgens Allen (1995): 'organizational flexibility', welke zou moeten bestaan uit:

- de mogelijk tot herschikking van personeel voor nieuwe diensten en structuren
- mogelijkheden tot het ter discussie stellen van standaard procedures en productiemiddelen, en
- het gebruik van beloningsmechanismen voor de deelname aan, en initiatie van, innovatie-projecten.

Wellicht ten overvloede zij vermeld dat aan een aantal van de genoemde voorwaarden aan eigen instelling reeds voldaan wordt. Niet alleen participeren diverse medewerkers vanuit diensten in meerdere gezamenlijke projecten maar is er op management niveau ook regelmatig overleg. Gewezen zou ook kunnen worden op het feit dat de noodzakelijke markt oriëntatie van bibliotheken in het elektronische tijdperk de diensten automatisch al wat korter bij elkaar zal brengen.

### **1.3 Expertise ontwikkeling.**

Een laatste algemene consequentie die genoemd zou kunnen worden, is expertise-ontwikkeling. Alhoewel veel bibliotheken al voorzien in bijscholing in het gebruik van allerlei applicaties zal er in de naaste toekomst meer aandacht naar beeldverwerkings-apparatuur en -software moeten uitgaan. Niet alleen vanwege het eigen gebruik en onderwijs erin maar ook vanwege de samenwerking met de rekencentra die van oudsher een voorsprong hebben in het gebruik van nieuwere applicaties.

Een tweede gebied waarop extra expertise ontwikkeld zal moeten vormen een aantal commerciële vaardigheden. Meer dan in het verleden zal er aandacht gegeven moeten worden aan:

- markt-onderzoek (intern en extern) en inkoopstrategieën
- bemiddeling tussen rechthebbenden en eindgebruikers
- opstelling en zorg voor de naleving van licentie- en gebruiks-overeenkomsten
- kostenberekeningen met betrekking tot eigen opslag, digitalisatie en diverse migratie-strategieën.

## **2 Consequenties op acquisitie-niveau**

### **2.1 Kwaliteitsbeoordeling en verantwoordelijkheid van aanschaf**

In de huidige structuur ligt de verantwoordelijkheid voor de concrete bestellingen bij de faculteiten, c.q. de docenten, zelf. Hier zal met een toename in het elektronische aanbod weinig aan veranderen. De criteria voor aanschaf zullen immers ook voor elektronische documenten blijven bestaan. Het feit dat iets elektronisch aangeboden wordt verandert niets aan feit dat documenten op hun wetenschappelijke kwaliteit en inhoud beoordeeld moeten worden en hun aanschaf gezien moet worden in het licht van de collectionerings-gebieden en de toekomstige waarde van het document.

### **2.2 Attendering**

Een taak die in elektronische tijdperk, vanuit de bibliotheek, veel zwaarder aangezet zal moeten worden, is de attendering. Omdat er nog langere tijd sprake zal zijn van een parallel aanbod van papieren en elektronische versies, is het goed denkbaar dat de bestelverantwoordelijke geattendeerd zal moeten worden op het bestaan van varianten met elk hun respectievelijke voor- en nadelen. Met name bij elektronische versies zal er vanuit de bibliotheek gekeken moeten worden naar:

- de authenticiteit van het elektronische object (bibliografisch identificeerbare en geautoriseerde versies eventueel voorzien met digitale handtekeningen)
- de redundantie van het object (andere, gemakkelijkere, garandeerde

- en/of goedkopere vindplaatsen op het Internet)
- de toegankelijkheid in termen van beschikbare hard- en software, en
- mogelijke beperkingen voor de beschikbaarstelling, back-up's en conversies in auteursrechtelijke zin of voortvloeiend uit beperkende licentieovereenkomsten.

Bij de attendering moet niet alleen gedacht worden aan commercieel leverbaar materiaal. Minstens zo belangrijk, en zeker in het begin wanneer de database in een opbouwfase verkeerd, is de attendering op de mogelijkheid tot opname van analogo en digitaal materiaal dat aan de eigen instelling aanwezig is.

### **2.3 Samenwerking met inhoudsdeskundigen**

Zoals al in hoofdstuk 2 van Deel II gezegd, vraagt digitalisatie en het gebruik OCR-software om nauwe samenwerking met inhoudsdeskundigen. Zo zal dan ook door de bibliotheek, in samenspraak met interne of externe inhoudsdeskundigen en in het licht van de uniekheid, het gebruik en het doel van het materiaal, nagedacht moeten worden over mogelijkheden en noodzakelijke bewerkingen tot digitalisatie en/of OCR van veel gevraagd of uniek materiaal. Een bijzonder probleem vormen hier afbeeldingen. Niet alleen omdat bibliografische gegevens vaak ontbreken en alleen de maker bekend zijn, maar ook omdat bij de archivering van afbeeldingen en foto's een keuze gemaakt zal moeten worden tussen kwaliteit (document-grootte) en efficiënte opslag. Bij compressie vragen beeldresoluties en kleur-representatie dan ook bijzondere aandacht en is een verantwoorde keuze alleen te maken wanneer bekend is in hoeverre de inhoud, c.q. de uniekheid, van een document afhankelijk is:

- van de configuratie van de bits
- de structuur en het formaat van de representatie
- de ideeën die het document bevat, en
- voor welke doeleinden het document gemaakt is.

Te denken valt hierbij aan de verschillen tussen kunsthistorische documenten en medisch beeldmateriaal waar alsnog berekeningen en bewerkingen op mogelijk moeten zijn en fractale compressie de enige optie zou kunnen zijn.

### **2.4 Diversificatie van de markt**

Een ander gevolg van een toename in het elektronische aanbod is dat bibliotheken in de acquisitie-sfeer, naast de traditionele uitgevers, ook van doen krijgen met andere aanbieders die minder georganiseerd kunnen zijn, waaronder ook bijvoorbeeld docenten en onderzoekers. Hier zullen bibliotheken een efficiënte organisatie met duidelijke procedures en garanties tegenover moeten stellen willen ze als intermediar kunnen bemiddelen tussen de gebruikers en rechthebbenden. Omdat ze voor de feitelijke aanschaf van de producten verantwoordelijk zullen blijven ligt het voor hand dat er ook vanuit de feitelijke aanschaf gestreefd wordt naar een unificatie in de aan te leveren documentformaten alsook naar een uniformering van licentie- en gebruiks-overeenkomsten. Dit veronderstelt wederom een brede markt-oriëntatie, een zekere onderhandelingscapaciteit en technische expertise over de mogelijke documentconversies en beperkingen van beschikbare hard- en software.

## **3 Consequenties voor catalogisering**

### **3.1 De bibliografische ontsluiting**

Een van de belangrijkste taken van het bibliograferen in het elektronische tijdperk is gelegen in de vaststelling van de integriteit van een elektronisch document. Omdat vorm en inhoud bij deze documenten minder verweven zijn, zijn er meerdere criteria voor een eenduidige beschrijving noodzakelijk. Wil men een citeerbaar document overhouden dan zal de catalogus niet alleen de titel, auteur, uitgever of producent en het jaar van uitgifte van het document moeten vermelden, maar ook informatie moeten verschaffen over:

- de eventuele aard van het origineel (analoog of digitaal) en een verwij-

- zing hiernaar
- het originele formaat van het elektronische document waarmee de layout van het document (paginering, hoofdstukken etc.) vastgelegd werd
  - met welke viewers en hardware het document op een correcte wijze te raadplegen en te citeren is
  - de grootte van het document evenals de wijze van compressie en decompressie en eventuele conversie-mogelijkheden
  - het moment waarop het document in de database beschreven en elektronisch opgeslagen werd
  - bij verwijzingen naar externe databases het moment van beschrijving en de wijze en autorisatie van de toegang
  - de geschiedenis van het document, c.q. de voorafgaande vindplaatsen op diverse servers of tertiare opslagmedia zoals cd-rom's etc.
  - welke cryptografische technieken ter identificatie van een origineel gebruikt worden
  - welk soort gebruik voor welke toepassingen toegestaan is, en
  - de plaatsen waar eventuele toestemming voor anderssoortig gebruik te verkrijgen is

Wanneer deze noodzakelijk informatie niet met het document gegeven is dan zal het bibliograferen geen eenvoudige taak zijn en veel extra bibliografisch onderzoek vergen dat zich ook uitstrekt tot technische aspecten omtrent het formaat, viewers en andere technische vereisten.

Een apart probleem voor het bibliograferen vormen zogenaamde gelinkte objecten. Wanneer de integriteit, de heelheid en uniekheid van de documenten in de relaties gelegen is, dan zal men de gehele verzameling gelinkte objecten moeten kunnen archiveren en fixeren in de tijd. Procedures hiervoor moeten nog ontwikkeld worden. Hetzelfde geldt ook voor bibliografisch moeilijk identificeerbaar materiaal waar beslissingen genomen moeten worden over de kenmerken waaraan het document voor opname in database heeft te voldoen. Wanneer er bovendien een toekomstige koppeling van de lokale database met de database van Pica nagestreefd wordt, dan zal er tevens overleg met Pica noodzakelijk zijn over de te gebruiken kenmerkcodes.

### 3.2. De inhoudelijke ontsluiting

Vakreferenten worden door multi-media op zowel een inhoudelijk als op een meer technisch vlak geraakt.

De belangrijkste verandering op het inhoudelijke vlak betreft de uitbreiding ervan.

Bij beeld- en video-materiaal zal niet volstaan kunnen worden met het benoemen van het onderwerp en/of concept, c.q. het thema of abstract begrip dat de maker van het document heeft willen uitdrukken, maar zullen ook waarneembare objecten, personen, gebeurtenissen en lokaties benoemd moeten worden. Omdat beeld- en video-materiaal veelal interdisciplinair van aard is, zal men bij de beschrijvingen al snel tegen de vakinhoudelijke grenzen aanlopen en voor een goede inhoudelijke ontsluiting op samenwerking met andere inhoudsdeskundigen, vanuit de bibliotheek of faculteit, aangewezen zijn. Hetzelfde geldt voor de bepaling van relevante details voor scanning, conversie en/of content-based-retrieval. Hiertoe is niet alleen een zekere technische expertise vereist maar ook een bredere oriëntatie op meerdere kennisgebieden dan doorgaans bij teksten het geval is. Vakreferenten zullen in ieder geval moeten kunnen constateren dat het beeldmateriaal dat zij toevallig ter ontsluiting aangeboden krijgen vanuit meerdere perspectieven ontsloten dient te worden en welke deskundigheid hiervoor noodzakelijk is. Gedacht kan hier worden aan beeldmateriaal dat bijvoorbeeld vanwege zijn medisch-historische betekenis voor opname in aanmerking zou komen maar dat ook kunst-historische kenmerken bevat die een de opname in de database zou rechtvaardigen. Voor een goede inhoudelijke ontsluiting van beeldmateriaal is het dus zaak dat vakreferenten al vroeg bij de acquisitie van beelden betrokken worden zodat de reden tot opname, het bijzondere, van het document bekend is en er eventueel meerdere inhoudsdeskundigen ingeschakeld kunnen worden ter bepaling van de diepte van ontsluiting.

Een tweede punt waarop multi-media de vakreferenten zullen raken is de frequentie waarmee documenten op relevantie gescreend moeten worden. Vanwege de systeembelasting die sommige media met zich meebrengen, het beslag dat ze op de serverruimte kunnen leggen en/of geconverteerd moeten worden omdat de soft- of hard-ware veranderd, zullen elektronische documenten vaker op hun relevantie en daadwerkelijke gebruik gescreend moeten worden dan hun papieren tegenhangers. Aan de hand van feitelijke gebruiksgegevens, die het systeembeheer te leveren heeft, zal er een terugkoppeling met de bestelverantwoordelijke moeten plaatsvinden over de migratie naar een minder belastend opslagmedium en/of besloten moeten worden tot afvoer. Wanneer in het kader van de coördinatie van collectievorming afspraken lopen over de beschikbaarheid van het bestand dan zal tevens overleg met met betrokken instellingen geboden zijn.

Op het technologische vlak worden vakreferenten geraakt door de expertise die voor scanning, conversie en content-based-retrieval noodzakelijk is. Ze zullen de bestelverantwoordelijke in ieder geval moeten kunnen wijzen op de (on-)mogelijkheden van bepaalde documentformaten en duidelijk moeten kunnen maken in hoeverre beeldresoluties en kleur-representaties door compressie beïnvloed worden. Verantwoorde keuzes zijn immers alleen te maken wanneer, ook bekend is in hoeverre de inhoud van het document afhankelijk is van de configuratie van de bits, de structuur en het formaat van de representatie.

Naast de zojuiste geschetste taak hebben de vakreferenten ook een taak in bijhouden van de ontwikkelingen op het gebied van de automatische beeldontsluiting. Ze zouden de onderzoekers minimaal erop moeten kunnen wijzen dat deze mogelijkheden bestaan en, nog beter, welke methoden zich voor welk soort beeldmateriaal het beste eigenen.

### 4 Consequenties voor ISD (Informatiesystemen en Dienstverlening)

De afdeling die het meest direct door de 'on-demand'-diensten getroffen zal worden is de afdeling ISD en met name uitleen. Omdat de informatie geleverd moet worden op een wijze die de gebruiker het meest schikt, zou dit kunnen betekenen dat ze:

- van analoge originelen, digitale copieën moeten kunnen leveren
- van digitale originelen, voorzover mogelijk, analoge copieën moeten kunnen leveren
- over de noodzakelijke soft- en hardware daartoe moeten beschikken
- de noodzakelijke expertise in de bediening ervan moeten ontwikkelen
- de kosten op een efficiënte manier moeten kunnen doorberekenen, en
- verantwoordelijk gesteld kunnen worden voor de afdracht van de rechten die met deze handelingen gemoeid kunnen zijn.

Wanneer gebruikers deze conversies zelf zouden kunnen uitvoeren, dan zou ISD moeten zorgen voor:

- de autorisatie en verstrekking van passwords voor de toegang tot bestanden
- de berekening en afdracht van gelden die hiermee gemoeid zijn
- de opstelling en naleving van eventuele gebruiksverklaringen
- de terugkoppeling van informatie, in samenwerking met systeembeheer, over veel gevraagd materiaal dat voor digitalisatie en/of analoge copieën aanmerking zou kunnen komen
- de bemanning van een helpdesk waar men met vragen over het gebruik van, en problemen met, de apparatuur terecht zou kunnen.

## **5 Consequenties voor het systeem-beheer en -onderhoud**

### **5.1 Monitoring**

Een belangrijke, en nieuwe, taak in het elektronische tijdperk voor het systeembeheer zal de 'monitoring' worden. Zij zullen migratie-strategieën moeten ontwikkelen en aangeven wanneer het gebruik van een bestand aanleiding geeft tot migratie naar een ander snel of minder snel opslag-medium. In samenspraak met de inhoudsdeskundige, over de relevantie van het object, kan dan tot migratie besloten worden. Wanneer hiervoor conversies noodzakelijk zouden zijn, zal ook de integriteit van de document in overweging genomen moeten worden en zal ook overleg met de catalogus-afdeling geboden zijn. Betreft het bovendien uniek materiaal waar, in het kader van de coördinatie van collectievorming, afspraken met andere instellingen over gemaakt zijn, zal ook terugkoppeling naar de instellingen noodzakelijk zijn. De verantwoordelijkheid hierover zou bij de vakrefereren kunnen liggen.

Een ander beslispunt bij 'monitoring' vormt het moment waarop er besloten wordt tot acceptatie van nieuwe soft- of hard-ware. De systeemtechnische consequenties van deze vernieuwingen kunnen van dien aard zijn dat er een verdubbeling van de systemen noodzakelijk is waarmee de consequenties in de praktijk getest kunnen worden.

### **5.2 Back-up's**

Een andere kritische taak vormen de noodzakelijke systeem-back-up's. In het elektronische tijdperk zullen deze niet alleen frequenter uitgevoerd moeten worden maar ook omvattender van aard zijn. Een gevolg is dat, wanneer er iets mis zou gaan, de schade ook vele malen groter zal zijn en elektronisch uniek materiaal voor altijd verloren kan gaan. Er zal dan ook niet alleen nagedacht moeten worden over de back-up-mediums maar ook over 'fail-safe'-mechanismen die schade kunnen voorkomen.

### **5.3 Formaat, applicatie en conversie ondersteuning**

Omdat de 'on-demand'-diensten vrij ver kunnen gaan, zal er ook geanticipeerd moeten worden op de mogelijkheid dat er uitvoer-formaten gevraagd kunnen worden waar ingewikkelde conversies voor nodig zijn. De expertise hiertoe zal ontwikkeld moeten kunnen worden. Wanneer er tot een beperking van de ondersteuning besloten wordt, dan zal dit bekend gemaakt moeten worden. Ontwikkelingen op het gebied zullen echter gevolgd moeten worden evenals de technische ontwikkelingen op het gebied van de automatische beeld- en video-ontsluiting.

### **5.4 Database-ontwerp, -onderhoud en -beheer**

Omdat kant-en-klare multi-media-database-management-systemen nog niet te koop zijn, zal het systeembeheer niet alleen voor onderhoud en beheer van de database verantwoordelijk zijn (record-structuren, bijwerking van trefwoord-indexen, autorisaties, etc.) maar ook een belangrijke stem moeten hebben bij het ontwerp en definitie ervan. Hiertoe zullen ze zich ook moeten bekwamen in meer theoretische zaken zoals technieken voor effectiviteits-verhoging van grote databases (tree-indexing, etc.) en alternatieve algorithmes waarmee de 'precision' en 'recall' van 'content-based-image-retrieval'-systemen bepaald kunnen worden.<sup>23</sup> Daarnaast zal het systeembeheer een stem moeten hebben in de bepaling van de netwerktopologie en moeten kunnen aangeven waar, welk gebruik te verwachten valt en met welke netwerkbelastingen dit gepaard zal gaan. Een laatste gebied waar de taken voor het systeembeheer in het elektronische tijdperk uitgebreid zullen worden betreft de aanschaf, het beheer en het onderhoud van de randapparatuur voor de uitvoer van digitaal materiaal.

## **6 Consequenties voor het bibliotheekonderwijs**

Een laatste gebied waarop de multi-mediale ontwikkelingen de bibliotheken zullen raken betreft het bibliotheekonderwijs. Veel auteurs zijn van mening dat de taak zich in het elektronisch tijdperk niet mag beperken tot louter de aanschaf en beheer van allerlei extra (rand-)apparatuur en software maar zich ook tot het onderwijs hierin zal moeten uitstrekken. Volgens Melin (1996), Corral (1995), Vries & Minnigh (1995), White (1995), Richard (1995), Sever & Harel (1995) en Sever (1995) zouden niet alleen rekencentra, maar ook bibliotheken verantwoordelijk gesteld kunnen worden voor de overdracht van de expertise in het gebruik van deze apparatuur zodat de gebruiker ook zelfstandig in staat zal zijn, zijn informatie te zoeken, te vinden, te verkrijgen en te verwerken. Niet alleen een actieve rol m.b.t., de informatieverzorging over mogelijke informatiebronnen, informatiefilters, wijzen van catalogiseren en indexering zou hiertoe geboden zijn, maar ook een actieve onderwijs-rol van bibliotheken in het gebruik van apparatuur en software.

## **7 Mogelijke consequenties voor het onderwijs-leerproces**

---

<sup>23</sup>. Voor voorbeelden zij verwezen naar Furht e.a. (1995), p. 267.

Zoals in deel I al gesteld werd, zou beschikbaarstelling van multi-mediaal materiaal een flexibilisering van het (afstands-)onderwijs met zich mee kunnen brengen. Niet alleen de toegankelijkheid van beeldmateriaal wordt met een multi-media-database-management-systeem vergroot maar ook de mogelijkheden tot de ontwikkeling van computerondersteunde onderwijsvormen die qua effectiviteit sterk afhankelijk van de mogelijkheden waarmee beeld, tekst en geluid geïntegreerd kunnen worden. Een ander voordeel is dat met de beschikbaarstelling van beeldmateriaal ook docenten gaandeweg met het gebruik van beeldmateriaal vertrouwd zullen raken en sneller geneigd zullen zijn om zelf aan COO-programma's te werken of bij de ontwikkeling te assisteren. Een gevolg hiervan is weer dat docenten ook oog zullen krijgen voor andere didactische principes dan alleen doceren, en een rol zullen aannemen die meer bij student-gecentreerde onderwijsvormen aan zullen sluiten. Kortom: multi-media kunnen als hefboom fungeren bij de noodzakelijk geachte verandering van hoger onderwijs dat meer onderwijs-kwaliteitsbewust zou moeten worden en betekenis gerichte leerstijlen zou moeten ontwikkelen.<sup>24</sup>

## 8 Mogelijke consequenties voor het onderzoek

Over de mogelijke consequenties voor het onderzoek aan de instelling kunnen we vrij kort zijn. Niet alleen kunnen afbeeldingen die in het kader van de reguliere patientenzorg gemaakt zijn, vrij snel ter diagnose onder de specialisten verspreid worden, maar een uniforme opslag en ontsluiting maakt ook vergelijkend onderzoek mogelijk op een grotere verzameling beeld-materiaal. Bovendien laat elektronische opslag ook meerdere berekeningstechnieken toe. Naast het medisch onderzoek moet echter ook aan cultuurwetenschappelijk onderzoek gedacht worden. Zo zou het systeem voor FdCW een goed stuk gereedschap kunnen vormen voor de bestudering van allerlei soorten beeld en video materiaal in het kader van Visuele Cultuur. Een object-georiënteerde database-structuur biedt hiervoor voldoende functionaliteit. Voor een overzicht van het mogelijk gebruik van IT-voorzieningen binnen de humaniora zij verwezen naar van den Berg (1995).

## 9 Samenvatting

Wanneer men de organisatorische consequenties bekijkt dan is het meest opvallende dat, als er daadwerkelijk 'on-demand'-diensten geleverd moeten gaan worden, de bibliotheek tijd en ruimte vrij zal moeten maken voor de ontwikkeling van de noodzakelijke technische kennis. Dit is deste opmerkelijker wanneer men bedenkt dat in het elektronische tijdperk waarin de vorm en inhoud van een document loskoppeld is en waarin bibliotheken aanvankelijk een minimale rol toebedeeld kregen, het toch de bibliotheekmedewerkers zijn die voor de handhaving van de integriteit van het document, de vorm en inhoud weer bij elkaar moeten brengen en daartoe flink bijgeschoold moeten worden.

Omdat de bijscholing dient te gebeuren op een gebied waar de rekencentra traditoneel sterk in waren, zal ook tussen de diensten een taakverschuiving plaatsvinden. Het meest duidelijk is dit bij het bibliotheekonderwijs dat in het elektronische tijdperk ook het gebruik van beeldverwerking apparatuur zal moeten omvatten en een afstemming met de rekencentra veronderstelt omdat deze een parallel cursusaanbod zullen hebben. Afstemming tussen de diensten is ook nodig voor de realisering, het onderhoud en beheer van een werkbare infrastructuur. Beide diensten kunnen immers op beschikbaarheid van documenten aangesproken worden en zijn in deze volledig op elkaar aangewezen.

Een nauwere samenwerking is ook geboden met de faculteiten, c.q. docenten. De vakreferenten zullen al vrij vroeg bij de keuze tot opname van beeldmateriaal betrokken moeten worden, zowel voor de inhoudsontsluiting als voor de technische vereisten en specificaties van het materiaal, en voor de voorlichting omtrent de mogelijkheden die het database-management-systeem heeft.

Op afdelingsniveau gebeurt ook het een en ander. Zo zal de vakreferent, die van oudsher vrij zelfstandig op zijn eigen vakgebied kon opereren, voor de onderwerpsontsluiting veel korter bij de feitelijke acquisitie en verwerking van multi-media en de technische aspecten ervan betrokken moeten worden. Bovendien krijgt hij extra verantwoordelijkheden met betrekking tot de attendering en samenwerking met andere inhoudsdeskundigen. Voor de afdeling acquisitie betekent dit dat ze zich veel meer op de commerciële taak kan richten en verantwoordelijkheden krijgt bij de uniformering van aan te leveren documentformaten en bij de uniformering van licentie- gebruikers-overeenkomsten. Voor de catalogusafdeling betekenen

---

<sup>24</sup>. Voor een nadere beschrijving van student-gecentreerd hoger onderwijs zij verwezen naar Franssen (1993).

elektronische documenten alleen maar meer en moeilijker werk. Veel meta-informatie over documenten zal zelf opspoord moeten worden omdat de documenten anders helemaal niet bibliografisch te beschrijven zijn. Hetzelfde geldt voor de afdelingen 'ISD' (uitleen) en 'systeembeheer' (of zo men wil het bureau). Ook voor deze afdelingen betekenen elektronische documenten meer werk. Zowel in het bedrijfs-informatie-systeem alsook in administratie-sfeer waar veelvuldig met gebruikers afgerekend zal moeten gaan worden. Een gevolg van dit alles is dat deze afdelingen ook nauwer zullen moeten gaan samenwerken omdat ze nog sterker van elkaars informatie afhankelijk worden. Samenwerking en een goede communicatie zal ook geboden zijn tussen de vakreferenten, de catalogusafdeling en het systeembeheer. Deze zijn niet alleen vanwege de bedrijfsinformatie op elkaar aanwezen maar deze moeten ook op elkaars expertise kunnen terugvallen omdat beschikbaarheid van elektronische document zowel de bibliografische als de inhoudelijke ontsluiting omvat alsook de opslag en het transport over het netwerk.

Op managementniveau zal er rekening mee gehouden moeten worden dat de beschikbaarstelling van multi-media vrij diep in de organisatie zal ingrijpen en een aantal taken zullen ontstaan die nieuw zijn. Vrijwel niemand in de organisatie zal om de ontwikkeling van een zekere technische expertise heen kunnen komen en taken toebedeeld moeten krijgen die vrij nieuw zijn en of uitgediept moeten worden. Hiertoe zal ook de 'organizational flexibility' noodzakelijk zijn die Allen (1995) bij de samenwerking met de rekencentra voorstelt.



## IV Aanbevelingen voor het vervolgtraject

### 1 Fase I.

Gegeven het uiteindelijke doel van het project: "de daadwerkelijke implementatie van een multi-mediaal systeem dat aan onderscheiden wensen tegemoet kan komen en waarbij een landelijke geaccepteerde, bibliografische en inhoudelijke, ontsluiting gewaarborgd kan worden", en wanneer men:

- het huidige gebruik van AV-materialen relateert aan de slechte toegankelijkheid en beperkte mogelijkheden tot integratie in het onderwijs en onderzoek
- bedenkt, dat juist het gebruik van AV-materialen de leerstof aanschouwelijker zou kunnen maken en de kwaliteit van het onderwijs zou kunnen bevorderen (zie b.v. het Kwaliteitsmanagementsplan (1996) van de Universiteit van Maastricht),
- met een goed en centraal multi-media management-systeem, de in het Kwaliteitsmanagementsplan (1996) gewenste, flexibilisering van het (afstands-)onderwijs beter vormgegeven kan worden
- de andere argumenten tot anticipatie serieus neemt,

dan zal er een begin gemaakt moeten worden met de opname, de beschrijving en de beschikbaarstelling van multi-mediaal materiaal. Omdat er op tal van punten nog de nodige kennis en expertise ontbreken is wellicht een experimentele en gefaseerde opzet van het vervolgtraject de meest aangewezen weg. Ook wanneer niet meteen een centrale beschikbaarstelling nagestreefd wordt en men niet gehinderd wordt door de beperkte bandbreedte van het huidige netwerk, zou men bij de juiste keuze van op te nemen documenten in de experimenteerfase voldoende ervaring kunnen opdoen. De voordelen die zo'n bescheiden begin met zich meebrengt zijn de volgende:

- dat de noodzakelijke communicatie tussen de belanghebbenden bij de vulling van de database beperkt kan worden tot enkele gesprekspartners en zich zou kunnen toespitsen op een of twee vakgroepen
- dat de procedures voor de opname en beschrijving van documenten geleidelijk aangepast en toegespitst kunnen worden
- dat de knelpunten, op technisch, juridisch en organisatorisch vlak zichzelf zullen aandienen en vanuit een praktijk-situatie verholpen worden
- dat expertise ontwikkeling zich in eerste instantie zou kunnen beperken tot de direct betrokkenen
- dat de diepte van de organisatorische consequenties en noodzakelijk expertises op de afdelingen geleidelijk zullen blijken en bijgestuurd kunnen worden.

Een eerste aanbeveling voor het vervolgtraject zou dan ook zijn dat er begonnen wordt met een soort experimentele fase, waarin er voor de opname, de digitalisatie en de beschrijving voor een beperkt aantal soorten documenten gekozen wordt. Omdat de mogelijkheid tot een diepe inhoudelijke en bibliografische ontsluiting van beeldmateriaal, tot op vakgroep niveau, een van de uitgangspunten van het project is, zou men ook in de experimenteerfase moeten denken aan documenten waar niet zozeer tekst maar beelden op de voorgrond staan. Niet alleen omdat tekst-georiënteerde documenten, en video's als geheel, via het pica-systeem beschreven kunnen worden maar ook vanwege het feit dat er bij de verwerking van beeldmateriaal de grootste achterstand bij het hoger onderwijs bestaat en tekst-georiënteerde documenten niet de vragen oproepen waar in de toekomst een antwoord op gegeven zal moeten worden. Het zou jammer zijn wanneer men zich in de experimentele fase beperkt tot eenvoudige diatheek- of 'video-on-demand'-systemen die met relationele databases en file-servers vormgegeven kunnen worden. Vragen omtrent de ontsluiting van afbeeldingen en video-segmenten worden hiermee vooruit geschoven. Hoewel bij het huidige gebruik van beeldmateriaal een object-georiënteerde database niet direct noodzakelijk is, zou met een beperking tot een systeem waarbij afbeeldingen en video's alleen als geheel ontsloten hoeven te worden, een waardevolle onderzoeksmogelijkheid op beeldmateriaal uitgesloten worden. De keuze voor beeld- en/of video-materiaal waarbij beeld-elementen benoemd moeten worden, dwingt in ieder geval tot een nadere bezinning over:

- aan te schaffen digitaliseerapparatuur als gevolg van de document-analyses
- de soort en het merk server die voor de opslag van beeld-materiaal in aanmerking komt,
- welke database-software het beste te gebruiken is,
- de vereiste opslagstructuur van het gedigitaliseerd materiaal die van de database-software afhankelijk kan zijn
- de database record-structuur
- het mogelijk gebruik van 'content-based-image-retrieval'-technieken
- de diepte van de inhoudelijke ontsluiting van beeld-materiaal, en
- welke soort beeld-trefwoord-thesaurus voor de inhoudelijke ontsluiting het beste gebruikt kan worden en of deze zich zowel voor cultuurhistorisch als medisch beeldmateriaal eigenen of laten integreren.

Het gebruik van OCR-software in de experimenteerfase is niet aan te bevelen. De vereiste expertise is zo omvangrijk en specifiek dat dit beter uitgesteld kan worden tot een later stadium wanneer in de experimenteer-fase duidelijk geworden is wat we wel zelf kunnen.

Waar in de beginfase wel voor gezorgd moet worden is voldoende technische expertise met betrekking tot de noodzakelijke database-technieken en de automatische beeldinformatie-extractie-technieken.

## 2 Fase II

### 2.1 Maasnet-2

Wanneer in een vervolgfase de beschikbaarstelling niet beperkt wordt tot een experimentele 'stand-alone'-configuratie dan vormt de bekabeling van netwerk de meest kritische factor in het vervolgtraject. Minder kritisch is de randapparatuur met de bijhorende software die in verhouding minder kostbaar zijn, snel verouderen maar ook snel vervangen kunnen worden. Welk soort randapparatuur de voorkeur verdient en aangeschaft moet worden zal afhangen van het soort materiaal dat men aan de eindgebruiker zou willen aanbieden en de wijze waarop men hem het materiaal zou willen aanbieden. Waar men bij een centrale beschikbaarstelling over het netwerk niet omheen zal komen is de juiste netwerktechnologie.

In het rapport "MAASnet, Van toen naar straks" (ICA 1996), waarin de toekomstige netwerkvoorzieningen aan de eigen instelling geschetst worden, is men een vergelijkbare mening als Schenau (1995a) in hoofdstuk 4 van deel II, toegedaan. Ook in het Maasnet-rapport wordt er voor de backbone gekozen voor de 155 Mb/s ATM-technologie (met optische bekabeling) en voor een UTP-bekabeling voor de gebouwen. Zij het dat, vanuit kostenoverwegingen, vooralsnog alleen nieuwe gebouwen met UTP voorzien zullen worden.

Problematischer is dat, in het rapport voor de werkstations in de studenten- en cursuszalen van het nieuwe UB-gebouw (Grote Looiersstraat), alleen gesproken wordt over Ethernet-aansluitingen en traditionele functionele eisen, zoals WP, spreadsheets etc.. Het uitgangspunt voor de studie-zalen is:

*"...100 PC's die per 7 geclusterd een 10 Mbps poort op een 16-poorts switch krijgen",*  
(ICA 1996, p. 33).

Voor multi-media applicaties zal dit wat aan de krappe kant zijn. Bij een zeven-tal gebruikers van 10 Mbps zou gemiddeld 1.4 Mbps per gebruiker ter beschikking staan. Een gemiddelde dat voor digitale video niet toereikend is. Een mogelijk oplossing zou een lagere clustering vormen, maar verstandig is dit niet. Met het oog op "market-pull"-effecten zal men moeten overdimensioneren en zeker wanneer de vraag naar steeds hoger resoluties zal toenemen en men voor video, MPEG-2 zou moeten gaan gebruiken dat gemiddeld maar liefst 20 Mbps vraagt. De vraag dringt zich dan ook op of men voor de studentenzalen niet meteen voor Fast-Ethernet (100 Mb/s) zou moeten kiezen dat via UTP-bekabeling mogelijk is. Deze datasnelheid ondersteund niet alleen een bescheiden MPEG-2-gebruik, maar laat, bij 1.5 Mbps video, een 50-tal video-gebruikers toe. Een aantal dat bij het huidige videogebruik weliswaar hoog is, maar in de toekomst niet uitgesloten mag worden. Bovendien biedt een grotere bandbreedte van lokale netwerksegmenten meer vrijheid in de plaatsing van de video-server die bij een kleine bandbreedte vrij kort bij de werkstations geplaatst zal moeten worden. Hiermee zou meteen de locatiegebondenheid van video-gebruik, zoals dat met de AV-groepskamers het geval is, ondervangen kunnen worden. Een ander voordeel is dat, wanneer men video op meerdere plaatsen in de instelling ter beschikken zou willen

stellen, het aantal video-servers beperkt zou kunnen worden en meervoudige opslag van video-files voorkomen kunnen worden. Dit zou wederom weer een besparing van beheers- en onderhoudskosten betekenen.

Hoewel het momenteel nog onduidelijk is hoe snel op bepaalde ontwikkelingen gereageerd moet worden, is het wel duidelijk dat er na de experimentele fase een antwoord gegeven moet kunnen worden op vragen als:

- wat er op bepaalde plaatsen ter beschikking gesteld zou moeten kunnen worden
- welke netwerktopologie hier het beste voor geschikt zou zijn en,
- gegeven een bepaalde topologie, hoe zwaar de traditionele functionele eisen het netwerk feitelijk zullen gaan belasten.

In ieder geval zal, voor de bekabeling van het nieuwe gebouw en aanschaf van apparatuur, hierover met het rekencentrum overeenstemming verkregen moeten worden en een inschatting gemaakt moeten worden hoe snel de geschetste ontwikkelingen, beschreven in hoofdstuk 3 van Deel I, hun invloed zullen doen laten gelden.

## **2.2 Centrale organisatie en coördinatie**

Een laatste aanbeveling die voor de implementatie van een multi-media-management-systeem gedaan zou kunnen worden is: centrale organisatie en coördinatie. Er zullen niet alleen meerdere diensten moeten participeren maar de gehele instelling zal van de ontwikkeling op hoogte gebracht moeten worden. Dit is de enige manier waarop er voorkomen kan worden dat er te veel eigen initiatieven op instelling zullen ontstaan die de uitwisselbaarheid van gegevens in gevaar kunnen brengen. Bovendien hoeft niet iedereen weer het wiel uit te vinden. Een centrale organisatie en coördinatie is ook de enige manier waarop garandeerd kan worden dat:

- er een grotere hoeveelheid elektronische documenten voor langere tijd en voor grotere groepen voor onderzoek en onderwijs beschikbaar zal blijven
- er geen catalogi ontstaan waar geen feitelijke documenten aanhangen zoals dat vaak bij internet-directories het geval is
- er een grotere consistentie gewaarborgd kan worden bij de inhoudelijke en bibliografische ontsluiting van elektronische documenten dan in het geval dat iedereen zijn eigen database gaat opbouwen en via het netwerk beschikbaar zou stellen
- de hoeveelheid documentformaten en applicaties binnen de perken gehouden kan worden en ondersteuning gemakkelijker te organiseren en te geven zal zijn
- elektronische documenten ook snel toegankelijk zullen zijn omdat snelle machines en een snel netwerk alleen centraal gefinancierd kunnen worden en aanpassingen meerdere diensten zou kunnen treffen
- bibliotheken hun archief- en bewaar-functie voor elektronische documenten kunnen waarmaken en in het kader van de coördinatie van collectie-vorming bindende afspraken kunnen maken en het nationale bezit beter en veiliger geconserveerd kan worden
- het beheer en onderhoud van het multi-media-management-systeem goed geregeld wordt.

De feitelijke organisatie van een centraal multi-media-management-systeem zal niet gemakkelijk zijn. Meerdere diensten zullen mensen, tijd en geld moeten vrijmaken omdat niet alle werkzaamheden naast de reguliere taken uit te voeren zullen zijn. Op management-niveau zullen duidelijke afspraken en toezeggingen gedaan moeten worden over hoeveel mankracht er vanuit de verschillende diensten ingezet zullen worden en voor welk deel van de ontwikkeling en het onderhoud welke dienst verantwoordelijk zal zijn. Daarnaast speelt nog het probleem dat er meerdere projecten aan de instelling lopen die een directe relatie met het nog te ontwikkelen systeem hebben. Wil men dubbel werk en frustraties voorkomen dan zullen de resultaten van deze projecten op een centraal niveau bijeen gebracht en teruggekoppeld moeten

worden. Wellicht dat de sommige projecten in elkaar moeten opgaan of dat mensen tussen de projecten uitgewisseld moeten worden. Een ander probleem is dat de projecten een tijdelijk karakter hebben en nog niet duidelijk is hoe de ontwikkelingen op de vele gebieden op een structurele wijze het beste bijgehouden kunnen worden. Hier zal de belangrijkste taak van het management van de verschillende diensten liggen.

## Literatuurlijst

- Ahanger, G. & Little, T. (1996). *A survey of Technologies for Parsing and Indexing Digital Video*. In: Journal of Visual Communication and Image Representation, special issue on Digital Libraries, V 7, No 1, 1996, p. 28-43.
- Allen, B. (1995). *Academic Information Services: A Library Management Perspective*. In: *Library Trends*, V 4, No 4, p. 645-662.
- Ardizzone, E. & Cascia, M. (1997). *Automatic Video Database Indexing and Retrieval*. In: *Multimedia Tools and Application*, V4, 1997, p. 29-56.
- Barker, P. & Yeates, H. (1985). *Introducing Computer Assisted Learning*. New York.
- Baxter, G. & Anderson, D. (1995). *Image indexing and retrieval: some problems and proposed solutions*. In: *New Library World*, V 96, No 1123, 1995, p. 4-13.
- Beek, N. van der (1996). *Imaging in opkomst; technieken voor geautomatiseerde beeldopslag*. In: *Open*, V 28, No 7/8, 1996, p. 176-178.
- Berg, J. van den (1995). *Informatietechnologie in het onderzoek in de humaniora*. In: Schenau (1995b), p.101-131.
- Bide, M. (1996). *How can we create universally acceptable standards for electronic publishing?*. In: Neubauer (1996), p.94-101.
- Brichford, M. & Maher, W. (1996). *Archival Issues in Network Electronic Publications*. In: *Library Trends*, 43 (4) 1995, p. 701-712.
- Brown, D. (1996). *Electronic Publishing and Libraries. Planning for the Impact and Growth to 2003*. Bowker, London 1996.
- Clark, T. (1995). *On the Cost Differences between Publishing a Book in Paper and in the Electronic Medium*. In: *Library Resources & Technical Services*, 39 (1) 1995, p. 23-28.
- Collecties op Achterstand: buitenlandse wetenschappelijke literatuur in Nederlandse bibliotheken: bronnenboek*. Koninklijke Bibliotheek, Den Haag, 1996.
- Corral, S. (1995). *Information Specialists of the Future: Professional Development and Renewal*. In: Helal & Weiss (1995), p.1-11.
- Dahlmans, I.(1996). *De bekeekfrequenties van video's op de studielandschappen FdEW/FdCW*. Universiteitsbibliotheek UM.
- Dannelly, D. (1995). *Resource Sharing in the Electronic Era: Potentials and Paradoxes*. In: *Library Trends*, 43 (4) 1995, p. 663-678.
- Dempsey, L. & Heijne, M. (1995). *Netwerkinformatiesystemen in de steigers*. In: Schenau (1995b), p.175-206.
- Dempsey, L & Mumford, A. & Robiette, A & Rusbridge, C. (1996). *eLib standards guidelines*. URL: [http://ukoln.bath.ac.uk/elib/wk\\_papers/stand.html](http://ukoln.bath.ac.uk/elib/wk_papers/stand.html) (18-6-1996)
- Dieleman, E. & Schenau, A. van (1995). *IT-voorzieningen in 1995*. In: Schenau (1995b), p.1-28.
- Elderink, K. (1996). *Bibliotheken en Internet*. In: *Open*, V 28, No 4, 1996, p. 88-91.
- Encarnacao, J. & Foley, J. (eds.) (1994). *Multimedia: system architectures and applications*. Springer-Verlag Berlin Heidelberg, 1994.

- Flerackers, E. e.a. (1996). *Draft report "Technical design of the service (IMM-DB)"*. Expertise Centre for Digital Media Diepenbeek, 1996. Geschreven in het kader van het Electra-project.
- Franssen, J. (1994). *Informatie-Technologie in het Hoger Onderwijs*. Rapport in het kader van het IT-project UB/ICA/O&O. Maastricht, 1994.
- Furht, B. & Smoliar, S. & Zhang, H. (1995). *Video and image processing in multimedia systems*. Kluwer, Dordrecht 1995.
- Giles, M. (1996). *From Gutenberg to Gigabytes: Scholarly Communication in the Age of Cyberspace*. In: The Journal of Politics, 58 (3) 1996, p. 613-626.
- Girod, B. (1994). *Integration of Motion Video into Multimedia Computers*. In: Encarnação & Foley (1994), p.139-154.
- Guenther, R. (1996). *The Challenges of Electronic Texts in the Library: Bibliographic Control and Access*. In: Peek & Newby (1996), p.251-275.
- Graziadei, W. (1995). *Putting MultiMedia into the Teaching-Learning-Scholarship Environment*. URL: <http://wings.buffalo.edu/faculty/fact/cit95/papers/CIT95-PA.GRAZIA-1> (13-03-1996).
- Hanjalic, A. & Lagendijk R. & Biemond, J. (1995). *Achievements and challenges in visual search of video*. URL: [http://www.tudelft.nl/pda/smash/public/benelux\\_vs.html](http://www.tudelft.nl/pda/smash/public/benelux_vs.html) (7-1-97)
- Harnad, S. (1996). *Implementing Peer Review on the Net: Scientific Quality Control in Scholarly Electronic Journals*. In: Peek & Newby (1996), p.103-118.
- Hayes, B. (1996). *The Economic Quandary of the Network Publisher*. In: Peek & Newby (1996), p.108-112
- Hurtado, L. (1996). *A Consortium for Refereed Electronic Journals*. In: Peek & Newby (1996), p.201-213.
- Heijne, M. & Rexwinkel, R. (1995). *Laat honderd bloemen bloeien. Eindrapport SURFdoc: opslag, inzage en verwerking van elektronische documenten*. SURFnet bv, Houten, 1995.
- Helal, A. & Weiss, J. (ed.)(1995). *Information Superhighway: The Role of Librarians, Information Scientists, and Intermediaries*. Essen, 1995.
- Hjelsvold, R. & Langorgen, S. & Midstraum, R. & Sandstra, O. (1995). *Integrated Video Archive Tools*. URL: <http://www.idt.unit.no/~videodb/artikler/ACM-MM95/paper.html> (7-1-97).
- Hogenboom, J & Voort, J. van der (red.) (1994). *Voor de zoeker; handleiding voor het registreren en uitwisselen van gegevens over fotocollecties*. Den Haag: NBLC, 1994)
- Horik, R. van (1996). *Scanning en het conversieproces*. In: Open, V 28, No 2, 1996, p. 46-48.
- ICA (1996). *MAASnet, Van toen naar straks*. Dienst Informatie, Communicatie en Automatisering, Rijksuniversiteit Maastricht, 1 mei 1996.
- King, D. & Griffiths, J. (1995). *Economic Issues Concerning Electronic Publishing and Distribution of Scholarly Articles*. In: Library Trends, 43 (4) 1995, p. 713-740.
- Kwaliteitsmanagementplan*. Rijksuniversiteit Limburg, Maastricht, 8 maart 1996.
- Lancaster, F. (1995b). *Attitudes in Academia Toward Feasibility and Desirability of Networked Scholarly Publishing*. In: Library Trends, 43 (4) 1995, p. 741-752.
- Lee, J. & Li, Q. & Xiong, W. (1997). *VIMS: A Video Information Management System*. In: Multimedia Tools

and Applications, V4, 1997, p.7-28.

Leskien (1996). *Dienstleistungen der Bibliotheken bei Zugang und Distribution elektronischer Publikationen*. In Neubauer (1996), p. 46-54.

Luksch, P. & Schultheiss, G. (1996). *Authentisierung, Integrität und Standard-formate elektronischer Publikationen*. In: Neubauer (1996), p. 83-93.

Lynch, C. (1996). *Integrity Issues in Electronic Publishing*. In: Peek & Newby (1996), p.133-145.

Menil, C. (1996). *Die soziale Funktion der Bibliotheken im Zeitalter der elektronischen Publikationen*. In: Neubauer (1996), p. 125-133.

Mirande, M. (red.)(1994). *De kwaliteiten van computerondersteund onderwijs: voorbeelden uit het hoger onderwijs*. Bussum, 1994.

Neubauer, W. (1996). *Elektronisches Publizieren und Bibliotheken*. Frankfurt am Main, Klosterman, 1996.

Niemegeers, I. (1995). *Trends in de communicatienetwerken en -diensten*. In: Schenau (1995b), p.133-154.

Okerson, A. & O'Donnell (eds)(1995). *Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing; An Internet Discussion about Scientific and Scholarly Journals and Their Future*. Association of Research Libraries, Washington 1995.

Optibase (1996a). *Digital Video Primer, Pt.1*. URL: <http://www.optibase.com/primer.html> (13-6-1996).

Optibase (1996b). *MPEG Primer*. URL: <http://www.optibase.com/primer2.html> (13-6-1996)

Peek, R. (1996). *Scholarly Publishing, Facing the New Frontiers*. In: Peek & Newby (1996), p. 3-15

Peek, R. & Newby, G. (ed.)(1996). *Scholarly Publishing: the Electronic Frontier*. MIT, Cambridge, 1996.

Pica (1995). *Richtlijnen voor het catalogiseren van online resources*. DN 54/0895. Pica; Centrum voor Bibliotheekautomatisering, Leiden, 1995.

Pilot, A. (1994). *Potenties van de computer in het hoger onderwijs*. In: Mirande (red.) (1994), p.40-52.

Raghavan, V. & Gudivada, V. (1994). *Picture Retrieval Systems: A Unified Perspective and Research Issues*. URL: <http://www.usl.edu/~raghavan/raghavan.bib.html#Gudi94d> (13-6-1996).

Rawlins, G. (1992). *The New Publishing: Technology's Impact on the Publishing Industry over the Next Decade*. In: The Public-Access Computer Systems Review, V 3, No 8, p.5-63.

FTP: lib3@uhupvm1.uh.edu

Rawlins, G. (1993). *Publishing over the Next Decade*. In: Journal of the American Society for Information Science, V 44, No 8, p.474-479.

Roes, H. (1996). *Electronic Journals: A Short History and Recent Developments*. URL: [http://cwis.kub.nl/~dbi/users/roes/articles/ej\\_1996.htm](http://cwis.kub.nl/~dbi/users/roes/articles/ej_1996.htm) (15-4-97).

Rowe, L. & Boreczkey, J. & Eads, C. (1994). *Indexes for User Access to Large Video Databases*. URL: <http://www.bmrc.berkeley.edu/papers/VodsDB94.html> (7-9-96).

Schenau, A. van (1995a)(red.). *Trends en visie: informatietechnologie in het Hoger Onderwijs. Deel 1: Bestuurlijk document*. Rapport van de Wetenschappelijk Technische Raad SURF. Cramwinckel, Amsterdam.

Schenau, A. van (1995b)(red.). *Trends en visie: informatietechnologie in het Hoger Onderwijs. Deel 2:*

*Onderzoek en visie*. Rapport van de Wetenschappelijk Technische Raad SURF. Cramwinckel, Amsterdam.

Seaman, D. (1996). *The Electronic Text Center & On-Line Archive of Electronic Texts*. In: Neubauer (1996), p.55-57.

Sever, I. (1995). *Academic Library Users and Electronic Retrieval Systems*. In: Helal & Weiss (1995), p.383-394.

Sever, S. & Harel, C. (1995). *Managing the Virtual Library: Issues and Challenges*. In: Helal & Weiss (1995), p. 369-382.

Shen, K. & Delp, E. (1996). *A Fast Algorithm for Video Parsing using MPEG Compressed Sequences*. URL: <ftp.skynet.ecn.purdue.edu/pub/dist/delp/icip95-parsing> (30-9-1996)

Silverman, R. (1996). *The Impact of Electronic Publishing on the Academic Community*. In: Peek & Newby (1996), p. 55-69.

Starre, J. van der (1996). *Verraderlijk, er staat niet wat er staat; ontsluiting van afbeeldingen*. In: Open, V 28, No 3, 1996, p. 68-71.

SURFnet BV (1996). *"Premium". PProduktie van Elektronische Materialen op basis van Internationale en Uniforme Methodes*. Laatst bijgewerkt 6-2-1996.

URL: <http://www.nic.surfnet.nl/surfnet/projects/premium> (8-10-1996)

URL: <http://www.nic.surfnet.nl/surfnet/projects/premium/plan-nl.html> (8-10-1996)

URL: <http://www.nic.surfnet.nl/surfnet/projects/premium/rapport.html> (8-10-1996)

URL: <http://www.nic.surfnet.nl/surfnet/projects/premium/products/verantwoording.html> (8-10-1996)

*SURFdoc* (1995). Zie: Heijne & Rexwinkel (1995)

Task Force on Digital Archiving (1996). *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information*.

URL: <http://www.rlg.org/ArchTF/tfadi.index.htm#contents> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.execsum.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.intro.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.challeng.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.objects.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.randr.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.recommen.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.notes.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.reference.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.apx1.html> (25-6-1996)

URL: <http://www.rlg.org/ArchTF/tfadi.apx2.html> (25-6-1996)

Tenopir, C. (1995). *Authors and Readers: The Keys to Success or Failure for Electronic Publishing*. In: Library Trends, 43 (4) 1995, p. 571-591.

Tuck, B. (1996). *Document Ordering and Delivery Systems in Europe: Projects of the European Commission, services, conditions and prices*. In: Neubauer (1996), p. 58-69.

Turner, J. (1994). *Indexing film and video images for storage and retrieval*. In: Information Services & Use, V 14, 1994, p. 225-236.

Vorst, G. van der & Looijen, M. (1995). *Opbouw van een facultaire iT-infrastructuur*. In: Schenau (1995b), p.231-257.

Vries, J. de & Minnigh, L. (1995). *Information Specialists in the Future. Academic Library: Flexible Tightrope Walkers*. In: Helal & Weiss (1995), p. 13-26.



Werkgroep AV-media (1996). *Toekomst AV-media*. Geschreven in opdracht van de Bibliotheekcommissie van de Faculteit der Rechten UM.

White, H. (1996). *The Role of Information Intermediaries and the Superhighway: Crucial, Important, Trivial, or Non-Existent?* In: Peek & Newby (1996), p. 47-60.

Wiel, T. van de & Vennix, W. (1994). *Imagesproject 1994. Eindrapportage*. URL: <http://www.nic.surfnet.nl/surfnet/projects/surfdoc/imagesproject/docs/Eindrapport-hse.txt>

Wolf, H. de (1995). *Informatietechnologie in het Hoger Onderwijs*. In: Schenau (1995b), p.29-52.

