

---

## Informatiegedragkunde: pijler van de moderne informatieverzorging \*

---

Reactie op de voordracht van J.A. Michon

G.A.M. Kempen

Professor Michon heeft een bewonderenswaardig heldere en onderhoudende inleiding in de cognitiewetenschap gegeven. Zijn theoretische beschouwing onderschrijf ik vrijwel volledig. Op zijn vraag of de ballonnetjes die hij heeft opgelaten moeten worden doorgeprikt, antwoord ik dan ook ontkennend. Toch blijft er wel het een en ander te wensen over. Ik mis namelijk de aansluiting op de praktijk van de informatieverzorging. In plaats van ballonnetjes door te prikken zal ik er enkele toevoegen, in de hoop dat het resultaat nog beter oogt dan het origineel.

### 1. Informatie: inhoud versus vorm

Professor Michon eindigt zijn betoog met de dringende aanbeveling om 'toegangstructuren' te ontwikkelen tot de 'gigantische hoeveelheden ongeïnterpreteerde kennis die in bibliotheken en databanken is opgeslagen'. Ik ben het er volstrekt mee eens dat hier een der hoofdtaken ligt van de moderne informatieverzorging. Maar, met alle respect, dat is gemakkelijker gezegd dan gedaan. Het ontwerpen van een 'gracieuze koppeling' tussen toegang- en zoekstructuren enerzijds en de 'cognitieve processen van de gebruiker' anderzijds heeft zich ontpopt als een zeer hardnekkig probleem. Professor Michon legt zelf de wortel van het kwaad bloot bij zijn behandeling van de fameuze Engelse zin *Dogs dogs dogs bite bite bite*. Veronderstellen we dat deze voorkomt in een biologisch artikel (Ik weet wel dat deze aanname irreal is want alleen taalkundigen kunnen zoiets herkennen). Een geautomatiseerd systeem dat deze zou willen herkennen als een uitspraak over de pikorde bij een bepaalde diersoort, kan niet volstaan met het extraheren van de trefwoorden *dogs* en *bite*. Het zal minstens door moeten dringen tot wat de linguïst Noam Chomsky 'dieptestructuur van de zin' heeft gedoopt. Als het systeem bovendien weet hoe de betekenissen van *hond*, *dier*, *bijten*, *straf*, *macht*, *hierarchie*, *orde* en dergelijke samenhangen, kan het via een ingewikkelde redenering wellicht afleiden dat de zin iets met pikordes van doen heeft.

En dat is nog maar een kant van het probleem. Aan de andere kant -van het probleem en van het beeldscherm- zit de gebruiker die niet weet in welke termen hij zijn informatiebehoefte moet formuleren teneinde contact te maken met zoveel mogelijk relevante teksten (en een minimaal aantal irrelevante). Inderdaad is heel wat kunstmatige intelligentie nodig om een ook maar enigszins gracieuze koppeling tot stand te brengen tussen het informatiesysteem en cognitieve processen van de gebruiker.

Het algemene probleem dat hier aan de orde is, vloeit voort uit het

feit dat informatiesystemen meestal werken met 'oppervlaktestructuren' terwijl de gebruiker niet geïnteresseerd is in oppervlaktestructuren maar juist in 'dieptestructuren'. En helaas bestaat er tussen oppervlaktestructuur en dieptestructuur tussen de verschijningsvorm van informatie en de informatie zelf geen eenvoudige correspondentie. Hoezeer dit ook een open deur lijkt, het kan geen kwaad hem nog maar eens in te trappen. Want bij het ontwerpen van nieuwe informatiesystemen wordt nog steeds teveel uitgegaan van de concrete vorm waarin men de informatie aantreft. Wie bijvoorbeeld een boekuitgave overzet naar een magnetische of optische informatiedrager en de tekst toegankelijk maakt via een index in de vorm van een zogenaamde inverted file, kiest voor de meest 'oppervlaktestructuur' methode.

Meer diepgang is onder meer als volgt te bereiken:

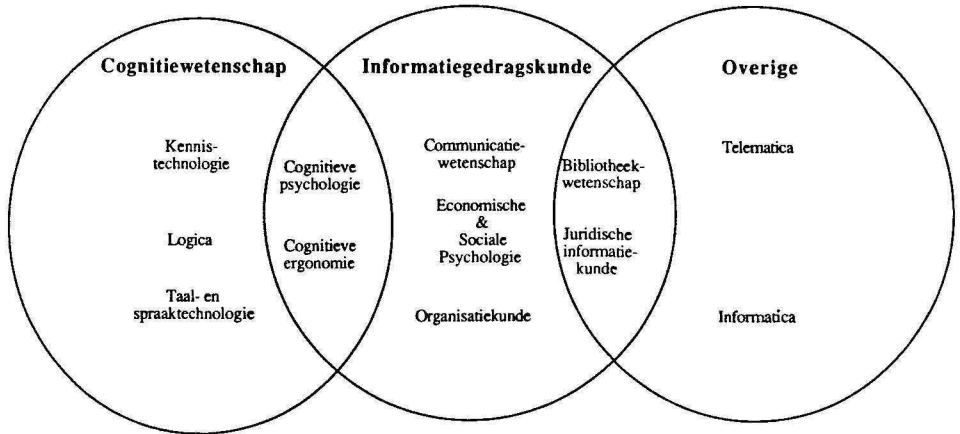
- Selectie van alleen de relevante trefwoorden met behulp van statistische en andere methoden uit het vak Information Retrieval.
- Toevoeging van synoniemen, woordvelden, verbuigingen en vervoegingen, spellingvarianten; ontleding van samenstellingen; opvang van type- en spelfouten. Dit maakt het systeem minder afhankelijk van de toevallige woordkeus van auteur en gebruiker.
- Segmentering van de tekst, volgens een overzichtelijk schema, tot korte passages die met behulp van eenvoudige commando's afzonderlijk en in elke gewenste volgorde oproepbaar zijn (cf. hypertext). Om een diversiteit van redenen lezen mensen niet graag grote lappen tekst vanaf het beeldscherm.
- Indien de inhoud van de tekst voornamelijk bestaat uit opsomming van een groot aantal feiten, kunnen de zinnen worden omgezet in (quasi-)logische formules die passen in een relationele databank. (Op dit principe berust Lexitron, het door Van Dale Lexicografie uitgebrachte woordenboek annex encyclopedie op CD-ROM).
- Teksten met een meer betogend karakter, bijvoorbeeld onderzoeksrapporten, kunnen semantisch geanalyseerd worden met behulp van technieken uit de kennistechnologie (Kunstmatige Intelligentie). De inhoud kan vervolgens opgeslagen worden in een kennisbank, bijvoorbeeld in de vorm van kennisgraphen. Zo'n bestand zou de kern kunnen uitmaken van een expertsysteem dat zelfstandig vragen interpreteert en beantwoordt. (Deze techniek staat nog in de kinderschoenen. Nederlands onderzoek erover is gepubliceerd door onder anderen Bakker 1987, en Van Cuijlenburg et al., 1988).

Ik besef terdege dat informatiesystemen net als zeeschepen bij grotere diepgang moeilijker op gang te krijgen zijn. Het inrichten van een 'moederbestand' waar de informatie in een abstracte notatie ligt opgeslagen, kost immers zeer veel tijd en geld. Op den duur betaalt deze investering zichzelf evenwel terug; uit het moederbestand (dieptestructuur) kan de informatieleverancier tegen lage kostprijs nieuwe informatiesystemen of -producten (oppervlaktestructuren) afleiden. Het moederbestand biedt immers sterk verruimde mogelijkheden tot efficiënte actualisatie, aanpassing aan speciale gebruikersgroepen, vertaling, overzetting naar andere informatiedragers, toevoeging van extra functionaliteit (intelligentie, gebruikersvriendelijkheid), enzovoorts.

## 2. Informatiegedragskunde

Tot voor enkele jaren maakten alle niet-mondelinge vormen van informatieverzorging gebruik van een en hetzelfde medium: drukwerk. Zowel vormgeving als inhoud van de te verspreiden informatie waren erop afgestemd. De eerste elektronische informatiediensten boden niet veel meer dan (eenvoudig) drukwerk op beeldscherm - net zoals een eeuw terug automobielen leken op koetsen met een motor in plaats van het paard. Nu het spectrum van nieuwe elektronische en optische media zich snel verbreedt, moet de informatieleverancier zich veel intensiever dan vroeger gaan bezighouden met de vraag welke aanbiedingsvorm het gunstigste effect zal sorteren. Het antwoord blijkt onverwacht lastig. De reactie van de informatieconsument op nieuwe media laat zich nauwelijks voorspellen, en de informatiewetenschap staat met lege handen. Kennelijk zijn de determinanten van informatiegedrag nog grotendeels onontgonnen gebied. We zijn hard toe aan een vak 'informatiegedragskunde' dat zich richt op het ontwarren van het kluwen van sociale, organisatorische, juridische, economische, motivationele, attitudinale, cognitieve en andere factoren die informatie(zoek)gedrag bepalen. (Zie bijvoorbeeld Andriessen, 1988, en Sepstrup, 1980.)

# INFORMATIEVERZORGING



Figuur 1 Wetenschappelijke fundering van de informatieverzorging.

Ik acht het van groot belang dat dit vakgebied in brede kring erkenning vindt als een van de wetenschappelijke pijlers die de sector van de informatieverzorging schragen. De Europese Gemeenschap is op de goede weg door in het RACE-onderzoeksprogramma voor telecommunicatie grote nadruk te leggen op 'upsability - engineering'.

Figuur 1 brengt de relaties in beeld die in deze visie tussen enkele betrokken vakgebieden bestaan. Tevens maakt de figuur duidelijk waarom ik het niet eens kan zijn met professor Michon wanneer hij de documentatie- en informatiewetenschap als "een serieuze tak van de cognitiewetenschap" kenschetst. Uiteraard wens ik niet te bestrijden dat de informatiewetenschap een serieuze tak van wetenschap is! Wel ben ik van oordeel dat, zoals tot uitdrukking gebracht in Figuur 1, de informatiewetenschap zich in belangrijke mate buiten de cognitiewetenschap afspeelt. Dit geldt bijvoorbeeld voor de tak die ik als informatiedragkunde heb aangeduid.

Tot slot zou ik het voorgaande enigszins willen concretiseren in de vorm van een aanbeveling voor onderzoek. Inzicht in de wetmatigheden achter informatiegedrag van personen en organisaties is verkrijgbaar via empirische studies waarin drie typen factoren onafhankelijk van elkaar worden gevarieerd: (1) de informatie-inhoud (diepestructuur), (2) de aanbiedingsvorm van de informatie (oppervlaktestructuur), en (3) de informatieconsument. Dergelijk 'vergeleekend multimedial gebruikersonderzoek' leidt op den duur wellicht tot betrouwbare wetenschappelijke kennis over gunstige combinaties van informatie-inhoud, informatiesysteem en informatiezoeker, en aldus tot daling van het aantal kostbare mislukkingen bij de introductie van nieuwe media.

- \* *Ik dank de leden van de RABIN en de medewerkers van het RABIN-bureau te Den Haag voor de uitvoerige discussies over de onderhavige problematiek. Niettemin komen de hier weergegeven opinies geheel en al voor mijn rekening.*

## Literatuur

Andriessen, J.H.T.H. (1988). Telematica in bedrijf. Een organisatie-psychologische beschouwing over elektronisch communiceren. Inaugurele Rede, Vrije Universiteit, Amsterdam.

Bakker, R.R. (1987). Knowledge graphs: representation and structuring of scientific knowlegde. Dissertatie, Universiteit Twente, Enschede.

Cuilenburg, J.J.van, J.Kleinnijenhuis & J.A. de Ridder (1988). Tekst en betoog. Naar een computergestuurde inhoudsanalyse van betogende teksten. Muiderberg: Coutinho.

Sepstrup, P. (1980). Consumption of mass communication. Construction of a model on information consumption behavior. Research in marketing, 3, 105-142.