

*BIJLAGE II. COMPUTERISERING BIBLIOGRAFIE BETREFFENDE
NEDERLANDSE POLITIEKE PARTIJEN*

DOOR R.A. KOOLE

1. inleiding

In de tweede helft van 1978 is op het Documentatiecentrum Nederlandse Politieke Partijen van de Rijksuniversiteit Groningen een begin gemaakt met de computermatige verwerking van de daar aanwezige bibliografie betreffende Nederlandse politieke partijen.

De noodzaak daartoe laat zich raden: de hoeveelheid literatuur inzake Nederlandse politieke partijen is groot en breidt zich enorm uit, terwijl van de kant van de gebruikers van deze bibliografie steeds meer gespecificeerde vragen komen, die bovendien vaak actuele gebeurtenissen raken. Deze situatie brengt een aantal eisen met zich mee, waaraan ten dele ook door de 'conventionele' kaartsystemen kan worden voldaan. Deze kaartsystemen dienen daartoe zeer verfijnd te zijn en bovendien uitgebreide verwijzingsystemen te bevatten. Dit heeft tot gevolg dat ze zeer arbeidsintensief zijn en tamelijk veel ruimte in beslag nemen.

De computermatige verwerking van de bibliografische gegevens biedt echter meer voordelen dan alleen besparingen op ruimte en arbeid: er is geen scheiding noodzakelijk tussen verschillende soorten rangschikkingen van de bibliografie (alfabetisch op auteur, systematisch of anderszins); bij de raadpleging van het bibliografisch bestand kan op veel meer onderdelen geselecteerd worden dan alleen op auteur of classificatie, zonder dat dit uitgebreide verwijzingsystemen vergt (men kan b.v. selecteren op de taal, het jaar van uitgave, etc.); de toegankelijkheid van het bestand wordt verder vergroot, doordat de (universitaire) gebruiker niet locatie-gebonden is (het is b.v. mogelijk om vanuit Leiden rechtstreeks gebruik te maken van het in Groningen

opgebouwde bestand); de snelheid van informatie-verwerking ligt hoger dan bij 'conventionele' kaartsystemen, waardoor informatie omtrent recente gebeurtenissen sneller beschikbaar is; de computermatig verwerkte literatuur biedt meer mogelijkheden voor inhoudsanalytisch onderzoek; en bovendien zou het bibliografisch bestand 'gekoppeld' kunnen worden aan andere bestanden met b.v. data betreffende partijfunctionarissen of data inzake de interne organisatie van de verschillende politieke partijen.

Voor het dagelijks verwerken van de bibliografische informatie is het noodzakelijk te beschikken over een computer-eindstation ('terminal') met een daarbijbehorende afdruckinstallatie, de 'printer'. Met deze apparatuur kan via de gewone telefoonlijn verbinding verkregen worden met de 'hoofdcomputer' in het Rekencentrum van de Rijksuniversiteit Groningen. Daar worden de via de 'terminal' aangevoerde gegevens verwerkt en bewaard op magneetbanden, zodat ze steeds beschikbaar blijven.

Alvorens over te gaan tot de aanschaf van dergelijke apparatuur, besloot men op het Documentatiecentrum gedurende een 'proefperiode' de computermatige verwerking van de bibliografische gegevens op de praktische uitvoerbaarheid ervan te toetsen.

In de volgende paragrafen wordt een overzicht gegeven van de organisatie en de resultaten van de computermatige verwerking, gedurende de proefperiode tot eind december 1978.

2. RIQS

Het voor de computermatige verwerking van bibliografische gegevens noodzakelijke programma kan in principe zelf ontworpen worden. Gekozen is echter voor het in de Verenigde Staten (Northwestern University) ontworpen en uitgetestte RIQS-pro-

gramma, dat eind 1977 door de Rijksuniversiteit Groningen (in de zng. 'batch'-versie) werd aangeschaft.

"Het Remote Information Query System (RIQS) is een computerprogramma voor 'information retrieval', d.w.z. het stelt de gebruiker in staat om uit een verzameling informatie die stukjes te halen die hij nodig heeft. Een zeer gebruikelijke toepassing van een 'information retrieval' systeem is de catalogus van een bibliotheek. Bibliotheekgebruikers zijn in staat om (via de alfabetische catalogus) alle aanwezige boeken van een bepaalde schrijver op te zoeken, en vaak ook (via de systematische catalogus) alle aanwezige boeken over een bepaald onderwerp. RIQS stelt de gebruiker in staat om dit zoekproces aanzienlijk sneller te laten verlopen en heeft tevens een aantal additionele mogelijkheden boven de kaartenbak methode." (Brouwer, E. en B. Niemöller. Inleiding RIQS. Amsterdam, Universiteit van Amsterdam, 1977, p. 1).

3. *'record definition' en invoering gegevens*

De bibliografische en eventuele andere gegevens van een eenheid literatuur (boek, artikel, brochure, etc.) moeten op een bepaalde manier aan de computer ter verwerking worden aangeboden. De computer voegt dan deze eenheid gegevens in de vorm van een 'record' toe aan de bestaande verzameling 'records', de zng. 'RIQS-file'. De structuur van een 'record' is van tevoren gedefinieerd. De 'record definition' voor de bibliografie betreffende Nederlandse politieke partijen ziet er als volgt uit:

- (1) classificatie
- (2) auteur(s), hoofdwoord
- (3) titel, ondertitel, druk
- (4) plaats van uitgave, uitgever, jaar van uitgave
- (5) aantal pp., ISBN-nummer
- (6) tijdschriftinformatie
- (7) illustraties, register, literatuuropgave

- (8) bijlagen, notenapparaat, taal
- (9) aard van het geschrift
- (10) aanvullende bibliografische informatie
- (11) zie ook (event. verwijzing)
- (12) recensie in:
- (13) bibliotheek
- (14) abstract (trefwoorden)
- (15) jaar van uitgave
- (16) t/m (20) ...

Enkele 'items' wil ik hier noemen. De classificatie, bedoeld bij 'item' (1), is de classificatie, zoals die gebruikt wordt in het (nieuwe) bibliografische systeem van het Documentatiecentrum. 'Item' (14) is belangrijk omdat daar de meer inhoudelijke gegevens betreffende een eenheid literatuur kunnen worden vermeld. In de proefperiode is gekozen voor een systeem van trefwoorden, dat met het invoeren van nieuwe eenheden literatuur steeds uitgebreider wordt, maar dat d.m.v. een zgn. 'KWIC'-opdracht ('keyword in context') aan de computer getoetst kan worden op zijn validiteit (zie voorbeeld 1): over het algemeen is het zo dat hoe vaker een bepaald trefwoord voorkomt, des te functioneler het is. 'Wildgroei' van trefwoorden kan zodoende voorkomen worden.

Naast of in plaats van een trefwoordensysteem kan men natuurlijk ook andervormige inhoudelijke informatie geven b.v. in de vorm van 'indicatieve referaten', waarin in enkele zinnen de hoofdzaken van een publicatie worden weergegeven of in de vorm van de meer uitgebreide 'informatieve referaten'. De keuze van de vorm, die hierbij gebruikt wordt, hangt af van de financiële mogelijkheden, de wensen van de gebruiker en de eisen, die een inhoudsanalytisch onderzoek stelt. Voor zover het zich nu laat beoordelen, biedt het trefwoorden-systeem bij de bibliografie betreffende de Nederlandse politieke partijen voldoende mogelijkheden voor een gericht literatuuronderzoek.

'Item' (15) is een herhaling van het onder het 'item' (4) of 'item' (6) genoemde jaartal. Het karakter is echter verschillend. Terwijl de overige 'items' een alfanumeriek karakter hebben (d.w.z.: zowel letters als cijfers kunnen worden ingevuld) heeft 'item' (15) een numeriek karakter, waardoor zoekacties mogelijk zijn, waarin meer condities betreffende relaties kunnen worden gesteld. B.v.: d.m.v. 'item' (15) kan men vragen naar literatuur verschenen tussen 1960 en 1970, terwijl zonder dit 'item' alleen naar literatuur uit een bepaald jaar gevraagd zou kunnen worden.

De 'items' (16) t/m (20) zijn reserve-'items', met verschillende karakters.

De organisatie van de invoering van deze gegevens is op het Documentatiecentrum zo, dat door de medewerkers op stencils met voorgedrukte 'record definitions', de relevante gegevens inzake een eenheid literatuur worden genoteerd. Vanaf de stencils worden deze gegevens (na de noodzakelijke 'stuurkaarten') op de terminal ingetikt en ter verwerking aan de computer aangeboden, die ze weer toevoegd aan de 'RIQS-file' op de magneetband. Het is natuurlijk mogelijk om de fase van het invullen van stencils over te slaan, maar vooralsnog biedt deze fase een uitstekende mogelijkheid om eenheid in de wijze van invoering te verkrijgen en te behouden.

4. mogelijkheden

De zo ingevoerde gegevens bieden een scala van mogelijke zoekopdrachten. Hoe meer voorwaarden men in de opdracht stelt, des te geringer het aantal publicaties zal zijn, dat de computer 'rapporteert'.

Men kan b.v. de computer opdracht geven om alle publicaties, volgens alfabetische volgorde op auteursnaam ('item' (2)) af te drukken, waardoor men een lijst krijgt vergelijkbaar met een alfabetische catalogus. Men kan echter ook een alfabetische

BEGIN SEARCH
IF (1) EQ "29" THEN PLACE IN SET 1..



RIJK

NO. OF RECORDS IN SET 1 BEFORE SEARCH = 0
NO. OF RECORDS PUT IN SET 1 BY SEARCH = 13
NO. OF RECORDS IN SET 1 AFTER SEARCH = 13
NO. OF REPORTS GENERATED BY SEARCH 1 = 0

THERE ARE 13 ENTRIES TO BE PRINTED
INDEX ON SET 1
AUTHOR ON (14).
PRINT (2) * (3) * (4) * (6).

ALGRA, H.

ALGRA, H.
MIJN WERK, MIJN LEVEN
ASSEN, VAN GOE CUYA, 1970

JONKMAN, J. A. • INDONE

JONKMAN, J. A.
NEDERLAND EN INDOONESISIE REIDE VRIJ: GEZIEN VANUIT HET NEDERLANDSE PARLEMENT: MEMOIRIES
ASSEN, VAN GORCUM, 1977

KAPPEYNE VAN DE COPPE

RIJK, M. DE
VVD-KAMERLID ANNELIEN KAPPEYNE: AAN EEN COMPROMIS OVER DE ABORTUS DOEF IK NIET MEE
HAASSE POST, 6-1978, NO. 3A-23 SEP, RLZ, 6-R

KOEKOEK, H.

RAAFTMANS, K. (P.S. VAN CHARLES PHYS)
HOEF KOEKOEK
AMSTEDAM, DE STANDAARD, 1966

PAIS, A.

BERKEL, M.
MINISTER APTE PAIS: "MACHT? NEE: TIJDELHEID? JA"
HAASSE POST, 6-1978, NO. 3A-23 SEP) • RLZ, 18-27

Voorbeeld 2

volgorde op andere 'items', b.v. de titel of het eerstvermelde trefwoord aanbrengen, verlangen.

Bij het afdrukken kan men verlangen dat alle gegevens van een publicatie (een heel 'record') afgedrukt worden, men kan ook slechts enkele 'items' laten afdrukken.

Het kan zijn dat men slechts geïnteresseerd is in een bepaalde groep literatuur (b.v. alle literatuur betreffende de ARP of alle biografische werken of literatuur uit een bepaalde periode) en slechts binnen deze groep de gespecificeerde zoekopdrachten wil laten verrichten. Dit is mogelijk door deze groep eerst op een aparte 'set' te plaatsen en daarna de computer op deze 'set' te laten 'rekenen'.

In voorbeeld 2 zijn een aantal van de boven vermelde mogelijkheden gecombineerd. Eerst zijn alle werken met biografische gegevens (classificatienummer 29) in set 1 geplaatst. Het voorbeeld laat zien dat de computer 13 werken vond met biografische gegevens (het bestand bevatte toen nog slechts 55 werken). Daarna is volgens het commando "index on set 1 - author on (14)." opdracht gegeven om binnen set 1 alle werken in alfabetische volgorde te plaatsen volgens de in het werk behandelde persoon, wiens naam als eerste trefwoord onder 'item' (14) stond vermeld. Vervolgens is met het commando "print (2), (3), (4), (6)." opdracht gegeven om alleen de meest belangrijke bibliografische informatie af te drukken (niet de gehele 'uitdraai' is in voorbeeld 2 afgedrukt).

In principe kan men op alle gegevens, die in de 'record definition' bepaald zijn, zoeken. Enkelvoudige voorwaarden of combinaties van voorwaarden zijn mogelijk. Het zou te ver voeren om alle mogelijkheden op te sommen. Men kan tenslotte de gekste vragen stellen: zo is het mogelijk alleen werken, die in tweede druk verschenen zijn, te laten vermelden. Het is aan de documentalist(e) ervoor te zorgen, dat de vragen

zo gesteld worden, dat alleen de relevante informatie door de computer wordt geleverd en dat onzinnige vragen worden geweerd.

Enkele mogelijke zoekopdrachten wil ik hier nog noemen:

Een eenvoudige, maar vaak verlangde, zoekopdracht is alle werken van één bepaalde auteur af te laten drukken. Voorbeeld 3 nu laat zien dat ook wanneer de desbetreffende auteur (in dit geval 'Dittrich') niet als eerste wordt genoemd, het werk wordt afgedrukt. Een systeem van verwijzingen is bij computermatige verwerking dus niet meer nodig.

In het geval dat iemand (uit het buitenland) informatie wil over Nederlandse politieke partijen, maar de Nederlandse taal niet beheerst, kan worden geselecteerd op de taal. B.v. alleen engelstalige werken kunnen worden afgedrukt.

Een ander voordeel van de computermatige verwerking voor bibliografisch onderzoek is dat geselecteerd kan worden op publicaties, die een (uitgebreide) literatuuropgave of een notenapparaat bezitten. Voorbeeld 4 laat dit zien. Gevraagd werd of onder 'item' (7) de code LLL (voor literatuuropgave) of onder 'item' (8) de code NNN (voor notenapparaat) voorkwam. Bovendien werd gevraagd dit alleen te doen bij werken betreffende een bepaald onderwerp en betreffende een bepaalde partij (onder 'item' (1) geclassificeerd). Zo wordt uit voorbeeld 4 nogmaals duidelijk dat zeer gerichte vragen mogelijk zijn.

5. voorlopige conclusies

Op basis van de in de 'proefperiode' opgedane ervaring kan men enkele voorlopige conclusies trekken:

a. technisch gezien is het mogelijk om een bibliografie aan te leggen, die geschikt is om op een snelle en trefzekere wijze informatie over literatuur aan de gebruiker te verschaffen. Daarbij zijn de mogelijkheden, die een 'gecomputeriseerde'

BEGIN SEARCH
IF (1) EQ "PVDA" AND (1) EQ "24" THEN PLACE IN SFT 2..



RUC

NO. OF RECORDS IN SFT 2 BEFORE SEARCH = 0
NO. OF RECORDS PUT IN SFT 2 BY SEARCH = 1
NO. OF RECORDS IN SFT 2 AFTER SEARCH = 1
NO. OF RECORDS GENERATED BY SEARCH = 1 = 0

OUTPUT OF SEARCH 2

BEGIN SEARCH OF SFT 2
IF (7) EQ "LLN" OR (8) EQ "NNN" THEN PRINT RECORD..

RECORD NUMBER 10

1. CLASSIFICATIE
AAA PVDA 24
2. TITEL/TYS/7 HOOFDWOORDEN
APADT
3. TITEL: ONDEFTITEL: BOK
APADT OF SAME: GEZICHTSDOEL EN OMTRENT MAATSCHAPPELIJK WERK EN VERZUILLING
4. PLATS VAN UITGAVE, UITGIVER, JAAR VAN UITGAVE
AMSTERDAM, DE ABDELINGSPEDS, (JAAR) AFCKMAN STICHTING, 1960
5. AANTAL PP. / STRUKTUUR
31
7. ILLUSTRATIES, ONTWERP, TITELBLADOPGAVE
NEE, NEE, ILL
8. BIJLAGEN, NOTEN/OPMANT, TAAL
NEE, NEE
9. AARD VAN AFT. GESCHAPPT
VAPRODT
14. TPF/WOORDEN
VERZUILLING, MAATSCHAPPELIJ WERK
15. JAAR VAN UITGAVE
1963

Voorbeeld 4

bibliografie biedt aanzienlijk groter dan de mogelijkheden, die een 'conventionele' bibliografie zou kunnen bieden, hetgeen voor de gebruiker van de bibliografie betreffende Nederlandse politieke partijen een aanzienlijke tijdsbesparing en een meer verfijnde informatie betekent.

b. doordat een publicatie slechts eenmaal verwerkt hoeft te worden in het bestand, is de gecomputeriseerde verwerking minder arbeidsintensief dan het kaartensysteem. Doordat echter in de in de vorige paragrafen geschetste opzet meer gegevens worden opgenomen (vgl. de 'record definition') en vooral doordat het 'oude' aanwezige bestand ook computermatig verwerkt moet worden, zal in de eerste jaren het arbeidsexcessieve element van de computermatige verwerking niet direct zichtbaar zijn.

c. voor het dagelijks documenteren van publicaties is de aanwezigheid van een 'terminal' (plus 'printer') onontbeerlijk. Dit vereist een investering van ca. fl. 16.000. Hierbij dient opgemerkt te worden, dat ook bij het handhaven van een kaartsysteem enkele investeringen in apparatuur noodzakelijk zijn.

d. op dit moment betaalt het Rekencentrum de computerrekenkosten van universitaire gebruikers/instellingen. Ook indien dit niet meer het geval zal zijn behoeven de kosten voor de gebruiker niet te hoog te zijn. De universitaire gebruiker zal dan de kosten bij zijn of haar faculteit kunnen declareren. De niet-universitaire gebruiker moet de kosten van de computerrekening wel zelf betalen (tenzij in het kader van de dienstverlening deze kosten geheel of gedeeltelijk gesubsidieerd zouden worden). Bij een combinatie van zoekopdrachten kunnen deze kosten zeer beperkt worden. Bovendien zouden 'uitdraaien' van veel gevraagde zoekopdrachten (gratis) ter inzage kunnen liggen.

e. de computermatige verwerking van bibliografische gegevens vereist een zekere kennis van het gebruik van de computer. Deze kennis zou bij de wetenschappelijke ambtenaar en bij voorkeur bij de documentalist(e) aanwezig moeten zijn.