

### I N S T A P

Dit artikel geeft een overzicht van moderne methodes om informatie op te sporen over een bepaald onderwerp, ter ondersteuning van management, besluitvorming, wetenschappelijk onderzoek...

#### HET BEGRIP 'INFORMATIE'

##### Informatie in onze wereld

Onze steeds complexere wereld wordt gekenmerkt door internationalisering en globalisering. De wetenschappelijke, technische vooruitgang en de maatschappelijke evolutie in het algemeen gaan snel en blijven zelfs versnellen. Economische producten zijn nu minder gebaseerd op natuurlijke bronnen en steeds meer op informatie en kennis. Ons antwoord op deze situatie vereist dan ook minstens meer kennis en meer vaardigheden, een toenemend aanpassingsvermogen en flexibiliteit en mobiliteit. De exploitatie van informatie en opleiding is in een dergelijke wereld belangrijker dan ooit.

##### Enkele bizarre eigenschappen van 'informatie'

Informatie verschilt erg van andere economische producten en bouwstenen van onze maatschappij.

Informatie geraakt nooit verbruikt en bederft niet. Informatie veroudert echter wel; de snelheid waartegen ze wordt geleverd kan daarom van cruciaal belang zijn. De context is belangrijk. Er bestaat geen genormeerde

maat van een informatie-eenheid. De prijs van informatie is niet gekoppeld aan de waarde ervan in een bepaalde situatie. Bovendien is het moeilijk om de kwaliteit en de waarde van informatie te schatten. Een gegeven kan voor verschillende mensen op hetzelfde moment ter beschikking worden gesteld. Informatie kan wel opnieuw worden aangemaakt, waardoor ze goedkoop wordt voor grootschalig gebruik. Niettegenstaande kunnen de copyrightkosten de prijs wel hoog houden.

##### Vormen van informatiebronnen

Informatie kan informeel worden vervaardigd door menselijke communicatie via een (telefoon)gesprek, een fax, een elektronisch postbericht..., of meer formeel worden opgeslagen in en verspreid via documenten zoals een afdruk of computerbestanden. Beide vormen van informatiestromen zijn belangrijk. We concentreren ons hier op informatie die in documenten wordt opgeslagen.

##### Criteria om de kwaliteit van informatiebronnen te evalueren

Gezien de sterk variërende kwaliteit van informatiebronnen enerzijds en de kosten die gepaard gaan met het gebruik van informatie anderzijds, moeten wij ons altijd kritisch opstellen en ons een aantal vragen stellen:

- is dit gegeven waardevol, betrouwbaar, origineel, authentiek, nauw-

In de rubriek 'Instap' worden technieken die relevant zijn voor het vakgebied van archivarissen, bibliothecarissen en documentalisten op een introducerende manier voorgesteld.

De voorliggende 'Instap' toont u de wegen om informatie op het spoor te komen.

keurig, correct, uniek, volledig, begrijpbaar, substantieel, objectief en actueel?

- bevat de bron veel primaire informatie die niet uit andere bronnen kan worden verkregen?
- heeft deze informatiebron een grote reikwijdte?
- biedt ze informatie in een duidelijke vorm en lay-out?
- wordt de informatie aangereikt door een gebruiksvriendelijk informatiesysteem? Kan de gebruiker zich goed binnen de informatiebron oriënteren en kan hij er gemakkelijk zijn weg in vinden? Geeft het informatiesysteem voldoende ondersteuning

voor de gebruiker, is het klantvriendelijk?

- is het verspreidingsmedium, zoals afdrukken, e-mail, on line verspreiding, het gekozen bestandsformaat... optimaal voor het beoogde gebruik?

### De stroom van documentaire informatie

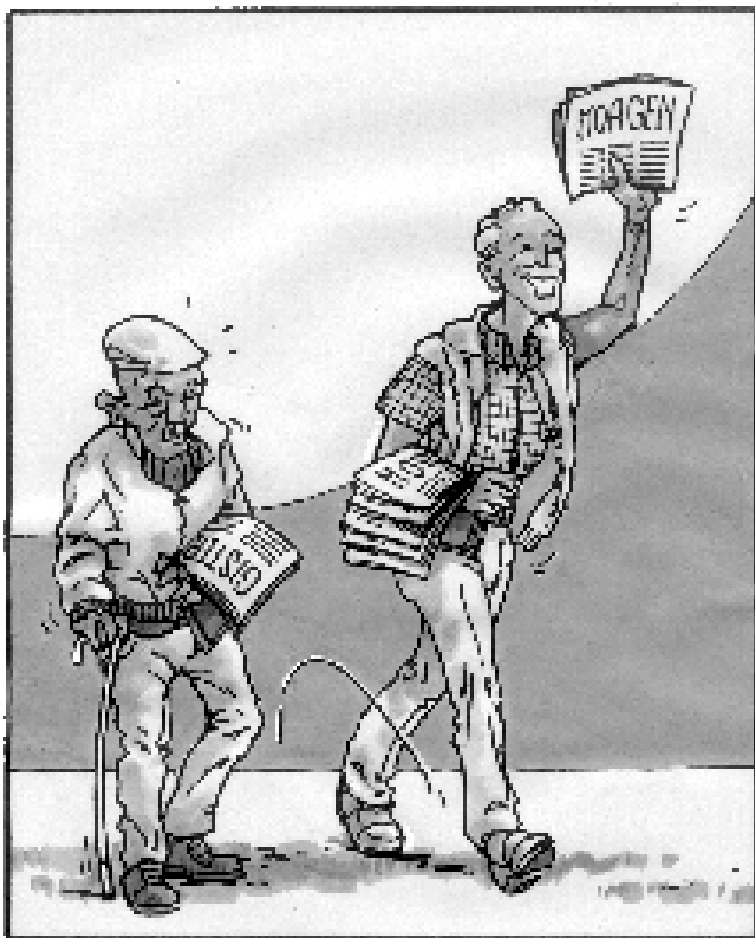
Auteurs zenden informatie naar lezers en gebruikers via verschillende kanalen en media. We kunnen een onderscheid maken tussen de primaire informatiestroom met primaire informatiebronnen, zoals wetenschappelijke artikels, en de secundaire stroom met secundaire bronnen zoals abstracttijdschriften, catalogi en bibliografische databanken. De secundaire informatiestroom is gebaseerd op de primaire stroom en is ontstaan omdat de grote

omvang van de primaire informatie de kans verlaagt om het geschikte informatiegegeven te vinden en te gebruiken. De secundaire informatie probeert enige orde te scheppen in de chaos.

### Informatiebronnen: de categorieën

Informatie kan op diverse manieren in categorieën worden ondergebracht:

- primaire en secundaire;
- tekst, beeld, geluid, software, gegevens (of een mengeling);
- hard copy (niet digitaal) ofwel digitaal bestand, computerondersteund, een computersysteem vereisend; de digitale bronnen kunnen op hun beurt verder worden onderverdeeld in off line (bv. informatie op cd-rom) en on line (bv. een databank die via het internet kan worden geraadpleegd).



INFORMATIE VEROUDERT SNEL

### Retrospectief vragen versus attentering

Om geschikte informatie op te sporen kunnen we een retrospectieve zoektocht uitvoeren, voornamelijk wanneer we nieuw, onbekend terrein betreden of een onderwerp aansnijden dat achtergrondinformatie vereist. Zodra we voldoende informatie hebben gevonden, staan we voor een nieuwe uitdaging, met name de voortdurende stroom van nieuw gegenereerde informatie en de veranderlijke wereld waarin wij leven en werken. Daarom willen mensen ook nieuwere informatie. Om zulke nieuwe bronnen te vinden, kunnen we een beroep doen op een attenteringdienst. Dat is een dienst die de ontvanger op de hoogte houdt van de laatste ontwikkelingen in zijn vakgebied of zijn interessegebied.

### De evolutie van gegevensdragers en distributie

De media die gebruikt worden om informatie te bewaren, te verspreiden

en op te zoeken blijven evolueren. Sommige mijlpalen hierin zijn de uitvinding van het drukken met behulp van verplaatsbare en hergebruikbare letters in de 15de eeuw, de eerste openbare, on line toegankelijke databanken in 1975, het ontstaan van het internet in de jaren 1970, de beschikbaarheid van cd-roms sinds 1985 en het ontstaan van het world wide web (www) rond 1990.

### Eindgebruikers versus informatiemakelaars

Mensen kunnen zelf rechtstreeks, als 'eindgebruikers', informatie terugvinden. Het informatielandschap is echter complex. Het kan dan ook veel tijd vergen om de juiste informatie te vinden. Verder kan het een dure zaak worden om informatie te zoeken en die dan in een geschikte vorm en layout te gieten zodat ze klaar is voor 'gebruik'. Daarom kan het belangrijk zijn om bijgestaan te worden door een tussenpersoon met expertise in dit domein, zoals een bibliothecaris of een informatiemakelaar. Hun diensten kunnen erg verschillen inzake expertise en prijs.

### Een vergelijking van enkele publicatiemedia

Papieren versies, cd-roms en bronnen die on line kunnen worden geraadpleegd, zijn drie voorbeelden van belangrijke informatiedragers en publicatiemedia. Belangrijke criteria om de kwaliteit en de geschiktheid van informatiemedia te beoordelen, zijn de hoeveelheid informatie waarmee ze efficiënt functioneren en de snelheid waartegen ze informatie kunnen actualiseren. Beide scores laag voor papieren versies, matig voor cd-roms en andere optische disks en het hoogst voor on line informatiesystemen. Elektronische informatiebronnen kunnen goedkoper worden aangeemaakt, getransfereerd en opgeslagen. Ze bieden betere zoekfuncties, verscheidene opmaken, en 'knippen en

plakken' naar andere documenten is snel. Samen zorgen deze eigenschappen ervoor dat grote, volumineuze documenten of databanken op een efficiëntere manier toegankelijk zijn. Verder kunnen elektronische media multimedia en hypermedia-inhouden aanreiken, zoals animatie, video, statische en dynamische virtuele realiteit (in plaats van enkel van opmaak voorziene teksten, cijfers en grafieken); ze kunnen 'actieve inhouden' aanreiken, dit wil zeggen verrijkt met begeleidende of zelfs ingebedde programma's om gegevens en informatieve inhouden te bekijken, te manipuleren, te sturen, te bestellen en te selecteren.

### Omschakeling van informatiemedia

Verscheidene informatiemedia zijn nu al op computers en digitale netwerken gebaseerd: cd-rom, dvd, hypermedia, het op afstand inloggen bij computers, het overbrengen van bestanden van en naar een computer via een netwerk, elektronische post, Usenet, het www... Andere, klassiekere media evolueren nu ook al in de richting van computer- en netwerkgestuurde technologie: telefoon, radio, TV, video, fax, tijdschriften, boeken...

### Elementaire problemen rond toegang tot informatie

Sommige fundamentele, elementaire moeilijkheden en problemen belemmeren de toegankelijkheid van informatiebronnen:

- zelfs als iemand al gebruiker is van een informatiezoeksysteem, is het soms niet helemaal duidelijk welke informatie er in feite nodig is;
- in een vraag die afgestemd is op een computersysteem kan men niet altijd duidelijk omschrijven welke informatie precies nodig is. Eén van de redenen daarvan is dat de context van de omschrijving van de informatie waarnaar men op zoek is, idealiter volledig zou moeten worden geformuleerd samen met de vraag, inclu-

sief de kennis en de achtergrond van diegene die informatie zoekt;

- ook al zijn computersystemen kunstmatig, toch gebruiken de meeste menselijke taal in hun interface met gebruikers; dat is bijvoorbeeld het geval voor de zoeksystemen van veel databanken. Dit kan problemen opleveren in verband met taal en woordenschat in het bijzonder. Mensen gebruiken verschillende talen en verschillende termen (woordenschat) om een gelijksoortig begrip te beschrijven. Begrippen, woordenschat en de betekenis van woorden en termen kunnen doorheen de tijd veranderen. De betekenis is soms ook afhankelijk van de context;
- informatie is gevat, ingesloten, verborgen in verschillende en onvolmaakte zoeksystemen. Om de informatie die in principe voorhanden is terug te vinden en er toegang toe te krijgen, moet de gebruiker dus over die verschillende zoeksystemen beschikken en er ook mee kunnen werken. Verder bestaat er (nog) geen volmaakte software inzake informatiezoeksysteem; wel is er sinds 1970 een snelle wetenschappelijke en technologische evolutie op het vlak van software in dit domein;
- het overaanbod van informatie: de gebruiker staat vaak versteld van de hoeveelheid beschikbare informatie en de grote toevoer van nieuwe informatie. Het is geen sinecure om hiermee permanent om te gaan;
- de prijzen van informatiebronnen: veel informatie is alleen mits betaling aan de producent en/of verdeler toegankelijk.

### Grasduinen en zoeken: basismethodes om informatie te vinden

De producent van een informatiesysteem kan twee manieren aanreiken om de gebruiker de juiste informatie in het systeem te laten opzoeken: grasduin- en zoeksystemen:

- bij grasduinen (of bladeren) volgt de gebruiker één of meerdere door

het systeem aangeboden zoekwe-  
gen. De informatie wordt bijvoor-  
beeld per onderwerp gerangschikt.  
De gebruiker hoeft zijn vraag niet  
in eigen woorden te formuleren om  
uit te drukken welke informatie hij  
zoekt. Om het organiseren en het  
opzoeken van informatieve gege-  
vens te ondersteunen, wordt door-  
gaans enige vorm van classificatie  
gebruikt. Voorbeelden daarvan zijn  
de inhoudstafel vooraan in een  
boek, enkele boeken die volgens  
een bepaald classificatiesysteem op  
rekken worden geplaatst, een direc-  
tory die hypertexten hiërarchisch  
ordent op het www, of hypermedia-  
systemen in het algemeen;

- bij zoeksystemen moet de gebrui-  
ker door een vraag specificeren naar  
welke informatie hij op zoek is,  
waarbij hij gewoonlijk een natuur-  
lijke of in sommige gevallen een  
meer formele taal gebruikt. In deze  
systemen is de informatie niet geor-  
dend volgens een bepaalde logica,  
maar gestructureerd als een compi-  
latie van items van een gelijksoor-  
tige vorm, zoals de gegevens van  
een elektronische databank. In veel  
gevallen wordt er een lijst van tref-  
woorden of een thesaurus gebruikt  
om moeilijkheden te vermijden die  
het gevolg kunnen zijn van het  
gebruik van een natuurlijke taal.  
Voorbeelden van dergelijke zoeksys-  
temen zijn de index (het register)  
achteraan in een boek, een cata-  
logus van een bibliotheek of een  
museum met een zoekinterface, en  
een zoekformulier op een website.

---

## INFORMATIE OP CD-ROM

### De technologie van de cd-rom

Gegevens kunnen digitaal worden  
bewaard op optische schijven zoals  
cd-roms. 'Cd-rom' is de afkorting

van Compact Disk-Read Only Memory.  
Vorm en structuur zijn afgeleid van de  
audio-cd. In 1985 verschenen de eerste  
systemen op de markt. De fysieke  
afmetingen zijn genormeerd: de dia-  
meter is 120 mm en in het midden is er  
een opening van 15 mm. Dit medium  
biedt directe toegang tot gegevens  
zoals andere schijven voor computers;  
de toegang is dus snel in vergelijking  
met media die sequentieel worden  
gelezen zoals cassettes. De gegevens  
worden in een drive gelezen door een  
laserstraal. De leeseenheid raakt het  
oppervlak van de schijf niet, waardoor  
er geen wrijving is en bijgevolg ook  
geen slijtage of beschadiging.

De geheugencapaciteit van een cd-rom  
is ongeveer 600 MB, 600 000 KB.  
Ter vergelijking: het gemiddelde blad  
papier kan informatie (gedrukte letter-  
tekens) bevatten die overeenkomt met  
2 KB geheugenruimte op een com-  
puter. Een cd kan dan ook ongeveer  
dezelfde hoeveelheid tekst bevatten  
als 300 000 A4-vellen. Duizend vellen  
papier vormen een stapel van onge-  
veer 10 cm. Eén cd komt dan ook  
overeen met een stapel van ongeveer  
30 meter A4-vellen papier, een stapel  
die zo hoog is als een gebouw van tien  
verdiepingen.

Cd-roms bieden heel wat voordelen ver-  
geleken met andere informatiedragers:

- de structuur is goed genormeerd en  
de technologie is stabiel; dit zorgt  
voor een hoge compatibiliteit;
- de informatiedichtheid is groot;
- de kosten van het opslaan van infor-  
matie per informatie-eenheid zijn  
laag;
- cd-roms kunnen gemakkelijk worden  
bewaard, vervoerd en verzonden  
met de post;
- bij normaal gebruik crasht de lezer-  
kop niet en de schijven zijn goed  
bestand tegen slijtage;
- er is directe en dus snelle toegang  
tot de informatie mogelijk;
- cd-romsystemen zijn gebruiksvrien-  
delijk.

Rond 1997 werd de dvd aangekon-  
digd. Dit medium biedt een hogere

geheugencapaciteit dan cd-rom. Dvd-  
spelers (lezers) kunnen ook cd's lezen.

### Op cd-rom gepubliceerde informatie

Cd-roms worden gebruikt om infor-  
matie te publiceren, kleinschalig te ver-  
spreiden, en op te slaan voor persoon-  
lijk gebruik. Er zijn al duizenden titels  
verschenen via dit medium. Hoe kunt  
u te weten komen of een cd-rom is  
gepubliceerd met relevante informatie  
voor u? Er bestaan overzichten, lijsten  
van cd-rompublicaties, bijvoorbeeld de  
*Gale Directory*, die in afgedrukte vorm  
en als doorzoekbare databank beschik-  
baar is; dit overzicht kan echter niet  
gratis worden gebruikt.

### De evolutie van cd-rom

We merken:

- een stijgend aantal gepubliceerde cd's;
- een verbetering van de software om  
informatie te lezen en te gebruiken  
en om met cd's te werken;
- een verhoogde toegang en trans-  
missiesnelheid van de drives (lezer-  
koppen, spelers);
- een stijgend aantal gepubliceerde  
titels die niet alleen eenvoudige tek-  
sten opleveren, maar ook multime-  
dia-inhoud;
- een dalende prijs voor de vereiste  
hardware;
- een grotere schijfruimte (van cd naar  
dvd);
- een toenemende in-huisproductie van  
cd's voor commercieel of persoonlijk  
gebruik, om informatie te bewaren.

---

## MULTIMEDIA EN HYPERMEDIA

### Multimedia

De toenemende mogelijkheden van  
computers (processorsnelheid, opslag-  
ruimte, grafische verwerking...) heb-  
ben ertoe geleid dat toepassingen niet  
alleen meer met tekst en cijfers werken  
maar ook met beeld, klank en zelfs van  
klank voorziene animatiefilmpjes en

video; dit alles valt onder de noemer multimedia.

## Hypermedia

Sinds het einde van de jaren 1980 worden de mogelijkheden van computers hoe langer hoe meer ingezet om verwijzingen tot stand te brengen tussen informatieve gegevens, en de gebruiker tot die informatie toegang te verlenen door hem te laten bladeren en vooraf ingestelde verwijzingen te laten volgen. Deze methode om toegang te verkrijgen tot informatie is niet zo vanzelfsprekend bij klassieker media (zoals gedrukte teksten). In plaats van het heel algemene woord 'documenten' worden de woorden 'hypertekst' en 'hypermedia' gebruikt om naar informatie te verwijzen die op die manier is georganiseerd.

## COMPUTERNETWERKEN EN HET INTERNET

Dezelfde soorten van data (informatie) die op een computer kunnen worden opgeslagen en beheerd, kunnen ook via een netwerk naar één of meerdere computers worden verstuurd. Die netwerken vormen een aanzienlijke uitbreiding van computers.

De voornaamste toepassingen van datanetwerken zijn het online gebruik van de processorsnelheid van een remote computer en – wat relevanter is in de context van dit artikel – de online toegang tot informatiebronnen. Concreet betekent dit dat deze netwerken online toegang verlenen tot catalogi van uitgeverijen, bibliotheken, boekenwinkels, informatiesystemen van een campus en de overheid, databanken (tekst of multimedia), op het netwerk gebaseerde tijdschriften en groepsdiscussies...

Elektronische netwerken lijden chronisch aan een te lage transmissiesnelheid en een te grote technische complexiteit, ondanks hun snelle evolutie. Een datanetwerk is sinds de jaren

1980 heel populair geworden en sterk gegroeid: het internet. Dit is een netwerk dat uit kleinere netwerken bestaat. Het is de wereldwijde aaneenschakeling van domeinen die lokaal met elkaar verbonden zijn, LAN en WAN (local area network en wide area network), die een welbepaalde verzameling (serie) normen (standaarden) hanteert, zogenaamde protocollen, om communicatie mogelijk te maken. Deze serie is de TCP/IP-suite van datacommunicatieprotocollen. Het internet verbindt niet enkel computers die door een zelfde fabrikant zijn gemaakt, maar ook verscheidene types van computers. Het beeld van een snelweg wordt soms aangehaald om het opkomende systeem van netwerkcomputers te beschrijven. Het internet heeft geen eigenaar en wordt niet door één bepaalde organisatie beheerd. Naast de basis-TCP/IP-protocollen gebruikt het internet ook veel andere vastgelegde normen om wereldwijde toepassingen (diensten) van het netwerk met verschillende doeleinden mogelijk te maken. Voorbeelden hiervan zijn:

- het verzenden van e-mails (met onder andere het smtp-protocol);
- het remote inloggen en gebruiken van de computer (het telnet-protocol);
- het overdragen van bestanden naar en van een computer op afstand (het ftp-protocol);
- het vragen en krijgen van informatie in de vorm van hypermedia (het http-protocol).

Om gebruik te kunnen maken van het internet moet men aangesloten zijn bij één van de vele Internet Service Providers (ISP).



KWANTITEIT, MAAR OOK KWALITEIT?

## HET WORLD WIDE WEB

Het world wide web is een dienst, een toepassing van het internet; het is gebaseerd op de infrastructuur van het internet. Het www is dus jonger dan het internet. Het concept van het www zag het licht aan het einde van de jaren 1980, lang nadat het internet werd opgericht. Het www legt hypertekst/hypermediaverbindingen tussen om het even welke documenten op het netwerk. Een document kan verwijzingen naar documenten van andere servers bevatten. Elke verwijzing zal het gerelateerde document opleveren, om het even waar het zich bevindt. De verwijzingen onderaan een blad kunnen bijvoorbeeld hypertekstverbindingen hebben met de feitelijke documenten die elders worden bewaard. Om toegang te krijgen tot het www moet je een browser installeren. Die browser leest de documenten en kan documenten uit andere bronnen halen. Informatieverstrekken richten hypermedia servers op en browsers kunnen documenten daar vandaan halen. De browser kan hypertekstdocumenten laten zien. Een hypertekst is een tekst met verwijzingen naar andere teksten. De browser maakt het

mogelijk op een transparante manier met de verwijzingen te werken: je selecteert de verwijzing en het document waarnaar verwezen werd, verschijnt op je scherm. Browsers zijn voor verschillende computerplatformen beschikbaar. Bij de totstandkoming van dit artikel waren Microsoft Internet Explorer en Netscape de populairste programma's voor toegang tot het www.

Het www boekte bijzonder veel succes in de jaren 1990. Het heeft het internet populair gemaakt. De programma's die toegang verlenen tot het www, de browsers, hebben veel functies overgenomen om ook andere door het internet geboden diensten te exploiteren. Daardoor is het verschil tussen het internet en het www voor veel gebruikers vervaagd. Het www is heel snel het populairste medium geworden dat toegang verleent tot informatie die wordt bewaard op verscheidene andere computers die op een netwerk aangesloten zijn.

Veel aspecten hebben bijgedragen tot dit succes:

- dankzij de client/server-architectuur die op het www wordt gehanteerd, kan de gebruiker vrij een client/server-programma kiezen;
- client-programma's (de browsers) zijn gratis;
- het www is toegankelijk via computers met de meeste besturingssystemen;
- het www biedt verscheidene complexe diensten aan via een gebruiksvriendelijke interface, die gebaseerd is op hypertext en hypermedia;
- de multimedia-inhoud is aantrekkelijk;
- de multimedia-inhoud is gemakkelijk toegankelijk, doordat browsers (in samenspraak met andere componenten van de computer) voor veel populaire bestandsformaten overvie-

wersoftware beschikken;

- het www maakt het mogelijk om op een transparante manier van de ene server naar de andere te gaan (bladeren, navigeren, surfen);
- de gebruiker kan relevante informatiebronnen vinden doordat er indexen en thematische overzichten gecreëerd zijn;
- veel informatie die met het www kan worden opgezocht is gratis;
- het www is vrij goed genormeerd (gestandaardiseerd), maar het wordt nog steeds verder ontwikkeld en verbeterd;
- het www is gebaseerd op een open, gepubliceerde norm/protocol voor communicatie van data ('http' genoemd);
- enkele goede http-server softwarepakketten zijn gratis beschikbaar;
- sommige programma's om informatiepagina's aan te maken die geschikt zijn om via het www, met html, te worden gebruikt, zijn gratis;
- het is vrij makkelijk om hypertext html-pagina's aan te maken;
- niet alleen teksten, maar ook multimedia-inhoud kunnen worden gepubliceerd;
- het www kan informatiebronnen (en diensten) van het internet die het www voorafgingen, opnemen en toegankelijk maken.

## ON LINE TOEGANKELIJKE INFORMATIEBRONNEN EN -DIENSTEN

### Inleiding

Om informatiebronnen die on line toegankelijk zijn te ontdekken, te vinden, op te zoeken en te gebruiken, hebben we een microcomputer, een systeem voor datacommunicatie en natuurlijk ook enig inzicht en vaardigheden nodig.

Informatiebronnen via netwerken groeien in stijgende lijn, niet alleen in volume, maar ook in belang. Veel bronnen zijn noodzakelijk voor onderzoek, andere zijn gewoon handig. Om concurrentieel te blijven zullen academici, onderzoekers, beheerders, zaken-

lui... voor snelle en efficiënte communicatie en toegang tot informatie steeds meer aangewezen zijn op het netwerk. Doen ze dat niet, dan zullen ze waarschijnlijk achterop raken, want wereldwijd zal de meerderheid van hun collega's zich daar wel mee inlaten.

Toch erkennen we enkele problemen, moeilijkheden die verband houden met informatie die on line beschikbaar is:

- enerzijds is er te veel informatie over bepaalde onderwerpen; veel informatie is met andere woorden overbodig en er zijn veel overlappingen. Anderzijds is er over andere onderwerpen weinig informatie on line toegankelijk;
- de informatie wordt niet geselecteerd en van de meeste bronnen wordt de kwaliteit ook niet gecontroleerd. Het is niet nodig om nieuwe informatie op het net te registreren. Is de informatie die je vindt echt, oprecht en authentiek? Het www verleent on line toegang tot een les over hoe de kwaliteit van op het www gebaseerde informatiebronnen kan worden beoordeeld; die is gratis te bekijken op <http://sosig.esrc.bris.ac.uk/desire/internet-detective.html>;
- informatiebronnen kunnen veranderen, groeien of verdwijnen;
- surfen op het www, en een meer algemeen gebruik van het internet, kan een tijdrovende in plaats van een productieve bezigheid zijn.

Bij de totstandkoming van dit artikel hadden we al toegang tot meer dan 1000 miljoen adressen (documenten) over het internet. Zo'n adres wordt aangeduid door een 'Universal Resource Locator' (URL). Op deze manier worden er meer dan 10 terabyte (10 000 gigabyte) tekstgegevens beschikbaar gesteld.

Veel informatie op het internet is gratis toegankelijk, maar voor bepaalde informatie moet je de producent en/of de distributeur betalen voor je er toegang toe krijgt. Sommige organisaties (instellingen, bedrijven) betalen die bijdrage voor een selectie van bronnen en

maken zo toegang mogelijk, zodat hun leden de informatie schijnbaar gratis kunnen opzoeken en exploiteren.

Veel van de informatie die gratis beschikbaar is, wordt door commerciële organisaties aangeboden en is dus meestal niet objectief of wetenschappelijk, maar in veel gevallen eerder subjectief of misschien zelfs misleidend. Vaak trekt ze ook meer aandacht dan meer objectieve wetenschappelijke informatie. Natuurlijk worden er ook veel informatiebronnen door commerciële organisaties aangeboden die tot de zogenaamde informatie-industrie behoren; die zullen heel waarschijnlijk kwalitatief hoge, objectieve informatie verschaffen, aangezien dit hun manier is om commercieel het hoofd boven water te houden.

In de volgende alinea's richten we ons op verscheidene types informatiebronnen die via het internet toegankelijk zijn. Waar vinden we goede en relevante informatiebronnen op het internet? Uitgaande van de moeilijkheden betreffende dit medium (zie voorgaande) is dit geen eenvoudige vraag. Om via het www potentieel interessante documenten en databanken te vinden, kunnen we gedrukte handleidingen en gidsen of systemen gebruiken die het www ons biedt. Enerzijds vertonen de gedrukte gidsen enkele nadelen: het is vaak niet makkelijk om snel aan een kopie te geraken, je moet ervoor betalen en ze geraken snel verouderd. Anderzijds vertonen de retrievalssystemen van het www voordelen: ze zijn rechtstreeks beschikbaar op dezelfde microcomputer als diegene die we willen gebruiken om informatie te raadplegen op het internet; veel van deze systemen zijn gratis en de meeste worden regelmatig geactualiseerd.

De zoekmachines die op het www worden aangeboden om informatie op het internet op te zoeken, behoren voornamelijk, zoals elk zoekstelsel, tot twee types: diegene die een overzicht bieden waarin kan worden gebladerd en diegene die de formulering van een zoekvraag vereisen.

## Overzichten van het internet om in te grasduinen

In de overzichten of lijsten kan je normaal gezien bladeren (grasduinen) door een boomstructuur (een eenvoudige hiërarchie) of een complexere structuur. De meeste structuren stemmen overeen met een classificatie die specifiek is voor dat bepaalde overzicht; met andere woorden, de meer bekende en klassieke, universele classificatiesystemen worden in de meeste overzichten van internet niet gebruikt. Het zijn virtuele bibliotheken met open rekken, die manueel worden beheerd en door de samenwerking van mensen tot stand zijn gekomen.

Veel van deze overzichten zijn voorzien van een index waardoor hun inhoud vlot kunnen worden opgezocht; de begeleidende classificatiestructuur wordt dan echter niet volledig benut en de gebruiker is soms slachtoffer van de moeilijkheden en de problemen die zoeksystemen met zich brengen als ze gebaseerd zijn op zoekvragen die in een natuurlijke taal geformuleerd worden. De mogelijkheid om het systeem op deze manier te gebruiken, kan bovendien verwarrend zijn aangezien deze overzichten niet echt volledige tekst-internetindexen zijn zoals diegene die andere zoekmachines aanbieden.

De bekendste overzichten behoren tot de populairste en de meest bezochte sites van het www.

Hier volgen enkele voorbeelden van dergelijke systemen die bekend zijn bij de totstandkoming van dit artikel:

Yahoo!, toegankelijk via <http://www.yahoo.com/> en via veel andere sites, is heel populair en gratis te gebruiken. In Yahoo! kan je niet alleen bladeren, grasduinen, maar ook een vraag formuleren om informatie te vergaren, hoewel de hiërarchische structuur op deze manier niet ten volle wordt benut. Wanneer zo'n vraag geen resultaten oplevert, gebruikt Yahoo! een veel grotere externe internetindex (die niet

door Yahoo! zelf is geproduceerd). De gekozen internetindex is doorheen de tijd veranderd. Deze methode is niet heel duidelijk gemaakt en zou de gebruiker kunnen verwarren.

Andere, soortgelijke en concurrerende overzichten die ook gratis toegankelijk zijn, kunnen als onderdeel op de volgende www-sites worden gevonden:

- <http://britannica.com/gecombineerd> met en geïntegreerd in een grote universele encyclopedie
- <http://bubl.ac.uk/>
- <http://www.dmoz.org/>
- <http://www.lii.org/>
- <http://www.looksmart.com/>; dit overzicht is ook te vinden via de populaire site <http://www.altavista.com>.

Het is onmogelijk om een volledige lijst van internetoverzichten wereldwijd te geven, want er zijn er talloze en ze evolueren doorheen de tijd; er zijn wel grote verschillen tussen deze sites. Daarom geven we een overzicht van mogelijke karakteristieken en evaluatiecriteria die voor deze systemen relevant zijn:

- is de site gratis toegankelijk?
- is de reikwijdte (de dekking) groot?
- is de site actueel? Wordt hij regelmatig geactualiseerd? Zijn slechts enkele verwijzingen gebroken ('dood')?
- is er een goede dekking van de bronnen in dat deel van de wereld waarin jij geïnteresseerd bent?
- weigert de beheerder in zijn overzicht een prominente plaats te geven aan sites die daarvoor willen betalen?
- is er een makkelijke gebruikersinterface?
- zijn de antwoordtijden kort?
- is de presentatie goed?
- zijn er dichterbij 'mirror sites' beschikbaar die sneller antwoord geven?
- wordt er voor elke vermelde site een beschrijving, notering, beoordeling en een controle aangeboden?
- wordt er gratis een vertaling van de documenten aangeboden?
- is er goede documentatie over het systeem en on line hulp beschikbaar?
- zijn de stabiliteit en de betrouwbaarheid van het systeem hoog?
- worden er door dezelfde site ook

andere diensten aangeboden? Wordt de 'subject directory' in andere diensten opgenomen? (Bijkomende diensten zijn bijvoorbeeld een internetindex of een www-index of een gateway naar een dergelijke index om op basis van een vraag informatie op te zoeken.)

Behalve de wereldwijde overzichten bestaan er ook die zich specialiseren in bronnen uit een bepaalde streek of land, of in een bepaald onderwerp. Veel van deze meer gefocuseerde overzichten kunnen worden gevonden via een algemener overzicht.

### Internetindexen

Behalve door te bladeren in overzichten kunnen we ook door meer directe zoekopdrachten (queries) veel items (adresseerbare bronnen) op het internet opzoeken en lokaliseren. Elk van deze zoeksystemen met queries is gebaseerd op een databank met verwijzingen naar URL's (pages). Door zoekvragen kan men geschikte URL's vinden via een grote, machinaal tot stand gekomen index van woorden en termen die op basis van de inhoud en de teksten van deze pages is aangemaakt. Om deze databank op te stellen en te blijven actualiseren, worden er continu pages van het internet verzameld door een softwaresysteem voor het 'vullen' van de zoekmachine. Dankzij zo'n zoekmachine met een gebruikersinterface die is opgenomen op het www kan de gebruiker deze databank doorzoeken.

Voorbeelden van dergelijke zoeksystemen op het internet zijn te vinden bij de volgende sites:

- <http://www.alltheweb.com/>
- <http://www.altavista.com/>
- <http://www.google.com/>
- <http://www.hotbot.com/>
- <http://www.search.msn.com/>

Naast de gemeenschappelijke bedoeelingen en eigenschappen van deze systemen, zijn er ook verschillen. Om deze varianten te illustreren en internetgebruikers te helpen bij het kiezen voor een bepaalde zoekmachine, kan de volgende lijst met eigenschappen en evaluatiecriteria nuttig zijn:

- is het gebruik van de zoekmachine gratis;
- hoe volledig is de reikwijdte (dekking);
- hoe goed is de dekking van bronnen voor een bepaald geografisch gebied;
- hoe goed is de dekking voor een bepaald soort documenten;
- is de doorzoekbare databank actueel; wordt de databank regelmatig geactualiseerd; bevatten de zoekresultaten slechts enkele gebroken (dode) verwijzingen;
- wordt 'spamming' gefilterd om meer normale pages meer kans te geven om in de zoekresultaten voor te komen;
- kan het systeem soortgelijke documenten in de resultaten groeperen; elimineert het systeem deze documenten uit zijn databank;
- werkt de databank met een volledige tekstindex van elk ascii en html-document dat in de databank is opgenomen, zodat er door de volledige tekst kan worden gezocht; of worden grote documenten slechts gedeeltelijk opgenomen in de index;
- is de inhoud van meta-tags ook geïndexeerd, met andere woorden, zijn ze te gebruiken als zoekleutel;
- indexeert het systeem buiten eenvoudige ascii en html-bestanden ook andere tekstbestanden op het web, zodat ook die op te sporen zijn; (bijvoorbeeld bestanden in het formaat van de verscheidene versies van Microsoft Word, Microsoft PowerPoint of bestanden in Adobe's Portable Document Format);
- biedt het systeem krachtige zoekopties zoals truncatie, booleaanse zoekcombinaties, het zoeken naar woorden die vaak samen voorkomen (termen), de automatische vertaling van het woord dat je zoekt naar verschillende andere talen, een

spellingcontrole van de gezochte termen...;

- kunnen de resultaten beperkt worden tot een bepaalde periode (bv. op basis van de datum van het bestand zoals die bij de server geregistreerd is of op basis van de recentste indexering van het bestand);
- is de gebruikersinterface duidelijk en gebruiksvriendelijk;
- wordt de gebruikersinterface aangeboden in een taal die je goed beheerst;
- rangschikt het systeem de resultaten volgens hun vermoede relevantie;
- presenteert het systeem elk resultaat (document/pagina/item) in een goede opmaak? Worden de gezochte termen bijvoorbeeld aangeduid (in het vet) in de resultaten.

### Multimediabestanden op het internet opzoeken

Er zijn gespecialiseerde systemen voorhanden om beelden (foto's, kunstwerken...), geluid (toespraken...) animatie of video, kortom multimedia, op het internet op te zoeken. Zo bieden de zoekresultaten van zo'n systeem bij het zoeken naar beelden niet alleen verwijzingen naar de beeldbestanden op het internet, maar ook onmiddellijk kleine versies van de beelden (de 'thumbnails'). Voorbeelden van zulke systemen zijn te vinden op de volgende sites:

- <http://gallery.yahoo.com>
- <http://images.google.com/>
- <http://www.altavista.com>
- <http://www.alltheweb.com>
- <http://www.go.com>

### Attenderingsdiensten gericht op pagina's van het www

Veranderingen opsporen in één of meer voor het publiek toegankelijke www-pagina's kan dankzij attenderingsdienstendie vernieuwingen opsporen en de gebruiker per e-mail daarvan op de hoogte brengen. Een voorbeeld: <http://informant.dartmouth>.



---

edu/. Deze systemen kunnen ook laten weten wanneer ze een nieuw document hebben ontdekt dat overeenstemt met je interesseprofiel.

### Databanken waarvoor moet worden betaald

Zoals eerder vermeld is er op het internet veel informatie die enkel toegankelijk is na betaling aan de producent en/of de verdeler. De eerste commerciële computersystemen die on line informatie verschaffen, ontstonden rond 1975. De meeste zijn nu terug te vinden op het internet. Een bekend en groot bedrijf dat allereerste informatie verschaft, bibliografische zowel als volledige teksten, is Dialog; zijn website is te vinden op <http://dialog.com/>; die verwijst naar verschillende databanksystemen, die originele Dialog (gelegen in de VS) en Data-Star (in Europa) inbegrepen.

De hoeveelheid informatie die door de grootste systemen wordt geboden is enorm, nu nog te vergelijken met alle pagina's van het www tezamen.

### On line toegang tot informatie

Er wordt steeds meer informatie ter beschikking gesteld op het internet en een steeds grotere hoeveelheid ervan is gratis. De kwaliteit van de clientsoftware neemt toe. Bijgevolg stijgt het aantal eindgebruikers dat on line zoekt naar informatie.

---

### WOORDENBOEKEN EN ENCYCLOPEDIËN

Woordenboeken en encyclopedieën worden boven veel andere informatiebronnen verkozen wanneer men geen gedetailleerde informatie nodig heeft over een onderwerp of wanneer men een meer gedetailleerde zoektocht naar een minder gekend onderwerp wil voorbereiden door te zoeken naar de juiste spelling, de context, synoniemen...

Sommige woordenboeken en encyclopedieën worden nu gratis op het www aangeboden. Enkele voorbeelden van bekende encyclopedieën:

- de *Encyclopaedia Britannica*, inclusief een verklarend woordenboek en verwijzingen naar www-sites die via <http://www.britannica.com/> kunnen worden bezocht; toegang tot de meeste gevonden teksten vereist echter toch betaling;
- de *Encarta Concise Free Encyclopedia* op <http://encarta.msn.com/> is inderdaad gratis, maar meer oppervlakkig.

Een lijst van de encyclopedieën op het internet is te vinden op <http://www.internetoracle.com/encyclop.htm>. Andere lijsten van elektronische encyclopedieën zijn te vinden als onderdeel van de internetoverzichten die gratis kunnen worden bezocht (zie boven).

---

### BOEKEN OPZOEKEN

Zelfs in dit tijdperk van op het internet gebaseerde informatiebronnen, wordt er nog veel informatie als gedrukt product verspreid. De inhoud van de meeste boeken is (nog altijd) niet beschikbaar op het internet. De meeste algemene zoekmachines van het internet die hoger werden vernoemd, laten ons niet toe meer informatie te vinden over boeken die voor ons nuttig kunnen zijn. Specifieke systemen om boeken op te sporen, kunnen dus handig en welkom zijn. Om actuele boeken te vinden, zijn de bibliografische databanken die grote boekenwinkels verzamelen erg interessant. Vele bieden een goede dekking en zijn gratis. Enkele voorbeelden:

- <http://www.amazon.com/> en lokale versies zoals <http://www.amazon.co.uk/>
- <http://www.bn.com/>
- <http://www.blackwell.co.uk/>

Zelfs vergelijkingen van het aanbod en de prijzen in boekenwinkels (en ook aanbiedingen op het vlak van muziek, film en andere producten) zijn gratis ver-

krijgbaar; bijvoorbeeld bij <http://www.dealtime.com/>.

Om oudere boeken terug te vinden, zijn vooral de catalogi van bibliotheken handig; vele grote zijn gratis on line beschikbaar.

---

### TIJDSCHRIFTENARTIKELS OPZOEKEN

Duizenden on line databanken bieden bibliografieën of zelfs volledige teksten uit tijdschriften over een bepaald onderwerp, maar toegang daartoe is doorgaans niet gratis. Slechts enkele organisaties bieden gratis grote databanken die summiere toegang verlenen tot bibliografieën van gepubliceerde tijdschriftenartikels over uiteenlopende onderwerpen. Enkele voorbeelden:

- <http://www.form.inist.fr/public/eng/constl.html>
- <http://www.ingenta.com/>
- <http://www.northernlight.com/>

---

### ELEKTRONISCHE NIEUWSBRIEVEN EN TIJDSCHRIFTEN

Sinds het einde van de jaren 1990 zijn elektronische tijdschriften een nieuw, niet meer weg te denken communicatiemiddel. We onderscheiden verschillende soorten, onder andere equivalenten van de papieren versie en de puur elektronische publicaties. Het www is het belangrijkste platform geworden voor toegang tot elektronische nieuwsbrieven en tijdschriften.

---

### ZOEKEN MET CITATEN

Naast gewoon zoeken met sleutelwoorden, kunnen we ook gebruikmaken van citaten, verwijzingen in wetenschappelijke artikels, om verwante en misschien relevante informatie te vinden. Het informatiesysteem bij uitstek dat deze aanpak mogelijk maakt, is het door het ISI geproduceerde, grote 'Web of Science'.

## INTERESSEGROEPEN GEBASEERD OP COMPUTERNETWERKEN

Eén van de methodes om informatie uit te wisselen tussen mensen met een gemeenschappelijke interesse of probleem, zijn de interessegroepen gebaseerd op computernetwerken en op internet in het bijzonder. Ze laten hun leden toe van nieuwe ontwikkelingen op de hoogte te blijven en toegang te krijgen tot de expertise van de groep. We onderscheiden verschillende systemen, onder andere diegene die gebaseerd zijn op elektronische post en diegene die gebruikmaken van Usenet News. Bovendien biedt het www sinds het eind van de jaren 1990 gateways naar deze en een basis voor soortgelijke systemen. We kunnen potentieel relevante groepen vinden (en de manier om lid te worden) via grasduinen door het www of meer direct door gratis gebruik te maken van databanken over zulke groepen (via grasduinen of zoeken met woorden). Voorbeelden daarvan vinden we bij <http://www.liszt.com/> en <http://groups.yahoo.com/>. Op <http://groups.google.com/> wordt sinds 2001 ook gratis een groot aantal van de door Usenet ver-

spreide teksten ter beschikking gesteld als up-to-date, uniek archief in de vorm van een doorzoekbare database.

## INTERBIBLIOTHECAIR LEENVERKEER EN DOCUMENTLEVERANTIE

Veel bibliotheken en meer commerciële organisaties bezorgen boeken of kopieën van documenten aan een persoon of een instituut dat hierom vraagt. Dergelijke diensten zijn handig als we weten dat een gezocht document wel bestaat, maar als we het niet kunnen bemachtigen.

## BESLUIT

De methodes om informatie ter beschikking te stellen – en dus ook om informatie te vinden – evolueren snel. Relatief eenvoudige vragen kunnen we nu al heel vlug bevredigend beantwoorden door gebruik te maken van gratis zoeksystemen via internet en het www. Meer exhaustief opsporen van informatie blijft echter ingewikkeld en tijdrovend, vooral als we zoeken naar informatiebronnen die handelen over een combinatie van concepten: enkele hinderpalen zijn immers nog niet uit de weg geruimd, zoals de basisproblemen van het zoeken naar informatie via natuurlijke taal en de verstrooiing van informatie over allerlei systemen.

## BIBLIOGRAFIE

- Frstrup, J.A., *Usenet netnews for everyone*. - Englewood Cliffs, New Jersey: PTR Prentice Hall, 1994. - 396 p. [Een van de weinige boeken die volledig gewijd zijn aan het groepscommunicatiesysteem Usenet news.]
- Large A., Tedd L.A., and Hartley R.J., *Information seeking in the online age: principles and practice*. - London: Bowker-Saur, 1999. - 308 p. [Een van de beste en meest recent geactualiseerde boeken over on line informatie zoeken. Het is algemeen zonder zich op een specifiek onderwerp te focussen.]
- Lynch C., *Searching the Internet*, in: *Scientific American*, (maart 1997); p. 44-48. [Een goed overzicht van internetzoekmachines voor leken.]
- *Gale directory of databases*. - Detroit: Gale Research Inc., 1994-... [de belangrijkste gids met beschrijvingen van duizenden databanken beschikbaar op cd-rom en/of on line. Hij bestaat in een gedrukte versie die twee keer per jaar geactualiseerd wordt en in een on line versie.]
- Notess G.R., *Search engines showdown*. [online] Beschikbaar op: <http://searchengineshowdown.com/> [geciteerd in 2001] [Deze website geeft gratis toegang tot een overzicht van algemene zoekinstrumenten om informatie op het internet te vinden. Werd tot nu toe regelmatig bijgewerkt.]
- Schwartz C., *Web search engines*, in: *Journal of the American Society of Information Science*, 49 (1998) 11; p. 973-982. [Een goed wetenschappelijk overzichtsartikel over zoekmachines op het web.]
- *Searching the Internet: recommended sites and search techniques*. [Online] Beschikbaar op: <http://www.albany.edu/library/internet/search.html> [Een vrij toegankelijke site op het www die een goede introductie geeft over informatie vinden op het internet.]

## SAMENVATTING

Dit artikel brengt een overzicht van moderne methodes om informatie op te sporen over een bepaald onderwerp, ter ondersteuning van het management, besluitvorming, wetenschappelijk onderzoek en dergelijke. De volgende onderwerpen worden kort besproken: het begrip 'informatie', informatie op cd-rom, multimedia en hypermedia, computernetwerken en het internet die de weg hebben geëffend voor het world wide web. Dit systeem geeft eenvoudig toegang tot vele on line informatiebronnen en -diensten. Voorts wordt aandacht besteed aan: woordenboeken en encyclopedieën, boeken en tijdschriftartikels, elektronische tijdschriften, zoeken met citaten, interessegroepen gebaseerd op computernetwerken, interbibliothecair leenverkeer en documentleverantie.

## ABSTRACT

This article presents an overview of contemporary methods to find information about a particular subject to support management, decision making, scientific research and so on. The following subjects are briefly discussed: the concept of 'information', information on CD-ROM, multimedia and hypermedia, computer data networks and the Internet that have paved the way for the World-Wide Web. This system now allows easy access to many on line access information sources and services. Next are discussed: dictionaries and encyclopaedias, books and journal articles, electronic journals, citation searching, computer-based network interest groups, interlibrary lending and document supply.