Het ALADIN-project

De computer als hulpmiddel bij de literatuurattendering ten behoeve van de landbouwwetenschap

H.C.Molster



Tk In. 772

Het ALADIN-project

41027

De computer als hulpmiddel bij de literatuurattendering ten behoeve van de landbouwwetenschap

H.C.Molster

Centrum voor Landbouwpublikaties en
Landbouwdocumentatie
Wageningen - 1977

ISBN 90 220 0648 4

© Centrum voor Landbouwpublikaties en Landbouwdocumentatie (Pudoc), Wageningen, 1977

Inhoud

Voorwoord 9

1 Inleiding en verantwoording 11
1.1 Inleiding 11
1.2 Verantwoording 12
1.2.1 Het faseringsschema 12
1.2.2 De bijsturingsnota 14
1.2.3 Eindrapportage 16

2 Historisch overzicht 17

3 De gebruikte magneetbanden met literatuurinformatie 23
3.1 Het CAIN-bestand 23
3.2 Het CAB-bestand 26
3.3 Het FSTA-bestand 28
3.4 Het AGRINDEX-bestand 29

4 De technische uitvoering van het project 32
4.1 Beschrijving en ontwikkeling van de gebruikte programmatuur bij IWIS-TNO 32
4.2 Produktie 36
4.3 Vraagverwerking op aanvullende bestanden bij andere instellingen 38
4.3.1 NOCI 38
4.3.2 BioMedical Documentation Center (BMDC), sinds 1 november 1975 Medical Information Center (MIC) 38
4.3.3 Andere contacten 39
4.4 Het opstellen van de zoekprofielen 41

5 Manuele en computerattendering vergeleken 44
5.1 Inleiding 44
5.2 De mate waarin met computer-attendering dezelfde titels gevonden worden als met manuele attendering 45
5.3 Tijdsverschillen in signalering tussen de manuele en computer-attendering en computerdiensten onderling 48
5.4 Andere aspecten bij de vergelijking van manuele en computer-attendering 52
5.4.1 Aard van het uitgangsmateriaal 52

5.4.2 Keuze van computerbestanden 53
5.4.3 Verdere voor- en nadelen van beide systemen 55
5.4.3.1 Verkrijgbaarheid van de originele documenten 55
5.4.3.2 Precisie 56
5.4.3.3 Speciale eisen aan de opdracht 57
5.5 Conclusies 57

6 *Waardering door de gebruikers* 59
6.1 Beoordeling van de relevantie en de nieuwswaarde van de gevonden titels 59
6.1.1 Het belang van de relevantie-parameter 60
6.1.2 De resultaten van de relevantiebeoordeling 61
6.1.3 De resultaten van de nieuwswaardebeoordeling 64
6.2 Continuering nadat betaling werd gevraagd 66
6.2.1 Continueringspercentages voor de verschillende bestanden 66
6.2.2 De continueringspercentages als tevredenheidsindicator 67
6.3 Redenen voor beeindiging 68

7 *Kosten en baten van een operationeel systeem* 70
7.1 Inleiding 70
7.2 De huidige produktiekosten 71
7.2.1 Computer- en abonnementskosten per profiel 71
7.2.2 Personele kosten 73
7.2.3 Totale kosten computer-attendering 74
7.3 Kosten manuele attendering 76
7.4 Algemene overwegingen bij financiering en prijsstelling 77
7.5 De tarieven voor literatuurattendering bij Pudoc 78

8 *Evaluatie* 82
8.1 De technische uitvoerbaarheid 82
8.2 De keuze tussen een computersysteem voor literatuurattendering en manuele attendering 82
8.3 De behoefte aan literatuurattendering bij de landbouwwetenschap 83
8.3.1 Het nut van een literatuurattenderingsdienst voor landbouwwetenschappelijk onderzoek 83
8.3.2 Gebruikerservaringen met de attenderingsdienst 84
8.3.3 Gebruikerscategorieën 85
8.4 Belemmeringen voor de externe literatuurattendering 87
8.4.1 De houding van de landbouwwetenschappelijke onderzoeker tegenover literatuuronderzoek 87
8.4.2 Financiële en organisatorische belemmeringen 88
8.5 Een eigen bestand met literatuurinformatie 89
8.5.1 Een nieuwe bibliografie op het gebied van de landbouw 90
8.5.2 De mogelijkheden voor het in computer-leesbare vorm brengen van bestaande Nederlandse bibliografieën op het gebied van de landbouw 90

8.5.3 Het samenvoegen van bestaande bestanden 91
8.6 Toekomstige ontwikkelingen 92
8.7 Slotopmerking 94

9 Samenvatting en conclusie 95

Lijst met gebruikte afkortingen 98

Verklarende woordenlijst 99

Bijlagen 101

Voorwoord

Pudoc heeft, overeenkomstig de doelstelling in de statuten, het steeds tot zijn taak gerekend bij te dragen aan de informatievoorziening ten behoeve van de landbouw en de landbouwwetenschap. In dit kader passen toepasbaarheidsonderzoeken van nieuwe methoden en technieken. De invoering van de computer in de literatuurinformatieverwerking gaf bij uitstek aanleiding tot een dergelijk onderzoek. Het initiatief hiertoe werd genomen door dr.ir. D.J. Maltha.

Het was een gelukkige omstandigheid dat de in 1970 tot stand gekomen Commissie voor de Landbouwkundige Documentatie en Informatie TNO, COLADIN, het brede forum bood om een voorstel uit te werken om de uitvoerbaarheid van literatuurattendering met behulp van computersystemen te testen. Niet minder gelukkig was het dat mede door de inbreng van COLADIN het Nederlands Orgaan voor de Bevordering van de Informatieverspreiding, NOBIN, bereid gevonden werd een onderzoek naar de inrichting en de toepasbaarheid van een dergelijke dienst te subsidiëren. Hierdoor werd het mogelijk dat in 1972 aan Pudoc opdracht werd gegeven een onderzoek ter hand te nemen. Het project kreeg de verwachtingsvolle naam mee: ALADIN, Automatisering Landbouwkundige Documentatie en Informatie in Nederland. Het resultaat van het onderzoek is in dit rapport neergelegd. De uitkomsten van de proefneming zijn van dien aard dat het Ministerie van Landbouw en Visserij hierin voldoende aanleiding zag om deze nieuwe methode van literatuurinformatie, onder de hoede van Pudoc, als operationeel systeem te doen voortzetten. De snelle uitbreiding van het aantal abonnees op deze dienst lijkt intussen de verwachtingen niet te beschamen. Wel moet worden geconstateerd dat er, met name bij de onderzoeksinstellingen van het Ministerie materiële, immateriële en wellicht structurele belemmeringen bestaan, die optimale benutting van de geboden mogelijkheden beperken. In het rapport zelf wordt hier nader op ingegaan.

De voortgang van het project en de totstandkoming van het eindrapport hebben vertraging ondervonden. De projectleider, ir. H.C. Molster, kon pas in april 1973 in dienst treden. In juni 1974 aanvaardde hij de functie van beleidsmedewerker bij Pudoc. Pas in november 1974 kon een nieuwe projectleider, drs. J. van den Burg, worden aangesteld. Wegens ziekte van drs. Van den Burg moest ir. Molster hem reeds na enkele maanden, naast zijn nieuwe taken, langdurig vervangen. Het eindrapport verschijnt, mede hierdoor, uiteindelijk ruim twee jaar na de afsluiting van het ALADIN-project. De vertraging in de verschijning heeft het voordeel dat ook ervaringen met het operationele systeem konden worden verwerkt. Het rapport is in de huidige vorm aan NOBIN aangeboden. Vergelijkingen met soortgelijke onderzoekingen en verwijzingen naar de literatuur zijn niet opgenomen.

Daar het rapport de ervaringen bij het opzetten van een nieuwe vorm van literatuurinformatie beschrijft en met name uitvoerig ingaat op de kosten van een gemechaniseerde attenderingsdienst, menen wij dat het nuttig is aan dit rapport een verspreiding te geven die breder is dan de kring van de direct bij het project betrokkenen.

mr. A. Rutgers, directeur Pudoc

drs. A.H.H.M. Mathijsen, voorzitter COLADIN

1 Inleiding en verantwoording

1.1 INLEIDING

Dit rapport bevat het verslag van de werkzaamheden, ervaringen, resultaten en de daarop gebaseerde conclusies, voortvloeiende uit het ALADIN-project. Het acronym betekent: Automatisering van Landbouwkundige Documentatie en Informatieverspreiding in Nederland. Het project betreft een evaluatie van een literatuurattenderingsdienst met behulp van computersystemen ten behoeve van het wetenschappelijk onderzoek in de landbouw. Het werd uitgevoerd door het Centrum voor Landbouwpublikaties en Landbouwdocumentatie (Pudoc).

Het rapport is allereerst bedoeld voor de subsidiegeefster, de Stichting NOBIN (Nederlands Orgaan voor de Bevordering van de Informatieverzorging), en het meest betrokken Ministerie, namelijk dat van Landbouw en Visserij.

Het inschakelen van de computer als hulpmiddel bij de literatuurselectie heeft een tweeledige achtergrond: enerzijds dienen zich nieuwe technologische mogelijkheden aan, die oorspronkelijk voor andere doeleinden zijn ontwikkeld, anderzijds wordt de gebruiker van wetenschappelijke literatuur geconfronteerd met een steeds groter wordende stroom literatuur, waarbij het steeds moeilijker wordt zelf al datgene te selecteren dat van belang is voor het eigen studie-onderwerp. Om de selectie van relevante literatuur te verbeteren en zo de gebruiker tijd te besparen kan de computer goede diensten bewijzen; die conclusie wordt althans bevestigd door de ervaringen opgedaan met dit project.

In dit rapport wordt vooral het gebruik van de computer bij de *literatuurattendering* onder de loep genomen. Onder literatuurattendering wordt verstaan het regelmatig op de hoogte houden van literatuur die op een bepaald gebied recent verschenen is. Attendering kan beschouwd worden als een onderdeel van literatuuronderzoek in meer algemene zin. Tot dat laatste behoort ook het opsporen van oudere literatuur (retrospectief literatuuronderzoek) en het vastleggen van de gegevens uit de literatuur, dat kan uitmonden in een literatuurrapport.[1]

Ook bij het opsporen van oudere literatuur kan de computer veel goede diensten bewijzen. Waar in dit rapport over het gebruik van de computer bij het literatuuronderzoek in algemene zin wordt gesproken, wordt dan ook meestal aan beide toepassingen gedacht.

Na deze inleiding volgt allereerst een beschrijving van de ontwikkelingen die tot de conceptie van het project leidden, met daarop aansluitend een historisch overzicht van het verloop van het project (hoofdstuk 2).

---

1. Zie voor een uitgebreide beschrijving: D.J. Maltha, 1972. Literatuuronderzoek en schriftelijk rapporteren. Pudoc, Wageningen.

In hoofdstuk 3 wordt een beschrijving gegeven van de technische uitvoering van de attenderingsdienst met de computer als hulpmiddel. Dit gedeelte bevat informatie over de inhoud van de gebruikte magneetbanden, de computerprogrammatuur waarmee deze magneetbanden werden doorzocht, de ervaringen met de verwerkingsinstellingen waarvan gebruik werd gemaakt en tenslotte een overzicht van de produktie.

In hoofdstuk 4 komt de vergelijking tussen de manuele attenderingsdienst van Pudoc en de attendering met behulp van het computersysteem aan de orde. Daarvoor werden de resultaten gebruikt van 25 literatuurinteresseprofielen, waarvoor met beide systemen naar relevante titels werd gezocht. In dit hoofdstuk wordt ook het opstellen van deze profielen besproken.

In hoofdstuk 5 volgen de kwantitatieve en kwalitatieve evaluatie van het systeem en de door de gebruikers geleverde resultaten.

Hoofdstuk 6 geeft een overzicht van de kosten van het systeem en een schatting van de te verwachten kosten bij een operationeel systeem voor het gehele landbouwwetenschappelijk onderzoek.

In hoofdstuk 7 wordt nader ingegaan op de kosten, de financiering en de prijsstelling van een operationeel systeem.

Hoofdstuk 8 bevat een aantal conclusies waartoe de ervaringen met het project aanleiding geven. Het bevat een beschouwing over de behoefte aan literatuurattendering bij de landbouwwetenschap, de belemmeringen voor externe literatuurattendering, de mogelijkheden voor de ontwikkeling van een eigen bestand met literatuurinformatie en de vooruitzichten voor de komende jaren, waarbij ook de noodzakelijk geachte veranderingen op financieel-organisatorisch terrein worden aangegeven.

1.2 VERANTWOORDING

In de volgende paragrafen zal nagegaan worden in hoeverre dit eindrapport uiteindelijk informatie bevat over al die onderwerpen die bij het begin van het project in het werkprogramma worden genoemd, met een aanduiding waar de desbetreffende informatie is te vinden. De belangrijkste documenten waarin het werkprogramma wordt omschreven, zijn het 'faseringsschema'[2] en de latere 'bijsturingsnota' [3].

1.2.1 Het faseringsschema

Het oorspronkelijk werkplan, zoals dat is neergelegd in het zgn. faseringsschema, veronderstelde drie fasen bij de uitvoering, nl.

Fase I - software-produktie,
Fase II - experiment met literatuurattenderingsdienst met behulp van de computer,
Fase III - eindrapport.

---

2. Fasering van het project Automatisering Landbouwkundige Documentatie en Informatieverspreiding in Nederland (ALADIN). TNO/72/24.
3. Bijsturingsnota ALADIN-project, waarin opgenomen de gewijzigde begroting 25 oktober 1973. (Ir. H.C. Molster, Pudoc, Wageningen 1973).

De ontwikkeling van de computer-programmatuur die in Fase I moest worden uitgevoerd, wordt besproken in hoofdstuk 3 van dit rapport.

In Fase II zouden de literatuurinteresseprofielen worden opgesteld en bijgestuurd, werden de resultaten beoordeeld, werden de resultaten van het gehele experiment geïnventariseerd, en zou worden nagegaan wat de mogelijkheden zijn om 'de computer retrospectief te laten zoeken'. Tevens moest over de voortgang van het project tussentijds aan NOBIN worden gerapporteerd.

De technische uitvoering van het project wordt in hoofdstuk 4 besproken. De resultaten komen in hoofdstuk 6 aan de orde.

De mogelijkheden voor het gebruik van de computer bij het retrospectief literatuuronderzoek [4] zijn in het kader van dit project niet nader onderzocht. Dit gebruik heeft in de projectperiode reeds een zo grote vlucht genomen, met name door het snel inburgeren van computereindstations die direct met computers kunnen 'converseren', dat de praktijk het nut van een onderzoek naar de mogelijkheden hiervan intussen heeft achterhaald. De mogelijkheden zijn er, alleen de toepasbaarheid in een bepaalde omgeving moet steeds opnieuw worden bezien. Pudoc kreeg reeds in 1974 door middel van een computereindstation aansluiting op de grote internationale informatienetwerken. Deze aansluiting werd direct door het Ministerie van Landbouw en Visserij gefinancierd en liep buiten het ALADIN-project om. In het specifieke kader van het ALADIN-project is daarom het retrospectief literatuuronderzoek niet meer aan de orde geweest.

Tussentijdse rapportering aan NOBIN heeft regelmatig plaatsgevonden. In het totaal werden aan NOBIN 14 kwartaalverslagen als 'voortgangsrapporten' en één financieel eindrapport aangeboden.

Volgens het faseringsschema moet het eindrapport uitmonden in een 'gefundeerde beschouwing over de *exploiteerbaarheid* van een centrum voor geautomatiseerde informatieverspreiding voor de agrarische en veterinaire wetenschappen, op een economisch en kwalitatief verantwoorde wijze'. Speciale aandacht zou ook moeten worden besteed aan de kosten en de mogelijkheid tot 'cost recovery' van een literatuurattenderingsdienst. Deze vraagstukken komen respectievelijk in hoofdstuk 6 en 7 aan de orde.

Wat betreft de statistische gegevens legt het faseringsschema de nadruk op precisie en relevantie van de geleverde titels (naast de financiële gegevens). De uitkomsten wat betreft deze parameters komen aan de orde in hoofdstuk 5.

---

4. Retrospectief literatuuronderzoek is het opsporen van zoveel mogelijk relevante literatuur over een bepaald onderwerp, die op enig eerder tijdstip is verschenen. Het is eenmalig en heeft meestal betrekking op een vrij beperkt onderwerp. Als men met een nieuw onderzoek start, zal men meestal eerst een retrospectief literatuuronderzoek (laten) doen om te weten te komen wat er over het onderwerp reeds bekend is, om zich daarna door middel van een attenderingsabonnement op de hoogte te laten houden van nieuwe literatuur die over het onderwerp verschijnt

1.2.2 De bijsturingsnota

De bijsturingsnota werd eind oktober 1973 aan de begeleidingscommissie van het project aangeboden. Het was het eerste produkt van de full-time projectleider, die in april 1973 in dienst was getreden. In deze nota wordt ten aanzien van de doelstelling de nadruk gelegd op de evaluatie van de 'mogelijkheden en de bruikbaarheid van en de behoefte aan een literatuurattenderingsdienst met behulp van magneetbanden'. Deze doelstelling is daarmee in zeker opzicht ruimer dan de vraag naar de 'exploiteerbaarheid van geautomatiseerde informatieverspreiding', zoals die in het faseringsschema wordt gesteld, maar het gaat nu specifiek om een 'literatuurattenderingsdienst' en niet meer om een 'centrum voor geautomatiseerde informatieverspreiding' (ook al is die omschrijving onzuiver).

In de bijsturingsnota worden daarbij nog een aantal nieuwe punten genoemd waar aandacht aan zou moeten worden besteed. Zo wordt er onder '3.1 Doel' gesteld dat er een vergelijking van de resultaten tussen manuele attendering en attendering met behulp van een computer moet plaatshebben. De uitkomsten van deze vergelijking worden in hoofdstuk 5 gegeven.

Ook moet de verhouding worden bepaald tussen de door het systeem geleverde titels en die welke de gebruiker zelf reeds langs andere weg vond. Dit is ten dele geschied door de gebruikers naar de nieuwswaarde van de gevonden titels te vragen. Hierover wordt in hoofdstuk 5 gerapporteerd. Slechts in een incidenteel geval is echter nagegaan hoerveel titels de gebruiker zelf nog vond die de attenderingsdienst miste, en waardoor. De redenen waarom dit aspect verder geen aandacht kreeg waren zowel van methodologische als van technische aard, terwijl ook de beschikbare tijd een rol speelde.

Het was van meet af aan duidelijk dat een literatuurattenderingsdienst, zeker indien geheel op computers en bestaande bibliografische bestanden op magneetbanden wordt vertrouwd, nooit in staat zal zijn om alle, of zelfs maar bijna alle documenten op te sporen die relevante informatie voor een bepaald onderzoek of interessegebied bevatten. De relevante informatie kan namelijk gemakkelijk 'verborgen' zijn in een artikel waarvan de titel te algemeen is om direct duidelijk te maken dat het deze specifieke informatie bevat, terwijl dan ook de eventueel toegevoegde trefwoorden of classificatiecodes dat niet direct doen vermoeden. Er is vanuit het project dan ook steeds gesteld dat een attenderingsabonnement het lezen van de belangrijkste vaktijdschriften en wellicht een paar zeer hoog gekwalificeerde meer algemene wetenschappelijke tijdschriften nooit kan en wil vervangen. Dat de dekking zeker geen 100% kan zijn, behoeft dus niet nog eens experimenteel te worden aangetoond.

Een methodologisch probleem werd gevormd door de tegenstrijdigheid die kan ontstaan als men enerzijds van de gebruikers vraagt gedurende het experiment gewoon door te gaan met het zelf verzamelen van literatuur zoals men gewend was, en anderzijds van hen wil weten of men zich verder op de attenderingsdienst wil verlaten als literatuurinformatiebron, naast het lezen van enkele belangrijke primaire (vak-)tijdschriften. Als namelijk deze vraag aan het eind van het experiment gesteld wordt, heeft iemand die tijdens het experiment beide methoden naast elkaar gebruikte, nog geen ervaring gehad met een situatie waarin hij zich inderdaad goeddeels verliet op een attenderingsdienst. Daarbij bleek ook dat lang niet alle proefpersonen over een kaartsysteem beschikken dat op precies hetzelfde

onderwerp was gericht als het literatuurinteresseprofiel. Het zou nog erg veel arbeid hebben gekost in al de privé-kaartsystemen van de onderzoekers die titels te selecteren die eigenlijk ook door de attenderingsdienst hadden moeten worden gevonden. Alle titels die alleen vanwege een zijdelingse belangstelling in het kaartsysteem zijn opgenomen, of die onderwerpen betreffen waarover men 'ook wel iets wil weten' of althans enkele referenties wil hebben, zouden moeten worden geëlimineerd. Ten slotte is het dan gewoon tijdgebrek dat het onmogelijk heeft gemaakt zo'n vergelijking verder uit te voeren. In de bijsturingsnota wordt ook gesteld dat de bereidheid bij te dragen in de kosten van de dienstverlening moet worden onderzocht. Dit is, zij het op kleine schaal, gebeurd; er wordt in hoofdstuk 7 over bericht. Voor de toekomst is de mate waarin de deelnemers aan het experiment er uiteindelijk toe overgingen betalend abonnee te worden in feite veel belangrijker. Deze cijfers komen in hoofdstuk 5 aan de orde.

De deelnemers aan het experiment zouden ook worden geënquêteerd over hun ervaringen met de dienstverlening van het project. Jammer genoeg kon ook dit slechts gedeeltelijk plaatsvinden. In maart 1974 werd een eerste enquête gehouden; deze was er vooral op gericht na te gaan welke respons kon worden verwacht indien betaling werd gevraagd voor een abonnement op de attenderingsdienst (6.2). Tevens werd hierbij gevraagd naar het belang van het afdrukken van de titel in de oorspronkelijke taal, het referaat, de naam en het adres van de auteur op de titelkaartjes (3.1 en 3.3).

In het voorjaar van 1975 werd een enquête gehouden onder diegenen die géén abonnement wilden nemen op de attenderingsdienst na afloop van de kosteloze experimentele periode. De resultaten van deze enquête worden nader in hoofdstuk 5 besproken. Zij geven een heel verdeeld beeld te zien, zonder dat er duidelijke tekortkomingen van het systeem uit blijken. De computerattendering is echter niet de panacee voor alle informatiebehoeften; er zullen natuurlijk altijd allerlei situaties blijven waarbij deze dienstverlening niet het geschikte middel is om aan informatie te komen.

De voorgenomen enquête onder degenen die hun abonnement voortzetten is uiteindelijk wegens tijdgebrek niet gehouden. Het moet echter worden betwijfeld of een dergelijke enquête veel essentiële nieuwe gegevens zou hebben opgeleverd. De ervaring met de enquête onder opzeggers (6.3) heeft geleerd dat voor betrouwbare kwantitatieve gegevens een zeer grote steekproefruimte nodig is gezien de grote variatie onder de gebruikers, hun huidige informatievoorziening en hun houding ten aanzien van informatieverkrijging.

De ervaringen die werden opgedaan door de intensieve informele contacten met de gebruikers, zijn op vele plaatsen in het rapport verwerkt. De kwalitatieve gegevens van een eventuele enquête zouden daaraan waarschijnlijk weinig toe- of afgedaan hebben.

Het marktonderzoek dat tenslotte in de bijsturingsnota nog wordt genoemd, moest achterwege blijven wegens gebrek aan tijd. Het staat bij Pudoc nog steeds hoog op de prioriteitenlijst.

Onder '3.2.4 Gebruikers' van de Bijsturingsnota wordt nog gesteld dat de gebruikers zullen worden geïnstrueerd over het zelf opstellen van literatuurinteresseprofielen. Dit is niet gebeurd, omdat gedurende de uitvoering van het project de mening is ontstaan dat het opstellen van dergelijke profielen toch een vrij grote ervaring vereist. Allereerst moeten gebruikers zich de mogelijkheden van het zoekprogramma en van de diverse elementen van de gebruikte bibliografische bestanden, zoals categoriecodes, eigen maken. Daarnaast

moeten zij ervaring opdoen in het samenstellen van een zo compleet mogelijke verzameling van termen, die een optimale zekerheid biedt dat alle titels gevonden worden die relevante informatie bevatten. Voor het vinden van in het profiel op te nemen termen is weliswaar intensief overleg met de gebruiker nodig, maar de informatiespecialist blijkt, als hij eenmaal begrepen heeft waar het om gaat, veelal beter de synoniemen en verwante termen te kunnen aandragen dan de gebruiker. Aangezien de meeste gebruikers in hun leven slechts enkele profielen voor een attenderingsabonnement zullen behoeven te schrijven, zullen zij in het algemeen niet de ervaring kunnen opdoen die voor het opstellen van een goed profiel nodig is.

1.2.3 Eindrapportage

De vertraging die het gereedkomen van het eindrapport heeft ondervonden, is voor een groot deel te wijten aan moeilijkheden bij de personele bezetting. Ik werd per 1 juli 1974 beleidsmedewerker bij Pudoc, terwijl de opvolger drs. J. van den Burg al snel na indiensttreding langdurig ziek werd. Nadat het financiële eindrapport in juni 1976 klaar was gekomen was er eindelijk tijd om aan het onderhavige verslag te beginnen. Diverse concepten hebben nadien voor commentaar bij de COLADIN-leden gecirculeerd, voordat de definitieve versie werd samengesteld.

In deze vorm is dit rapport allereerst bedoeld als een verslag van de werkzaamheden en de resultaten. Voor anderen dan de direct betrokkenen kan lezing van belang zijn om met de ervaringen die hierin zijn neergelegd hun voordeel te doen bij het opzetten van soortgelijke projecten. Het ligt in de bedoeling elders kortere beschouwingen te publiceren alsmede enige artikelen in de wetenschappelijke pers over onderdelen van dit project en de verkregen resultaten. Deze zullen ook een vergelijking met ervaringen en resultaten elders bevatten.

2 Historisch overzicht

Op 6 december 1968 werden de mogelijkheden voor mechanisering van landbouwkundige documentatie en informatie reeds besproken in een daartoe door de Algemeen Directeur Landbouwkundig Onderzoek van het Ministerie van Landbouw en Visserij opgerichte werkgroep. Dit overleg mondde uit in een nota getiteld 'Landbouwkundige projecten voor de informatieverspreiding', die de Algemeen Directeur voornoemd op 25 juni 1969 aanbood aan de Werkgroep Wetenschappelijke en Technische Informatie[5]. Hierin staat centraal 'de bevordering van de mechanische verwerking van de documentaire informatie op het gebied van de landbouw'. Als eerste vereiste daarvoor wordt gezien het aantrekken van een jong landbouwkundig ingenieur, die zou moeten worden opgeleid tot deskundige op het gebied van de informatieverspreiding. Hij zou met name kennis moeten nemen van de ontwikkelingen bij de National Agricultural Library (NAL) in de V.S., waar het mechaniseringsproject CAIN (CAtaloguing and INdexing) toen in statu nascendi verkeerde.

In de nota van 1969 wordt opgemerkt dat de Bibliography of Agriculture 'binnen vijf jaar in de vorm van magneetbanden ter beschikking zal komen'. Dit bleek echter een zeer voorzichtige uitspraak: de eerste banden werden reeds één jaar later, in 1970, verspreid.

De in de eerste alinea genoemde werkgroep kreeg in de Commissie LAndbouwkundige Documentatie en INformatie (COLADIN) van de Nationale Raad Landbouwkundig Onderzoek TNO (NRLO-TNO) een meer permanente opvolger. COLADIN werd op 9 oktober 1970 geïnstalleerd. Deze commissie kreeg tot taak 'de activiteiten op het gebied van de landbouwkundige documentatie en informatie te coördineren en te stimuleren en tevens op te treden als overlegorgaan voor internationale projecten waarbij Nederland betrokken is of wordt'.

COLADIN heeft vanaf haar eerste vergadering de mechanisatie van de landbouwkundige documentatie en informatie met voorrang behandeld. Men heeft ook van meet af aan hard gezocht naar een tot informatiespecialist op te leiden landbouwkundig ingenieur, maar het heeft tot april 1973 geduurd voordat deze pogingen met succes werden bekroond.

COLADIN heeft zich aanvankelijk vooral beziggehouden met het opstellen van een nota over de behoeften aan mechanisatie voor de landbouwkundige documentatie en informatie. Deze nota werd op 16 april 1971 aan NRLO-TNO aangeboden. Er wordt in geconcludeerd dat het noodzakelijk is een nader onderzoek te doen naar de uitvoerbaarheid van mechanische verwerking van landbouwkundige informatie in Nederland, waartoe allereerst een toepasbaarheidsonderzoek zal moeten worden verricht. Hierbij zullen de bestaande documentatie- en informatiesystemen op het gebied van de landbouw in Nederland moeten worden geanalyseerd.

5. Deze werkgroep werd voorgezeten door prof.dr. C.J.F. Böttcher. De werkgroep adviseerde tot de instelling van een centraal instituut voor de informatieverzorging. Dit advies resulteerde in de oprichting in 1971 van de Stichting NOBIN.

De analyse van de toenmalige situatie werd verricht door middel van een enquête onder de bibliotheken van instellingen voor landbouwkundig onderzoek. Een kort verslag van de resultaten van deze enquête is toegevoegd als bijlage 1. Binnen het kader van het toepasbaarheidsonderzoek wilde men daarna zo snel mogelijk het gebruik van banden van CAIN en FSTA (Food Science and Technology Abstracts) toetsen 'om na te gaan of resultaten van mechanische attendering in gunstige of ongunstige zin afwijken van die welke volgens de conventionele methoden worden verkregen' (4e COLADIN-vergadering, 27 juli 1971). Deze concrete voorstellen waren mede mogelijk geworden doordat de magneetbanden van het CAIN-bestand vanaf 1970 beschikbaar waren gekomen en die van FSTA vanaf 1969.

Op 25 oktober 1971 bood COLADIN aan NOBIN een nota aan getiteld 'Voorstellen inzake automatisering van de documentatie en de informatieverspreiding in Nederland, op het gebied van de agrarische en veterinaire wetenschappen'. Het projectvoorstel was vervat in een bijlage, getiteld 'Voorstel experimenteel current-awareness-service op CAIN-tapes en op magneetbanden van FSTA'. Aan NOBIN werd gevraagd of het de aanloopfase van dit project gedeeltelijk zou kunnen financieren. NOBIN accepteerde dit voorstel getuige zijn schrijven d.d. 30 november 1971. De doelstelling van het project wordt daarin als volgt omschreven:

'Gebruikmakend van het free-language-search programme van IWIS-TNO zal worden nagegaan in hoeverre de CAIN-tapes en de magneetbanden van Food Science and Technology Abstracts bruikbaar zijn voor een current awareness service. Zo mogelijk zal op beperkte schaal ook retrospectief literatuuronderzoek worden verricht op de magneetbanden van Food Science and Technology Abstracts. De beoordeling zal geschieden door antwoorden van het computersysteem te vergelijken met de resultaten van conventioneel (manueel) onderzoek. Het project zal resulteren in een rapport waarin de uitkomsten van het vergelijkend onderzoek kwalitatief en kwantitatief zullen worden beoordeeld en waarin een uitspraak zal worden gedaan over de toepasbaarheid van een dergelijk computersysteem voor de landbouwkundige informatie in Nederland.'

In de bijsturingsnota werd de doelstelling later samengevat als 'Het evalueren van de mogelijkheden en de bruikbaarheid van en de behoefte aan een literatuurattenderingsdienst met behulp van magneetbanden voor gebruikers, die ressorteren onder de vakgebieden welke vertegenwoordigd zijn in de NRLO-TNO'.[6]

De naam 'ALADIN' komt voor het eerst voor in een schrijven van COLADIN aan NOBIN d.d. 21 december 1971. Het door NOBIN gevraagde faseringsschema en de gedetailleerde begroting werden op 18 januari 1972 aan NOBIN aangeboden (zie 1.2.1). Het project zou worden uitgevoerd door het Centrum voor Landbouwpublikaties en Landbouwdocumentatie (Pudoc) te Wageningen, als de meest voor de hand liggende instelling met rechtspersoonlijkheid. COLADIN zou optreden als begeleidingscommissie.

De toenmalige directeur van Pudoc, ir. D.J. Maltha, nam in het begin zelf de uitvoe-

6. Met literatuurattendering wordt hier bedoeld het regelmatig selecteren van nieuwe publikaties die relevant lijken voor een specifiek onderzoeks- of interessegebied. Lopende bibliografieën, die op magneetband verkrijgbaar zijn, kunnen met behulp van een computer worden afgezocht op relevante titels. In het volgende hoofdstuk wordt op de techniek hiervan nader ingegaan.

ring van het project ter hand, daarbij gesteund door het hoofd van de manuele attenderingsdienst van Pudoc, ir. C.D. Voogd. Op 28 maart 1972 werden de eerste testprofielen gedraaid op banden van Chemical Abstracts. Deze banden werden reeds in opdracht van de Nederlandse Organisatie voor Chemische Informatie (NOCI) in Den Haag bij het Instituut voor Wiskunde, Informatieverwerking en Statistiek TNO (IWIS-TNO) verwerkt. Het lag daarbij in de bedoeling dat de CAIN- en FSTA-banden voor het ALADIN-project bij hetzelfde computercentrum en met dezelfde programmatuur te verwerken.

Vanaf mei 1972 werden iedere 14 dagen de resultaten van de verwerkte banden van Chemical Abstracts voor enige literatuurinteresseprofielen geëvalueerd. Ir. T. Eernstman, literatuuronderzoeker bij Pudoc, belastte zich vanaf juni 1972 met dit werk; hij stelde tevens trefwoordenlijsten voor een aantal nieuwe profielen samen.

Op 11 september 1972 werd de eerste 'eigen' band voor het ALADIN-project bij IWIS-TNO verwerkt, namelijk de FSTA-band van juli 1972. Op 21 september 1972 volgde de eerste CAIN-band (ook van juli 1972) en op 8 november 1972 de eerste band van de 'BoA'-versie[7] van het CAIN-bestand. In september kwam het eerste testresultaat van Biological Abstracts ter beschikking, door bemiddeling van NOCI verwerkt in Stockholm (zie 3.3.2).

De eerste tijd bleef het aantal profielen klein, vooral vanwege het uitblijven van de aanstelling van een projectleider. Toen deze tenslotte op 1 april 1973 in dienst kwam, liepen er echter toch reeds verscheidene maanden 12 profielen. (De resultaten van 11 hiervan werden door ir. Maltha meegedeeld op het eerste Regionale Europese IAALD Symposium, mei 1973, te Wageningen.)[8] Daarna kon het aantal profielen betrekkelijk snel worden uitgebreid, hetgeen in het volgende hoofdstuk nader wordt beschreven.

Door de zich wijzigende omstandigheden en de eerste ervaringen van de projectleider werd in overleg met COLADIN een bijsturingsnota opgesteld als aanvulling op het faseringsschema. Deze nota ging vergezeld van een gewijzigd begrotingsvoorstel. De nota is gedateerd 25 oktober 1973 en werd NOBIN op 1 november 1973 aangeboden. Na enig overleg werden de bijsturingsnota en de begroting in principe door NOBIN aanvaard.

In de bijsturingsnota worden als belangrijkste punten van 'bijsturing' genoemd: het beschikbaar komen van meer landbouwkundige bibliografische bestanden op magneetbanden, nl. alle referattijdschriften van de Commonwealth Agricultural Bureaux (CAB) en de nieuwe wereldomvattende bibliografie van de landbouw, AGRINDEX, van de Food and Agriculture Organization (FAO). Het gebruik van deze banden zou ook in het project moeten worden opgenomen. Daarnaast kon door een beter inzicht in de kosten een betere begroting worden opgesteld. Deze kwam aanzienlijk hoger uit dan de oorspronkelijke begroting: ƒ 500.500 tegenover ƒ 266.000. De verhoging werd grotendeels door twee posten veroorzaakt: de kosten voor

7. Deze werd geproduceerd door McMillan Information, New York, de uitgever van de Bibliography of Agriculture, welke gebaseerd is op het CAIN-bestand. De gegevens zijn echter anders gestructureerd op de band vastgelegd en het betreft een selectie uit het CAIN-bestand, o.a. zonder monografieën.
8. D.J. Maltha, 1973. Testing a scheme for machine-retrieval of information in agricultural science. In: Progress and prospects in agricultural librarianship and documentation. Papers presented at the Regional European Symposium of the International Association of Agricultural Librarians and Documentalists 14-18 May 1973, Wageningen, the Netherlands: 43-63. Pudoc, Wageningen.

een projectleider en de computerkosten. Nu de projectleider officieel bij Pudoc en dus bij het Ministerie van Landbouw en Visserij in dienst kwam, moest ook de norm voor een zgn. Z-formatieplaats worden gehanteerd, die aanzienlijk boven de oorspronkelijke raming lag. De computerkosten werden hoger, doordat een jaar lang twee verschillende versies van de CAIN-band naast elkaar werden verwerkt en doordat de periode waarin magneetbanden voor het project verwerkt werden, verlengd werd van $1\frac{1}{2}$ tot $2\frac{1}{2}$ jaar. Het verwerken van de CAB- en AGRINDEX-bestanden stond daarbij nog pro memorie op de nieuwe begroting, omdat de exacte datum waarop deze beschikbaar zouden zijn nog niet bekend was.

Een ander belangrijk punt uit de bijsturingsnota betrof het betrekken van eindgebruikers bij het project: naast de vergelijking met de manuele attendering zouden er voor maximaal 80 deelnemers met een zo goed mogelijke spreiding over het gehele terrein van de landbouwwetenschap (inclusief diergeneeskunde) literatuurinteresseprofielen worden opgesteld. De leden van deze experimentele gebruikerskring zouden gedurende de looptijd van het project de uitvoer voor hun profiel gratis ter beschikking krijgen, waarbij als tegenprestatie aan hen een beoordeling van de relevantie en de nieuwswaarde van de titels werd gevraagd. In het faseringsschema werd de beoordeling van de uitkomsten impliciet aan de projectleider overgelaten. In de bijsturingsnota werd tevens duidelijk gesteld dat, naast de eigen verwerking van CAIN, FSTA, CAB en AGRINDEX, profielen voor andere data bases elders konden worden verwerkt, overal waar het onderwerp zulks wenselijk maakte.

Ten slotte werd in de bijsturingsnota het hele project meer toegespitst op het *evalueren* van een *attenderingsdienst* met behulp van computersystemen dan op onderzoek naar de mogelijkheid tot *exploitatie* van een *centrum* voor geautomatiseerde informatieverspreiding, zoals dat in het faseringsschema stond aangegeven. Over de keuze van het te gebruiken computercentrum bleef nog enige tijd onzekerheid bestaan.

In het voorjaar van 1973 verzocht NOBIN de projectleider na te gaan of het financieel te verantwoorden was dat de computerverwerking ook voor de resterende tijd van het project door IWIS-TNO zou worden verzorgd. Er werden daartoe inlichtingen gevraagd bij 14 instellingen, waarvan er 6 uiteindelijk een offerte uitbrachten. De samenvatting van de offertes werd op 1 november 1973 aan NOBIN aangeboden. Bij vergelijking van de offertes bleek dat er enkele instellingen waren die lagere variabele kosten berekenden dan IWIS-TNO, maar de dan vereiste investeringen konden nooit binnen de periode van het project worden terugverdiend. Daarom werd besloten IWIS-TNO voor het resterende gedeelte van het project als verwerker te handhaven (brief NOBIN 9 april 1974).

Volgens de bijsturingsnota zou het project in april 1975 beëindigd moeten worden. Het eindrapport zou in de eerste drie maanden van 1975 moeten worden geschreven. De evaluatie van de verkregen resultaten zou tot eind 1974 moeten voortduren. Gedurende het schrijven van het eindverslag zou de computerverwerking van nieuwe banden wel moeten doorlopen op kosten van het project. De financiering zou na 1 april 1975 moeten worden overgenomen door het Ministerie van Landbouw en Visserij (MvLV), indien de resultaten van het project daartoe tenminste voldoende aanleiding zouden geven. Het was trouwens een informele voorwaarde geweest voor subsidiëring van het project dat in voortzettings-financiering kon worden voorzien bij een gunstig resultaat van het onderzoeksproject. Er was daartoe dan ook, met gunstig resultaat, overleg met de Directie Landbouwkundig Onderzoek van het MvLV gevoerd. De procedure die was voorzien was echter niet logisch sluitend: in

feite had de produktie op kosten van het project moeten kunnen doorlopen tot het moment waarop redelijkerwijs een beslissing over het operationeel maken van de dienstverlening kon worden genomen op basis van het eindverslag van het project. Dit moment zou nog wel 3-6 maanden na het verschijnen van het eindrapport moeten liggen. Gelukkig heeft de directeur van Pudoc het MvLV er al in een veel vroeger stadium van kunnen overtuigen dat het ALADIN-project een zaak was die waard leek gecontinueerd te worden. Voor 1975 werd Pudoc daartoe een speciaal krediet toegekend. In 1976 werd de 'computerattendering' een vast begrotingshoofd.

In januari 1975 kwam, na de experimentele band van 1974, de eerste operationele AGRINDEX-band ter beschikking, evenals de eerste CAB-band. Voor een vergelijking tussen deze algemeen landbouwkundige bestanden en CAIN was het intussen in het kader van het ALADIN-project te laat geworden. Er werden toen door Pudoc initiatieven ontwikkeld om evaluaties van deze nieuwe bestanden te doen plaatsvinden in het kader van het programma voor landbouwkundige informatievoorziening van de Europese Commissie. Voor wat betreft de vergelijking tussen AGRINDEX en CAIN werd tussen Pudoc en de Europese Commissie op 10 december 1975 een contract getekend voor de computerverwerking van 200 profielen met deze bestanden. Het opstellen van de profielen en het evalueren van de uitkomsten was voor de verantwoordelijkheid van de Duitse zusterinstelling van Pudoc: ZADI (Zentralstelle für Agrardokumentation und Information in Bonn-Bad Godesberg). Het eindrapport van dit project wordt in april 1977 verwacht.

Een vergelijking tussen CAB en CAIN kon uiteindelijk niet bij de EC worden ondergebracht. Er was hiervoor echter wel interesse bij de British Library. Er werd een plan opgesteld voor een gezamenlijk Brits-Nederlands project; bij NOBIN werd een aparte subsidie aangevraagd voor een bijdrage in het Nederlandse aandeel. NOBIN wees dit verzoek af omdat zo'n evaluatie 'niet meer geacht kon worden tot de experimentele fase (van een attenderingsdienst) te behoren'. NOBIN rekende het in het algemeen niet tot zijn taak subsidie te verlenen voor kwaliteitsonderzoek van nieuwe bestanden op magneetband. Gelukkig werd het Ministerie van Landbouw en Visserij bereid gevonden uiteindelijk het gehele Nederlandse aandeel te financieren. De Engelse projectleidster kwam op 1 augustus 1976 in dienst van Wye College (London University), dat als contracthouder optreedt, terwijl Pudoc weer verantwoordelijk is voor het verwerken van de profielen. Het eindrapport van dit 'CAB/CAIN-vergelijkingsproject' wordt begin 1978 verwacht.

In afwachting van deze studies leek het echter van belang in het kader van het ALADIN-project reeds met beide banden ervaringen op te doen. Bij het opstellen van de begroting was hier, zoals reeds genoemd, slechts door middel van posten pro memorie rekening gehouden, maar de (ongewilde) besparingen op de post 'projectleider' door het enige tijd ontbreken van deze functionaris (zie hierna) gaven nog enige financiële ruimte. Het huren en verwerken van het gehele CAB-bestand voor 1975 was financieel niet (meer) mogelijk. Uiteindelijk werd besloten van CAB alleen het diergeneeskundige gedeelte te gebruiken, omdat dit gebied door het CAIN-bestand slecht wordt gedekt (20e COLADIN-vergadering, 11 november 1974).[9]

9. Hoewel de eerste volledige CAB-band voor het CAB/CAIN-vergelijkingsproject pas per 1 januari 1977 verwerkt is, kon Pudoc toch reeds voor heel 1976 dit bestand huren. Gedurende de eerste drie maanden van 1976 ontvingen alle abonnees van de computer-attenderingsdienst CAB-uitvoer voor hun profiel gratis, waarna gevraagd werd of men een CAB-abonnement tegen betaling wilde voortzetten.

Over de ervaringen met de bestanden van CAB en AGRINDEX wordt in hoofdstuk 3 nader bericht.

De fase waarin magneetbanden voor het ALADIN-project werden verwerkt, werd dus voor wat betreft CAIN en FSTA afgesloten op 1 april 1975, maar voor CABVET (het veterinaire deel uit CAB) en AGRINDEX pas in begin 1976, nl. met de laatste banden van 1975. NOBIN ging in zijn brief van 3 oktober 1975 akkoord met de verlenging van de projectduur in deze zin.

3 De gebruikte magneetbanden met literatuurinformatie[10]

De ingebruikname van door een computer gestuurde zetmachines bij het vervaardigen van drukwerk heeft er oorspronkelijk voor gezorgd dat er ook bij het vervaardigen van titelbibliografieën en referaattijdschriften als tussenprodukt magneetbanden ontstonden, waarop de af te drukken tekst is vastgelegd. Ook de mogelijkheid om met behulp van een computersysteem tekst te rangschikken en automatisch indexen te maken gaf een extra stimulans tot het gebruik van magneetbanden bij het vervaardigen van dergelijke tijdschriften. Er wordt op deze wijze een bestand gevormd, waarmee met name bij accumulatie van de ingevoerde titelbeschrijvingen op allerlei wijzen gemanipuleerd kan worden. Bij bestanden die pas recent ontstaan zijn, is het vormen van het bestand vaak primair en wordt de gedrukte versie als afgeleid produkt beschouwd. Dit is onder andere bij het AGRIS-systeem van de FAO het geval. In 3.4 zal dit uitgebreid worden besproken.

Een magneetband die gemaakt is om een zetmachine te sturen, is niet zonder meer bruikbaar om met behulp van een computer-zoeksysteem selectie te plegen in de erin opgeslagen tekst. Daartoe is het nodig de informatie op de band te vertalen, zodat het gebruikte computer-zoeksysteem ermee aan het werk kan. Dit is er de oorzaak van dat versies die direct bruikbaar zijn voor computer-zoeksystemen soms pas aanzienlijk later zijn ontstaan dan de magneetband voor de zetmachine. Met name bij de Commonwealth Agricultural Bureaux is dit het geval geweest.

3.1 HET CAIN-BESTAND

De eerste magneetband die ter beschikking kwam op het gebied van de landbouw in het algemeen, was die van de National Agricultural Library (NAL) van het Amerikaanse Ministerie van Landbouw. Dit bestand is het produkt van haar eigen 'CAtaloguing and INdexing System' (waarvan het acroniem CAIN is afgeleid). Met behulp hiervan worden twee gedrukte bibliografieën vervaardigd, nl. de 'Bibliography of Agriculture' (in gedrukte vorm uitgegeven door Macmillen Information, sinds 1975 door Oryx Press) en de 'National Agricul-

10. In dit hoofdstuk wordt enige kennis omtrent het principe van het doorzoeken van informatie op een magneetband met behulp van een computerzoekprogramma bekend verondersteld. In het volgende hoofdstuk wordt in detail ingegaan op de werkwijze van het zoekprogramma dat bij IWIS-TNO wordt gebruikt. Tot een goed begrip van het onderhavige hoofdstuk diene hier alleen het volgende voorbeeld. Stel dat er 10.000 titels in willekeurige volgorde op een magneetband staan en men wil alleen alle titels afdrukken die betrekking hebben op het *oogsten van maïs*, dan zal men het computerprogramma bij iedere titel laten nagaan of de titel over maïs gaat en, als dat het geval is, of het dan ook over oogsten gaat. Daartoe zou in de titel het woord MAIS, dan wel MAYS, dan wel MAIZE of CORN moeten voorkomen, met daarnaast HARVEST, HARVESTER of HARVESTING. De genoemde trefwoorden in hun logisch verband (a_1 of a_2 of a_3 of a_4 EN OOK b_1 of b_2 of b_3) wordt het '(zoek)profiel' genoemd.

tural Library Catalog' (in gedrukte vorm uitgegeven door Rowman and Littlefield). De CAIN-magneetband wordt vanaf begin 1970 maandelijks geproduceerd. Het bestand is in eerste instantie een bibliotheekhulpmiddel, dat de beschrijvingen van alle primaire documenten bevat die in de collectie van de NAL worden opgenomen. Het bestand onderscheidt zich echter van een 'gewone' bibliotheekcatalogus, doordat het de meeste artikelen in de opgenomen tijdschriften apart beschrijft. Er komen per jaar ongeveer 140.000 titels bij.

Hoewel het CAIN-bestand de grootste bibliografie is die het hele gebied van de landbouwwetenschap bestrijkt, is de dekking zeker niet volledig: de totale produktie van wetenschappelijke landbouwkundige literatuur wordt wel op 250.000 titels per jaar geschat. De NAL is daarbij nationaal depot voor de Amerikaanse landbouwkundige literatuur. De actieve acquisitie is daarnaast gericht op de typische gebruikerskring van de bibliotheek. Bovendien wordt nog veel opgenomen dat de NAL ongevraagd wordt toegezonden en dat relevant lijkt. Het heeft nooit in de bedoeling van de NAL gelegen alle landbouwliteratuur van de gehele wereld te dekken. Dit is dan ook één van de achtergronden van het AGRIS-systeem van de FAO, dat straks aan de orde komt.

Het CAIN-bestand bevat van iedere opgenomen titel behalve de bibliografische beschrijving (inclusief de titel in de oorspronkelijke taal, en een Engelse vertaling) een aantal extra informatie-elementen. Zo wordt de titel in 15% van de gevallen verrijkt met enkele woorden die het informatiegehalte van de titel verhogen. (Als uit de titel niet blijkt dat het artikel bijvoorbeeld één specifiek gewas behandelt, of over de situatie in één land gaat, wordt de naam van het gewas of het land aan de titel toegevoegd.) Verder wordt door middel van een categoriecode aangegeven bij welk hoofdonderwerp het artikel is in te delen. Er zijn 70 beschikbare codes. (In de gedrukte versie worden de titels naar deze categoriecodes gerangschikt.) Voor elk document wordt meestal één categoriecode gebruikt, maar het kunnen er ook 2 zijn (in 10% van de gevallen). Verder wordt onder andere nog de oorspronkelijke taal van het document aangegeven, staat erbij of het een uitgebreide bibliografie bevat en of het een overzichtsartikel (review) betreft. Ook wordt het bibliotheeknummer van de NAL (volgens het systeem van de NAL-Library of Congress) opgegeven, hetgeen het direct opvragen van (fotocopieën van) de oorspronkelijke documenten bij de NAL zeer vergemakkelijkt. In tegenstelling tot vele andere bibliografieën is men er bij dit bestand zeker van dat de NAL het desbetreffende document ook bezit en dat een kopie ervan steeds kan worden verkregen.

Een CAIN-'record'[11] bevat dus wel iets meer informatie dan alleen de bibliografische beschrijving; met name de categoriecodes kunnen bij de selectie een belangrijke rol spelen. Anderzijds is men bij het zoeken toch hoofdzakelijk aangewezen op selecteren uit de titelwoorden. Bij het formuleren van een 'zoekprofiel' moet men zich afvragen welke titelwoorden een auteur zou kunnen hebben gebruikt bij het formuleren van een titel voor een artikel dat de zoeker interesseert, zulks enigszins in tegenstelling tot systemen met een veel diepergaande classificatie of met toegevoegde trefwoorden. In het geval van een zgn. gestructureerde thesaurus worden deze geput uit een tevoren opgestelde woordenlijst met hiërarchische structuur.

---

11. Een 'record' is het totaal aan informatie, in computer-leesbare vorm vastgelegd, die betrekking heeft op één document.

Men kan uit de titel alleen echter niet altijd opmaken of het artikel inderdaad van belang zal zijn; de aanwezigheid van een referaat maakt voor de gebruiker de selectie veel eenvoudiger. Het belang van referaten voor computerselectie is echter een omstreden punt (zie 3.3).

De Macmillan-versie van de CAIN-banden Van de CAIN-magneetbanden heeft enige tijd nog een afgeleide versie bestaan, die ook apart werd geëvalueerd. Deze versie werd op de markt gebracht door de toenmalige uitgever van de Bibliography of Agriculture, Macmillan Information. Op de band van Macmillan (verder in dit rapport de BoA-band genoemd) was alle informatie weggelaten die niet van belang was voor het zoeken naar titels voor literatuurinformatie en het vervaardigen van de gedrukte versie. Ook bevatte hij geen monografieën of nieuwe titels van series, omdat deze niet in de Bibliography of Agriculture werden opgenomen. Ten slotte werd bij een niet-Engelse publikatie alleen de vertaalde Engelse titel vermeld. De band was daardoor 'schoner' en gemakkelijker leesbaar. Daarnaast werd voor de index een speciaal programma gebruikt dat titelwoorden waar nodig verving door geprefereerde woorden. Dit programma zorgde ervoor dat bijvoorbeeld de titelwoorden 'pig', 'pigs', 'swine', 'hog' en 'hogs' allemaal onder 'pigs' terug te vinden waren. De andere genoemde woorden en woordvormen kwamen in de index niet voor. Aan een titel die 'swine' bevatte was op de band ook daadwerkelijk 'pigs' toegevoegd, zodat ook bij het zoeken met de computer in feite alleen 'pigs' als zoekwoord behoefde te worden opgegeven. Dit kon het zoeken natuurlijk vereenvoudigen; welk percentage termen van een CAIN-profiel overbodig werd, is echter niet geanalyseerd. Men moest van iedere term controleren òf hij wel weggelaten mocht worden, hetgeen heel wat arbeid vereiste. Uit een vergelijkende studie van beide versies[12] bleek dat onze CAIN-profielen op de BoA-band maar 1% extra relevante titels opleverde, titels dus die op de CAIN-band niet werden gevonden - een te verwaarlozen percentage. Er was echter bij de CAIN-profielen ook zoveel mogelijk voor gezorgd alle noodzakelijke woordvormen en synoniemen op te nemen.

Na in 1973 beide versies naast elkaar te hebben verwerkt, werd besloten in 1974 verder alleen nog de oorspronkelijke NAL-versies te gebruiken. De belangrijkste redenen hiervoor waren:
- het ontbreken van de titels in de oorspronkelijke taal bij niet-Engelse[13] publikaties (dit maakte identificatie van het oorspronkelijke document soms moeilijk),
- het ontbreken van monografieën (verlies aan relevante titels $\pm$ 10%)[14],
- het prijsverschil ($540 voor CAIN tegenover $1000 voor BoA).

12. H.C. Molster, 1974. Differences between the CAIN and Bibliography of Agriculture tapes. Quarterly Bulletin of the International Association of Agricultural Librarians and documentalists (IAALD) 19(2):111-115.

13. Het afdrukken van de titel in de oorspronkelijke taal werd bij navraag onder de gebruikers door geen enkele 'essentieel' geacht, door één 'erg handig maar niet essentieel' en door 12 'onbelangrijk'. Het probleem van het terugvinden van het originele document laten zij graag aan anderen over!

14. Uit de gehouden enquête onder gebruikers bleek dat 8 van de 13 respondenten monografieën 'erg belangrijk' vonden, 4 'soms van belang maar niet essentieel' en 1 'onbelangrijk'.

De enigszins gecontroleerde woordenlijst van de BoA-band had o.i. in zijn toenmalige vorm slechts geringe voordelen. Deze zouden echter groter kunnen worden als er in deze thesaurus meer structuur gebracht zou worden.[15] De exploitatie van de BoA-band werd in 1975 geheel gestopt toen de uitgave van de gedrukte versie van Macmillan Information overgenomen werd door Oryx Press.

3.2 HET CAB-BESTAND[16]

Naast het CAIN-bestand dekt het totale bestand van de Commonwealth Agricultural Bureaux de landbouwwetenschap ook in grote mate, terwijl het daarnaast nog gegevens bevat uit raakgebieden met de geneeskunde. Het CAB-bestand wijkt echter op een aantal essentiële punten af van CAIN: hij bevat meestal vrij uitgebreide referaten; de staf speurt intensief naar relevante primaire bronnen en er wordt streng geselecteerd bij de opname van titels om een hoge kwaliteit van het gerefereerde materiaal te handhaven. (Er is door CAB wel gesteld dat zij van 90% van alle relevante literatuur op de hoogte is, maar daarvan slechts 50% opneemt. De referaattijdschriften van CAB, zoals 'Field Crop Abstracts' en 'Soils and Fertilizers', genieten een hoog aanzien en worden in gedrukte vorm ook aanzienlijk meer gebruikt dan de Bibliography of Agriculture (zie Bijlage 1). Voor het merendeel van de gebruikers is deze voorselectie wellicht een voordeel; er wordt echter ook literatuur gemist die voor sommige gebruikers juist wel van belang kan zijn. Het bestand als geheel is echter nog weinig geïntegreerd: het zijn in feite 19 aparte bestanden, die gecentraliseerd tot die 19 referaattijdschriften worden verwerkt. Het materiaal wordt echter op de diverse bureaus van CAB geselecteerd en van toegevoegde trefwoorden, coderingen en referaten voorzien. Met name de systemen voor toegevoegde trefwoorden en coderingen zijn van bureau tot bureau verschillend.

Dat het bestand allereerst voor het vervaardigen van referaattijdschriften wordt gebruikt, waarbij nog nauwelijks rekening wordt gehouden met de eisen die computer-zoekprogramma's aan zo'n bestand stellen, blijkt bijvoorbeeld ook uit het feit dat niet alle 'records' een volledige bibliografische beschrijving bevatten. Dit kan voorkomen bij een reeks artikelen, die alle uit eenzelfde verzamelwerk komen en in het referaattijdschrift alle onder elkaar staan afgedrukt, zodat voor de bibliografische gegevens naar de eerste titel uit de reeks kan worden verwezen. Als echter met het zoekprofiel de eerste titel niet gevonden wordt en latere titels uit de reeks wèl (bijvoorbeeld omdat de eerste titel te algemeen is, en dus niet de specifieke trefwoorden waarmee gezocht werd bevat), krijgt de gebruiker niet de beschikking over de bibliografische gegevens en zijn de gevonden titels dus moeilijk voor hem op te sporen (tenzij hij de gedrukte versie alsnog raadpleegt).

---

15. De woordenlijst is nu gepubliceerd als 'Thesaurus of Agricultural Terms' (Oryx Press, 1976). Deze woordenlijst bevat géén structuur. Wel zijn voorkeurstermen met woordvarianten en een aantal verwijzingen opgenomen.
16. In het kader van het ALADIN-project werd alleen het veterinaire bestand van de CAB verwerkt. Het totale bestand werd pas in 1976 door Pudoc gebruikt, na beëindiging van de NOBIN-subsidie (zie pag. 21). Om een zo volledig mogelijk beeld te geven van de nu gebruikte landbouwkundige bestanden wordt hier ook aan het volledige CAB-bestand de nodige aandacht besteed.

Voor een bestand dat voor een computer-zoekprogramma gebruikt wordt, is het noodzakelijk dat iedere titel op zichzelf staat, zonder dat naar informatie elders behoeft te worden verwezen.

Een andere 'kinderziekte' betreft het nog niet beschikbaar zijn van een lijst met verklaringen van alle coderingen (deze bestaat in sommige gevallen alleen in handschrift op het desbetreffende bureau). De coderingen gaan in de meeste gevallen wel veel verder dan de 70 categorieën van CAIN. De classificatiesystemen zijn echter voor ieder tijdschrift verschillend. Een gebrek van het CAB-bestand is ook dat de dekking van de landbouw als geheel nog enkele hiaten vertoont; met name landbouwwerktuigen en gebouwen en veeteelt (behalve melkvee en fokkerij) worden niet gedekt. De verwerking van landbouwprodukten wordt hoofdzakelijk overgelaten aan Food Science and Technology Abstracts (waarin CAB participeert), maar daar valt de verwerking van niet-voedsel produkten weer buiten.

Een laatste belangrijk verschil met CAIN zijn de abonnementskosten van de band. Een titel (met referaat) op de CAB-band kost ca. 20 cent (zonder referaat 11 cent), terwijl één titel op de CAIN-band slechts 1,3 cent kost. Hierdoor en door de hogere computerverwerkingskosten van de langere records (aantal tekens per record) zijn de totale aan het bestand gebonden kosten voor de huur plus verwerking van de CAB-band, afhankelijk van het aantal profielen, 3 - 3,5 maal zo hoog als die voor CAIN.

Het volledige CAB-bestand is nu vanaf 1973 op magneetband voor het doorzoeken met behulp van een computersysteem (retrieval) beschikbaar, het diergeneeskundig bestand reeds vanaf 1972. De Agricultural Research Service van het Amerikaanse Ministerie van Landbouw was de eerste die in 1975 een abonnement op het gehele CAB-bestand nam, zij het alleen voor intern gebruik. Pudoc volgde in 1976 als tweede. De banden van de eerste jaren zijn pas later voor 'retrieval'-gebruik geschikt gemaakt. Het hele bestand is in begin 1977 d.m.v. Lockheed Information Systems in Palo Alto voor interactief terugzoeken met behulp van computereindstations ter beschikking gekomen, zodat Pudoc nu ook over deze mogelijkheid beschikt. (CAIN is al vanaf 1974 bij Lockheed beschikbaar, evenals bij Systems Development Corporation.)

In 1975 werd door Pudoc alleen het diergeneeskundige gedeelte van de band verwerkt (Index Veterinarius en Veterinary Bulletin). De uitvoering daarvan verliep zonder veel problemen. De enige tekortkoming van dit beperkte bestand is dat Veterinary Bulletin wèl categoriecodes heeft, terwijl die bij Index Veterinarius ontbreken. Voor het totale diergeneeskundige bestand zijn deze codes daardoor weinig bruikbaar. Bij dit bestand worden als unicum binnen CAB trefwoorden uit een gestructureerde woordenlijst toegevoegd[17], al wordt deze niet altijd consequent toegepast.

In 1974 werd een gedeelte van het diergeneeskundige bestand tijdelijk voor interactief gebruik d.m.v. computereindstations ter beschikking gesteld voor gebruikersevaluatie. Ook Pudoc maakte hier gebruik van. Onze bevindingen werden in aan apart rapport[18]

17. Gepubliceerd als 'Veterinary Subject Headings, compiled by Roy Mack. Commonwealth Agricultural Bureaux 1972'.

18. H.C. Molster, summer 1974. Some results from the use of the CABVET system. Pudoc, Wageningen.

vastgelegd. De bevindingen van de Engelse gebruikers werden gerapporteerd door T.D. Crane[19].

Toen Pudoc in 1976 het gehele CAB-bestand ging verwerken (behoudens enige onderdelen waarvoor weinig vragen verwacht werden), kwamen er vele problemen aan het licht zowel wat betreft het lezen als het gebruik voor retrieval van de band. De laatste kwamen vooral voort uit het gebrek aan integratie tussen de diverse bureaus en de geringe aandacht voor de speciale eisen die aan een retrieval-band gesteld worden. Daarnaast was de documentatie voor het lezen van de band ook uiterst gebrekkig. Uiteindelijk kon een bezoek aan de systeemanalist van IWIS-TNO aan CAB in oktober 1976 pas definitieve oplossingen voor de verschillende programmeringsproblemen brengen.

Op speciaal verzoek van de diergeneeskundige faculteit van de Rijksuniversiteit te Utrecht wordt gebruik gemaakt van het CAB-bestand inclusief referaten. Deze worden niet doorzocht maar wel afgedrukt. Hopelijk zal het CAB/CAIN-vergelijkingsproject (zie hoofdstuk 2) nadere informatie verschaffen over de waardering voor de referaten, welke eventueel in de prijs zou moeten worden ingecalculeerd. De abonnementsprijs voor de band mèt referaten is namelijk aanzienlijk hoger dan zonder referaten, terwijl deze de computerverwerkingskosten ook aanzienlijk verhogen.

3.3 HET FSTA-BESTAND

Food Science and Technology Abstracts komt tot stand door technische samenwerking van CAB en de 'Zentralstelle für Maschinelle Dokumentation' in Frankfurt.[20] In het Management Committee zijn naast CAB het Amerikaanse Institute of Food Technologists, het Duitse Institut für Dokumentationswesen en Pudoc vertegenwoordigd. Inhoudelijk wordt het bestand in Engeland samengesteld, terwijl de magneetband en het gedrukte tijdschrift in Duitsland worden geproduceerd. De eerste band kwam in januari 1969 gereed. In 1976 werden ongeveer 18.000 titels aan het bestand toegevoegd. Het totale bestand bevat nu ongeveer 123.000 titels.

FSTA bevat referaten en ook toegevoegde trefwoorden uit een gecontroleerde woordenlijst, die een min of meer hiërarchische structuur heeft. Dat dit het opstellen van een zoekprofiel in sommige gevallen erg kan vereenvoudigen is o.a. gebleken bij het profiel 'eetbare paddestoelen'. Bij FSTA kan hiervoor met één term worden volstaan: MUSHROOMS, terwijl bij CAIN (en CAB) alle soorten eetbare paddestoelen met name moeten worden genoemd; het aantal termen in het CAIN-profiel bedraagt 43. Tevens heeft FSTA een indeling in 19 hoofdgroepen, die in incidentele gevallen in zoekprofielen worden gebruikt.

In het begin werd bij FSTA het referaat doorzocht en afgedrukt. Het doorzoeken le leverde echter nogal wat extra irrelevante titels op, omdat nogal eens een volgens het zoekprogramma 'juiste' combinatie van trefwoorden in het referaat werd gevonden, terwijl

19. T.D. Crane, October 1974. Preliminary experiment with Index Veterinarius. Report to BLRDD en Project SI/G/101.

20. Voor een overzicht van het FSTA-systeem zie: E. Mann. The International Food Information Service - the first seven years. Paper read at the 5th World Congress of the International Association of Agricultural Librarians and Documentalists (IAALD) (Mexico DF, 14-18 April 1975).

deze woorden niet het gewenste logische verband met elkaar hadden. Door de extra informatie die de toegevoegde trefwoorden reeds bieden, werd verondersteld dat het doorzoeken van de referaten ook niet veel nieuwe relevante titels zou opleveren. Een Deens onderzoek hiernaar bevestigde dat echter niet.[21] Onzerzijds werd hiernaar echter geen kwantitatief onderzoek gedaan.

Het doorzoeken van de referaten is in 1974 gestaakt. Ook het afdrukken ervan werd stopgezet, omdat de meeste klanten de gedrukte versie toch bij de hand bleken te hebben. Uit een gehouden enquête bleek dat men in het algemeen het afdrukken van de referaten wel op prijs stelde, maar niet essentieel achtte.[22] Hopelijk levert de vergelijking tussen CAIN en CAB, die nu aan de gang is, meer gegevens over de waardering van referaten op, met name over de extra prijs die men ervoor over heeft. Mocht uit genoemde studie blijken dat men aan het afdrukken van referaten toch de voorkeur geeft, dan zal bij FSTA ook het referaat weer afgedrukt worden.

Op de FSTA-band is ook de instituutsnaam en -adres aanwezig. Nadat bij de enquête onder gebruikers bleek dat het afdrukken hiervan op prijs werd gesteld, werden deze gegevens later toegevoegd op de titelkaartjes.[23]

Van de gebruikte bestanden geeft FSTA qua inhoud de minste 'problemen': het vakgebied is duidelijk omschreven en wordt vermoedelijk goed gedekt (al is de overlapping met bijvoorbeeld CAIN niet onderzocht en wordt bij de huidige Pudoc-klanten FSTA eigenlijk nooit alléén gebruikt, maar meestal in combinatie met CAIN en/of Chemical Abstracts). Daarnaast is de informatie voor iedere titel relatief goed en geeft de verwerking verder geen problemen.

3.4 HET AGRINDEX-BESTAND

AGRINDEX is het eerste tastbare produkt van het nieuwe 'Agricultural Research Information System' (AGRIS) van de Food and Agriculture Organisation (FAO). Het systeem werd opgezet om de gebreken van de bestaande systemen voor landbouwkundige literatuurinformatie te ondervangen en ook om vooral de ontwikkelingsgebieden beter van deze informatie te voorzien, o.a. door deze landen zelf actief bij het systeem te betrekken. Over de organisatie en de merites van het systeem werden bij Pudoc twee nota's geschreven[24], terwijl

---

21. I.B. Hansen, 1974. Evaluation of Food Science and Technology Abstracts on tape. Internal Report, Documentation Dept., National Technological Library, Lingby, Denmark.
22. Bij de enquête onder de gebruikers over deze vraag antwoordde er slechts één dat het afdrukken van het referaat 'essentieel' voor hem was, negen 'erg handig maar niet essentieel' en drie 'onbelangrijk'. Gesuggereerd werd dat het direct bij de hand hebben van de gedrukte versie met referaten van grote invloed is op het belang dat gebruiker hecht aan het afdrukken van de referaten op de attenderingskaartjes.
23. Vier respondenten vonden het afdrukken van de auteursaffiliatie 'essentieel', zeven 'erg handig maar niet essentieel' en twee 'onbelangrijk'.
24. H.C. Molster, september 1975. AGRIS. Een beschouwing over het wereldomvattende Landbouw Informatie Systeem, mede naar aanleiding van het FAO Council document CL 66/12 van de 66e zitting in Rome, juni 1975. Pudoc, Wageningen.
H.C. Molster, juni 1976. AGRIS. Een evaluatie van het Agricultural Research Information System ten behoeve van de FAO Council bijeenkomst van juli 1976. Pudoc, Wageningen.

er ook reeds een uitgebreide literatuur over AGRIS bestaat[25]. De aandacht zal hier vooral worden gericht op de inhoud van de AGRINDEX-band, zoals die nu geproduceerd wordt.

AGRINDEX is in feite een titelbibliografie en in die zin dus het beste met CAIN te vergelijken. Op twee punten verschilt AGRINDEX nogal aanzienlijk van CAIN: de dekking naar land van publikatie is geheel anders en er zijn drie classificaties in plaats van één.

De dekking Ieder aan AGRIS deelnemend land is er verantwoordelijk voor dat van de wetenschappelijk van belang zijnde landbouwkundige literatuur die binnen zijn grenzen wordt geproduceerd (uitgegeven), een genormaliseerde beschrijving aan AGRIS wordt geleverd. Doet een bepaald land niet mee, dan ontbreekt de in dat land geproduceerde literatuur ook in het bestand. Er wordt namelijk (nog) geen poging gedaan om de literatuur uit de niet-deelnemende landen toch door de deelnemers te laten beschrijven. Gezien het feit dat er uit de V.S. nog steeds maar 1000 titels per maand komen, terwijl de werkelijke produktie veel hoger moet liggen, wordt de Amerikaanse literatuur momenteel slecht gedekt. Dit geldt ook voor de Russische. Daarentegen bevat AGRINDEX veel materiaal uit verscheidene ontwikkelingslanden en ook uit Japan, dat door andere lopende bibliografieën slecht wordt gedekt. De landen van de Europese Gemeenschappen leveren een belangrijk deel van de input. De literatuur uit deze landen wordt vrij volledig ingevoerd.

Het totaal aantal titels in AGRINDEX bedroeg in 1975 54.423; In 1976 77.015. Groei zit er duidelijk in, maar vooralsnog is de samenstelling vrij onevenwichtig. Gebleken is echter al wel dat AGRINDEX allerlei onverwacht interessant materiaal bevat, dat niet in CAIN voorkomt. Gezien de selectiecriteria van de CAB is het moeilijk een vergelijking te maken met betrekking tot de dekking van beide bestanden: CAB selecteert bewust, AGRINDEX wil echter juist *alle* artikelen opnemen die binnen zijn gebied (de 'scope') vallen. Het streven is zoveel mogelijk literatuur op te nemen die wetenschappelijk belang zou kunnen hebben.

De categorieën AGRINDEX heeft behalve algemene onderwerpscodes ook gewascodes en geografische codes. Van deze codes mogen er in feite zoveel worden toegevoegd als noodzakelijk lijkt, al wordt beperking aanbevolen. De indeling is iets fijner dan bij CAIN, hoewel de ideale indeling bepaald nog niet gevonden is. De toevoeging van een gewascode is een belangrijke verbetering ten opzichte van CAIN, omdat het hierdoor niet meer nodig is om van één gewas allerlei namen en schrijfwijzen op te nemen en het tevens erg gemakkelijk is om met één code een hele groep van gewassen te selecteren. (voor 'graangewassen' is bij CAIN altijd de hele lijst van graangewassen nodig; bij AGRINDEX kan men met de eerste twee cijfers van de 'object code' voor 'cereals' volstaan.)

Ten slotte is er een geografische code, zowel naar land als naar een aantal andere criteria (fysisch-geografische indeling, indeling naar intergouvernementale organisaties, naar klimaatzones, naar ontwikkelingsgebieden, etc.). Deze indelingen zijn wel hiërar-

25. Zie bijvoorbeeld Vol. 19 (1974) nr.1 van het Quarterly Bulletin of the International Association of Agricultural Librarians and Documentalists, dat geheel aan AGRIS is gewijd. Zie ook: G. Genie, 1975. Het AGRIS-project. Landbouwtijdschrift 28(3):497-511.

chisch, maar niet altijd op een consequente decimale wijze, zoals de geografische codes bij de UDC. Het is overigens onduidelijk waarom deze niet is overgenomen.

Als toegevoegde informatie kan een AGRINDEX-'record' dan nog woorden ter verrijking van de titel bevatten, zoals dat ook bij CAIN het geval is.

De opbouw van een AGRINDEX-'record' is betrekkelijk gecompliceerd: het aantal verschillende velden voor één 'record' kan in theorie enkele honderden bedragen. In de praktijk blijkt dat het voorkomen van het ene veld vaak het voorkomen van een ander veld uitsluit, terwijl ook vaak dezelfde informatie in verschillende velden voorkomt. (De titel van een jaarverslag is vrijwel gelijk aan de corporatieve auteur, aan het adres van de auteur en aan de naam van de reeks waarin het is verschenen.) Het geschikt maken voor efficiënt zoeken en afdrukken blijkt om deze redenen wel enige inventiviteit te vereisen.

Het hogere informatiegehalte maakt het AGRINDEX-bestand op zichzelf aantrekkelijker dan dat van CAIN, maar momenteel is de dekking van AGRINDEX minder evenwichtig dan die van CAIN, zodat men zich toch moeilijk alleen op AGRINDEX kan verlaten.

In de herfst van 1977 zal de FAO Conference besluiten of het AGRIS-project al dan niet moet worden voortgezet. Als men tot een positieve beslissing komt, lijkt één of andere vorm van integratie met CAIN onvermijdelijk (zoals ook met Nuclear Science Abstracts is gebeurd, toen het International Nuclear Information System - INIS - tot volle wasdom was gekomen). In hoeverre de CAB ook in het AGRIS-systeem kan en zal worden geïntegreerd, is vooralsnog niet te overzien. De geleverde produkten verschillen met name op twee punten: grote selectiviteit tegenover brede dekking en referaten met soms aanzienlijke vertraging tegenover geen referaten met (hopelijk) geringere vertraging. Deze verschillen lijken voorlopig groot genoeg om het bestaan van beide diensten te rechtvaardigen. De financiering door de gebruikers en door politici zal uiteindelijk de toekomst bepalen. Voor computerattendering is men althans in 1977 aangewezen op ten minste drie abonnementen, wil men van een goede dekking verzekerd zijn.

4 De technische uitvoering van het project

In begin 1971 werd het Instituut voor Wiskunde, Informatieverwerking en Statistiek - TNO (IWIS-TNO) betrokken bij de voorbereiding van het COLADIN-project voor mechanisatie en informatie. Dit instituut werd toen verzocht een werkschema op te stellen voor de eerste fase van de toepasbaarheidsstudie. IWIS-TNO beschikte reeds over een programmapakket voor de verwerking van magneetbanden met bibliografische gegevens voor literatuurinteresse-profielen, het zgn. FLIRT-programma (Free Language Information Retrieval Tool). Dit programma is ontwikkeld door drs. T. de Heer en J.R.E. Lieftinck. Drs. De Heer maakte vanaf 13 maart 1971 ook deel uit van COLADIN.

Het FLIRT-programma werd reeds voor het begin van het ALADIN-project door NOCI gebruikt voor de verwerking van Chemical Abstracts. De meeste wijzigingen, die later werden aangebracht, kwamen in overleg met NOCI tot stand en werden ook samen met deze organisatie gefinancierd.

4.1 BESCHRIJVING EN ONTWIKKELING VAN DE GEBRUIKTE PROGRAMMATUUR BIJ IWIS-TNO

Het FLIRT-programma beoogt het allereerst mogelijk te maken en na te gaan of in willekeurige teksten bepaalde tekencombinaties voorkomen om daarna de eenheid (record) die aan de gevraagde combinatie voldoet, af te zonderen, te rangschikken en eventueel af te drukken. Het feit dat in een tekst van willekeurige lengte, met een willekeurig aantal tekstelementen kan worden gezocht, onderscheidt 'free text'-programma's zoals FLIRT van een verwant maar eenvoudiger soort, waarbij alleen in enkelvoudige tekstelementen van vaste lengte kan worden gezocht (fixed format).

In feite worden door middel van het FLIRT-programma twee bestanden met elkaar vergeleken, namelijk het bestand met te doorzoeken teksten - in het onderhavige geval de bestanden zoals CAIN en AGRINDEX - en het bestand met de zoektermen van alle profielen. In het zoektermenbestand wordt een hiërarchische structuur aangebracht, waarna dit bestand vergeleken wordt met ieder 'record' van het tekstbestand totdat een overeenkomst wordt gevonden.

Voor iedere soort bibliografisch bestand zijn twee aparte besturingstabellen en een apart conversieprogramma nodig. Het conversieprogramma zorgt ervoor dat het oorspronkelijke 'format' wordt aangepast aan het interne FLIRT-'format'. Eén besturingstabel geeft aan welke velden doorzocht moeten worden en eventueel in welke combinatie. De tweede besturingstabel beheerst de wijze waarop de uitvoer wordt afgedrukt: wat de volgorde van de elementen moet zijn, wanneer een nieuwe regel moet worden begonnen, wanneer moet worden ingesprongen etc.

Sinds augustus 1973 worden de geselecteerde titels met bibliografische gegevens en eventueel met referaten afgedrukt op kaartjes (A06-formaat = 105 x 148 mm) in plaats van

op papiervellen, zoals aanvankelijk gebeurde. Deze overgang werd door de meeste gebruikers zeer gewaardeerd.[26] In Fig.1-5 zijn voorbeelden van kaartjes weergegeven zoals die bij dit programma worden geproduceerd.

Het FLIRT-programma kent sinds het begin twee *zoekmogelijkheden*, die ook in combinatie kunnen worden gebruikt. In het algemeen wordt een zoekprofiel samengesteld uit een aantal groepen van termen: de termen binnen de groepen zijn alternatieven. Tussen de groepen bestaat dan een EN OOK-relatie of een negatieve EN NIET. Tevens kan aan iedere term een gewicht worden toegekend. Hierdoor wordt bereikt dat alleen die 'records' worden afgedrukt waarvan de som van de gewichten der in de 'record' aangetroffen zoektermen groter dan of gelijk aan het aangegeven drempelgewicht is. Een voorbeeld van een zoekprofiel wordt gegeven in Fig.6.

Alle termen kunnen steeds rechts en/of links afgebroken (getrunkeerd) worden, waarbij het irrelevant is welke lettersequenties op de plaats van het trunkatie-symbool voorkomen.

Welke velden moeten worden doorzocht kan (per zoekterm) apart worden aangegeven (auteur, titel, codes, taal etc.). Termen uit verschillende zoekvelden kunnen direct met elkaar worden gecombineerd.[27]

Sinds de aanvang van het project vonden drie belangrijke wijzigingen in het programma plaats.

1. Sinds de zomer van 1974 wordt de CAIN-band in secties doorzocht. Dit houdt in dat voor iedere vraag aangegeven wordt in welke secties van de band gezocht moet worden, terwijl vroeger iedere vraag vergeleken werd met alle titels van de band. De indeling in 9 secties vindt plaats op grond van de categoriecodes die aan iedere titel zijn toegevoegd. Het doorzoeken per sectie betekent een belangrijke kostenverlaging. (Een veeteeltvraag bijvoorbeeld wordt nu alleen in de veeteeltsectie gezocht en niet in de bosbouwsectie etc.) De AGRINDEX- en CAB-bestanden werden later in dezelfde secties opgedeeld die voor CAIN werden gevormd. Het aantal doorzochte secties is voor CAIN gemiddeld 2,1, voor AGRINDEX en CAB 2,3.
2. In 1975 werden de zoekmogelijkheden uitgebreid tot (bijna) volledige Booleaanse logica met het ongelimiteerde gebruik van haakjes (nesting). Er bestaat alleen nog een maximum voor het aantal logische combinaties na herleiding, namelijk zes. Tevens werd het mogelijk gemaakt om bepaalde woordvormen over te slaan, terwijl de stam van het woord wel een zoekterm is (IGNORE facility).
3. De profielenbestanden voor de diverse bestanden werden samengevoegd. Oorspronkelijk was er een apart profielenbestand voor iedere band, waarbij iedere wijziging die voor meer dan één bestand gold ook steeds in ieder profiel apart moest worden doorgevoerd. Nu

---

26. Sommige gebruikers hadden liever het kleinere bibliotheekformaat 75 x 125 mm gezien, maar dit is technisch onmogelijk vanwege de vrij grote letters die de computer-regeldrukker gebruikt; er kunnen daardoor te weinig tekens op een klein kaartje. Hopelijk kan dit lettertype in de toekomst wel door een ander worden vervangen, dat prettiger leesbaar is, maar dat behoeft dan nog niet tot het gebruik van kleinere kaartjes te leiden.
27. Hierdoor is het bijvoorbeeld mogelijk *DUST* (rechts en links getrunkeerd) als zoekterm op te nemen en tevens de term INDUST* te vermijden. INDUSTRIAL DUST wordt dan wel gevonden, maar INDUSTRY alléén niet. Zonder deze IGNORE-mogelijkheid zou wellicht INDUST* als absolute uitsluiting zijn opgenomen, maar dan was een wellicht relevante titel met INDUSTRIAL DUST verloren gegaan.

CAIN VOL 77 NO 1 =BIBL OF AGRIC 42-02 DATE 01/03/77

215? 20 S 1 121 ZOUNDBLOOM $PCPFH

34 REFERENCES FOUND OUT OF 1?24? SEARCHED

TOTAL 115?8 REFERENCES IN THIS ISSUE

①

215? - 1

DELANEY, M E: WALKER, D A:

A RECONSTITUTED CHLOROPLAST SYSTEM FROM HELIANTHUS ANNUUS [SUNFLOWER].
(ENG)

OCT 1976
PLANT SCI LETT: 7 (4): 285-294. REF.
CAT CODE: 403500. NAL CALL NO: QK1.P5.
CAIN ID : 0112943.

②

CAIN VOL 77 NO 1 =BIBL OF AGRIC 42-02 DATE 01/03/77

215? - 12

DELANEY, M.E.: WALKER, D.A.
SHEFFIELD UNIV. (UK). DEPT. OF BOTANY

A RECONSTITUTED CHLOROPLAST SYSTEM FROM HELIANTHUS ANNUUS
(EN) (SUM.: EN)

PLANT SCIENCE LETTERS (NETHERLANDS)
OCT 1976
V. 7(4) P. 285-294
SUMMARY
NOTES: 23 REF.
CAT CODE: F6? 03J: *1337
SEQ NR: 135994 INPUT: NL

③

PUDOC - AGRINDEX VOL 77 NO 3 DATE 01/03/77

2150 - P

KONOPLEV, V. P.; ZUBAREV, I. G.; EGOROV, G. A.; ZOLOCEVSKIJ [ZOLOCHEVSKII], V. T.: UGHEBRISVILI [MYEBRISHVILI], T. V.: SCLIAEV [SHKLYAEV], E. V.; DILANIAN [DILANYAN], R. S.; ARUTIUNIAN [ARUTYUNYAN], N. S.;

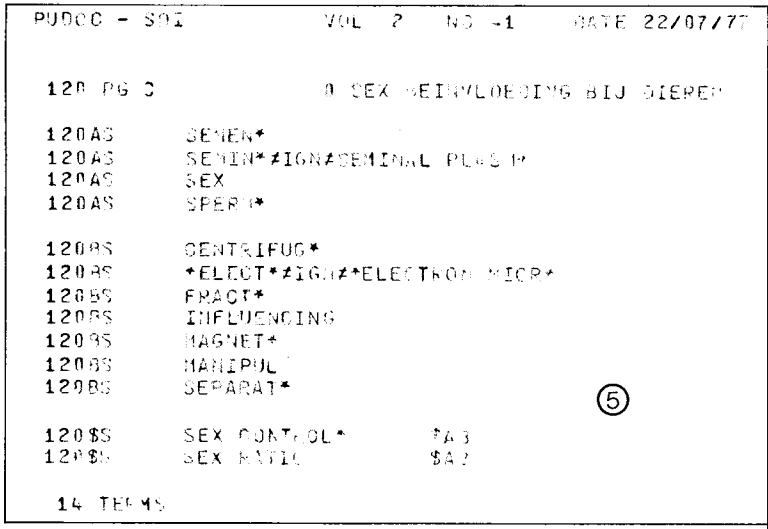
APPLICATION OF AN ELECTRIC FIELD IN REFINING SEED OIL MISCELLAE.
(IT ,SUM: EN)

*REFINING *ELECTRICITY *SUNFLOWER OILS
1976
OLI, GRASSI, DERIVATI: 12 (4) 29. [3 REF.]
[GROZ NII, USSR]
CAT CODE: N0113 .

④

PUDOC - ESTA VOL 77 NO 2 DATE 08/03/77

pudoc p.o. box 4 wageningen netherlands tel 08370 19548



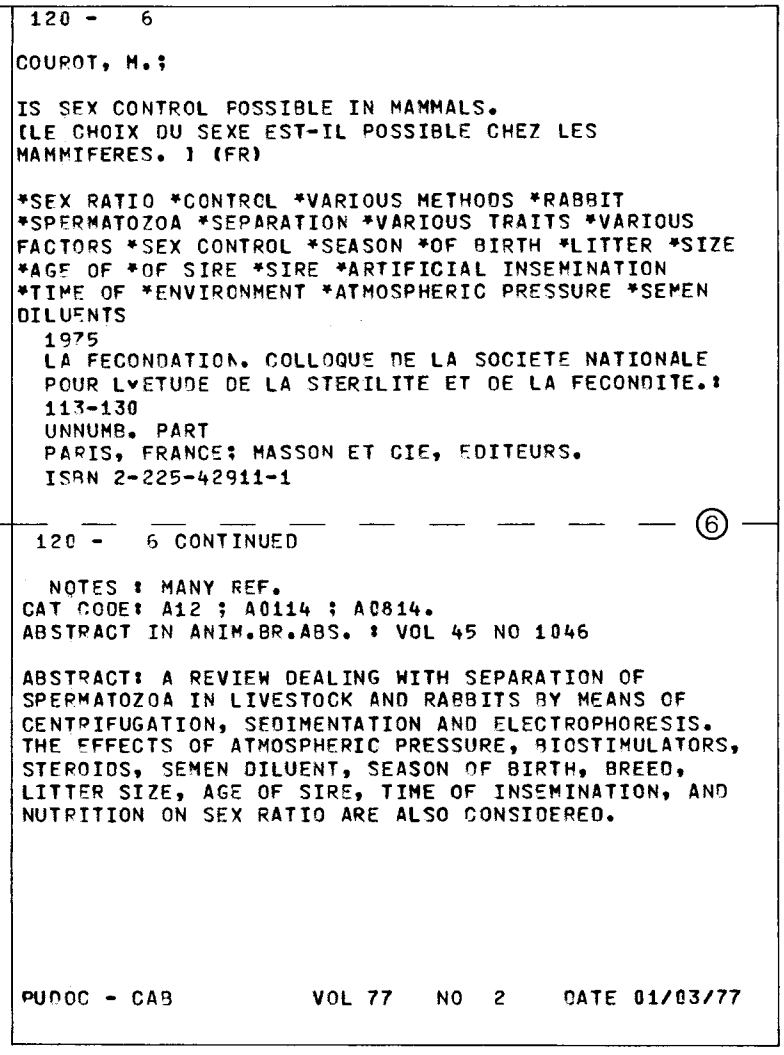


Fig.1. Voorbeeld van een CAIN-frontkaartje.

Fig.2. Voorbeeld van een CAIN-attenderingskaartje.

Fig.3. Voorbeeld van een AGRINDEX-attenderingskaartje (zelfde publikatie als CAIN-titel in Fig.2).

Fig.4. Voorbeeld van een FSTA-attenderingskaartje.

Fig.5. Voorbeeld van een profielopstelling.

Fig.6. Voorbeeld van een CAB-attenderingskaartje. (Deze treffer werd gevonden met het profiel van Fig.5 op de term SEX CONTROL.)

alle profielbestanden samengesmolten zijn, is slechts één wijziging nodig die voor alle relevante bestanden geldt. In de identificatiekaart van het profiel kan worden aangegeven welke bestanden met dat profiel moeten worden doorzocht.

Er bestaat een apart programma voor het aanvullen en wijzigen van het profielenbestand. Deze wijzigingen worden in het algemeen vóór het verwerken van een nieuwe band uitgevoerd, waarna eventuele fouten nog kunnen worden hersteld. (Het programma geeft bijvoorbeeld aan wanneer een te wijzigen term niet in het bestand kan worden gevonden; dit kan bijvoorbeeld voorkomen bij spelfouten.) Profielwijzigingen worden bij Pudoc op ponsformulieren gecodeerd en opgestuurd naar IWIS, waar ze geponst worden. In 1975 kwam bij IWIS een interactieve versie van het profielwijzigingsprogramma klaar, waarvan een gedeelte van de kosten ten laste van de begroting van het ALADIN-project kon worden gebracht. Hiermee is het nu mogelijk met behulp van een bij Pudoc opgestelde terminal direct profielwijzigingen door te voeren. Het gebruik van deze faciliteit lijkt echter nogal duur, zodat het tot nu toe beperkt blijft tot het verbeteren van fouten en het invoeren van wijzigingen die op het laatste moment binnenkomen. Nieuwe profielen en grote hoeveelheden wijzigingen worden nog steeds door IWIS geponst en verwerkt.

4.2 PRODUKTIE

In het kader van het project werden de volgende banden verwerkt:

CAIN	juli 1972	t/m maart 1975
BoA	juli 1972	t/m december 1973
FSTA	juli 1972	t/m maart 1975
AGRINDEX	januari 1975	t/m december 1975
CABVET	januari 1975	t/m december 1975

In Tabel 1 wordt het aantal profielen vermeld dat op een aantal peildata gebruikt werd. De getallen in deze tabel hebben betrekking op het aantal onderwerpen (intellectual concepts) waarvoor profielen waren opgesteld. Dit aantal is lager dan het aantal profielnummers, en wel om twee redenen. Ten eerste was het gedurende een bepaalde periode nodig om profielen die op méér dan drie maar niet op alle secties moesten lopen, 2 (of 3) maal in te voeren, omdat het maximum aantal specifieke secties dat in de titelkaart kon worden aangegeven, slechts 3 bedroeg. Deze beperking werd later gelukkig opgeheven. Ten tweede kan het efficiënter zijn één onderwerp op te splitsen in twee profielen, namelijk als de vereiste logica de vorm A EN OOK B OF, C EN OOK D had. Met invoering van volledige Booleaanse logica is ook dit nu niet meer nodig. (Zo'n geval kan nu als LOGIC 'AB + CD' worden aangegeven.) Ten slotte werden voor een aantal onderwerpen alternatieve profielopstellingen getest.

De peildata komen overigens niet overeen met de banden die die datum dragen. De verwerkingsdata van de CAIN- en FSTA-banden staan vermeld in Bijlage 2. De *datum van de band* heeft voor de verschillende banden een andere betekenis. Zo is dit voor CAIN de maand waarvan de invoer-produktie nog op de band is opgenomen; de band zelf komt echter pas in de loop van de daaropvolgende maand ter beschikking. Theoretisch moet het mogelijk zijn hem dan ook nog in die maand te verwerken, maar dat gelukt bijna nooit. Sommige CAIN-banden hebben zelfs zo lang op zich laten wachten, dat de verwerking vaak pas 3-4 maanden

Tabel 1. Aantal profielen van 8 bestanden dat op een aantal peildata in behandeling was.

Peildatum	CAIN	BoA	FSTA	CAC	BA	MED	AGRINDEX	CAB
1-10-72	9	-	4	3	3	-	-	-
20- 1-73	9	9	4	6	5	3	-	-
12- 4-73	9	9	4	6	5	3	-	-
2- 6-73	52	52	15	13	3	6	-	-
1-10-73	56	56	15	13	3	6	-	-
1- 1-74	69	69	22	24	3	4	-	-
1- 4-74	98	-	26	33	6	8	-	-
1- 7-74	109	-	32	33	5	8	-	-
1-10-74	116	-	32	35	11	7	-	-
1- 1-75	100	-	33	35	12	7	-	-
1- 4-75	101	-	31	(14)	11	4	-	-
1- 8-75	(91)	-	(27)	(14)	(4)	(2)	94	39
1-11-75	(89)	-	(24)	(16)	(4)	(1)	145	38
1- 1-76	(92)	-	(22)	(16)	(4)	(1)	131	31

De getallen tussen haakjes hebben betrekking op de verwerking van banden die buiten het eigenlijke ALADIN-project vallen. De Nobin-financiering liep voor CAB en AGRINDEX af met de laatste tapes van 1975; voor alle andere na de eerste drie van 1975.

CAIN - Cataloguing and Indexing tape of the National Agricultural Library, USA
BoA - Macmillan versie van CAIN
FSTA - Food Science and Technology Abstracts
CAC - Chemical Abstracts Condensates (verwerkt bij NOCI)
BA - Biological Abstracts plus Bio Research Index (verwerkt bij BMDC in Stockholm)
MED - Medlars (verwerkt bij BMDC in Stockholm)
AGRINDEX - Agricultural Research Information System - FAO
CAB - Commonwealth Agricultural Bureaux: Index Veterinarius en Veterinary Bulletin

na de band-datum kon plaatsvinden. De BoA-band was steeds één maand later gedateerd dan de overeenkomstige CAIN-band.

De verwerking is trouwens niet altijd onmiddellijk na binnenkomst uitgevoerd. Computerstoringen, overbelasting, vakanties en wachten op nieuwe profielopstellingen hebben soms tot uitstel geleid. Het hercoderen voor het samengevoegde profielenbestand leidde ook tot vertraging. Pas sinds juni 1976 kan min of meer gesteld worden dat alle banden direct na binnenkomst worden verwerkt.

Bij AGRINDEX heeft de directe toezending uit Wenen in plaats van via Karlsruhe, waar hij voor de EEG-landen apart werd gedupliceerd, tot twee weken tijdwinst geleid. De band komt nu steeds aan het eind van de maand voorafgaand aan de band-datum binnen. (Bij AGRINDEX is de band-datum de maand waarin de afgeleide gedrukte versie van de pers komt.) FSTA kan in het begin van de band-maand worden verwerkt. Oorspronkelijk werd getracht de CAIN- en FSTA-banden met dezelfde band-datum ook tegelijkertijd te verwerken, maar dit betekende voor FSTA een eigenlijk moeilijk te verdedigen vertraging. Bij een attenderingsdienst is de snelheid waarmee de klant geïnformeerd wordt over nieuwe literatuur toch een essentieel element van de dienstverlening.

4.3 VRAAGVERWERKING OP AANVULLENDE BESTANDEN BIJ ANDERE INSTELLINGEN

4.3.1 NOCI

De eerste testprofielen voor het ALADIN-project werden niet met een typisch landbouwkundig bestand in eigen beheer verwerkt, maar met dat van Chemical Abstracts[28] bij de in 1969 opgerichte Nederlandse Organisatie voor Chemische Informatie (NOCI). NOCI was op dat moment de enige instelling in Nederland die een computerattenderingsdienst aan het publiek aanbood, min of meer gelijkend op wat COLADIN voor ogen stond. (NOCI heeft eerst vanaf september 1969 een aantal profielen bij de Deense Technische Bibliotheek laten verwerken. Sinds oktober 1971 wordt CAC bij IWIS-TNO met het FLIRT-programma verwerkt.)

Het ALADIN-project heeft veel profijt gehad van de ervaring met computerattendering die reeds bij NOCI bestond; de projectleider heeft in feite in het begin de kunst bij NOCI kunnen afkijken. Bij de verdere ontwikkeling van het FLIRT-programma is er ook steeds een bijzonder vruchtbare samenwerking tussen NOCI en ALADIN geweest. Aangezien het zoekprogramma voor CAC hetzelfde was als voor CAIN, werden de CAC-profielen ten slotte meestal door de projectleider geschreven en bij NOCI na inhoudelijke controle ingevoerd.

Aan het einde van het project liepen er bij NOCI voor ALADIN 35 profielen, hetgeen 11% was van het totaal aantal CAC-profielen verwerkt door NOCI.

4.3.2 BioMedical Documentation Center (BMDC), sinds 1 november 1975 Medical Information Center (MIC)

Het BMDC in Stockholm heeft voor het ALADIN-project een aantal profielen op de magneetbanden van Biological Abstracts (Biosis)[29] en Medlars (Medical Literature Analysis and Retrieval System van de National Library of Medicine, USA) verwerkt. Oorspronkelijk liepen de contacten voor Biosis via NOCI, maar deze tussenschakel bleek later niet langer noodzakelijk. De bibliotheek bij de diergeneeskundige faculteit van de Rijksuniversiteit te Utrecht (later: 'afd. Biomedische Informatie van de Universiteitsbibliotheek Utrecht') fungeert als tussenschakel voor de Medlars-profielen; daar worden de profielen ook geformuleerd, aangezien dit veel ervaring in het gebruik van de hiërarchische Medlars-thesaurus vereist.

De BA-profielen werden door de projectleider geformuleerd; er werd hierbij door het BMDC jammer genoeg geen steun verleend, maar het gebruik van het retrieval-systeem werd al doende steeds beter beheerst. De communicatie met Stockholm verliep in het algemeen vrij traag: het duurde meestal vele weken voordat wijzigingen werden doorgevoerd. Hoewel de retrieval-resultaten uiteindelijk in het algemeen bevredigend waren, was het toch de

28. De profielen voor het ALADIN-project zijn door NOCI steeds met het bestand 'Chemical Abstracts Condensates' verwerkt. Deze bevat wèl toegevoegde trefwoorden en categoriecodes, maar geen referaten. (De gedrukte versie van Chemical Abstracts bevat wèl referaten. Het bestand 'Chemical Titles' bevat alleen bibliografische informatie.)
29. Het bestand van Biosis heet 'BA previews'. Dit is een combinatie van het referaattijdschrift 'Biological Abstracts', dat tweemaal per maand verschijnt, en de 'Bioresearch Index' (zonder referaten), die eenmaal per maand verschijnt.

wat moeilijke communicatie die er, min of meer onbewust, toe heeft geleid dat er minder profielen op Biosis werden verwerkt dan wellicht zinvol zou zijn geweest.[30]

De communicatie met Utrecht betreffende de Medlars-profielen is steeds bevredigend verlopen. Het was hierbij een voordeel dat al in een vroeg stadium op eenvoudige wijze retrospectief literatuuronderzoek kon worden uitgevoerd door stapel(batch)-verwerking in Stockholm. Deze gaf de mogelijkheid om snel tot een goed uitgangsprofiel te komen en direct met een uitgebreide literatuurlijst te kunnen beginnen. Voor andere bestanden was dit ook wel mogelijk, met name aan de University of Georgia in de USA, maar dit was aanzienlijk duurder en kostte veel meer tijd vanwege de ingewikkelde communicatie. Met de komst van de interactieve retrospectieve retrieval-systemen in 1974 werd dit bezwaar ondervangen; zowel Pudoc als NOCI en de Rijksuniversiteit te Utrecht kregen toen al snel met hun eigen terminal toegang tot diverse bestanden die ook voor profielopstelling werden gebruikt.

4.3.3 Andere contacten

International Labour Organization (ILO) Door de ILO werden twee retrospectieve onderzoekingen gedaan ten behoeve van het ALADIN-project, uitgevoerd op hun eigen bestand. Voor een attenderingsabonnement bleek het rendement echter te gering. De communicatie was plezierig, de formulering door de ILO-staf goed. De gevonden titels waren veelal uniek: het ILO-bestand bevat veel materiaal dat niet elders in machine-leesbare vorm ter beschikking is.

Danmarks Tekniske Bibliothek (DTB) Bij de DTB in Kopenhagen werd lange tijd één profiel verwerkt op Compendex (Engineering Index), nl. 'vliegtuigen in de landbouw'. De opbrengst aan relevante titels was steeds gering, hoewel de gevonden relevante titels in het algemeen niet elders werden gevonden. De communicatie met het centrum was plezierig; men was ook steeds zeer bereidwillig het profiel te verbeteren. In 1969 trad IWIS-TNO al in contact met de DTB uit interesse voor het bij de DTB gebruikte retrieval-programma. Het huidige IWIS-FLIRT-programma bouwt enigszins voort op dit oudere Deense programma. In één geval werd door de DTB voor ons een retrospectief onderzoek op FSTA uitgevoerd; dit betrof het samenstellen van een index op instellingen die onderzoek op het gebied van aroma's doen. Door het grote aantal verschillende schrijfwijzen van de naam van hetzelfde instituut was het resultaat niet direct bruikbaar, maar het was voor de geïnteresseerden toch nuttig.

Excerpta Medica Op 10 juli 1970 begon Excerpta Medica reeds met een computerattenderingsdienst. Mede gezien de hoge kosten en de verwachte geringe opbrengst voor landbouwkundige vragen werd voor het ALADIN-project van deze attenderingsdienst echter geen gebruik gemaakt. Via de Rijksuniversiteit van Utrecht werd in de loop van 1973 met het retro-

30. In 1976 werden de Biosis-profielen voor Pudoc van BMDC overgebracht naar de Biological Information Service (G.B.) vanwege de veel lagere prijs en de uitvoer op kaartjes i.p.v. op vellen.

spectieve on-line-systeem wel een aantal proeven genomen, die echter door technische mankementen weinig bevredigend waren. In 1975 werd ten slotte nog een aantal maal retrospectief onderzoek uitgevoerd voor de zich daartoe lenende profielen. De profielopstelling hiervoor werd geheel door de staf van Excerpta Medica verzorgd; Pudoc heeft er ook geen inzage in gekregen. De resultaten kwamen met grote vertraging binnen; ze bevatten wel allerlei nieuw materiaal, hoewel de relevantie in het algemeen laag was. (De profielopstelling zal er wel op gericht zijn geweest om in ieder geval zoveel mogelijk van de beschikbare, mogelijk relevante titels te vinden.) De overlapping door Medlars was over het algemeen gering.

ENDS/INIS Het profiel betreffende het effect van radioactieve neerslag op levende organismen werd bij de Europese Commissie in Luxemburg verwerkt tegen het bestand van het 'European Nuclear Documentation System' van de Europese Commissie (ENDS). De uitvoer kwam traag binnen en bevatte hoofdzakelijk oud materiaal. Het Europese ENDS-bestand werd met 1972 afgesloten, in afwachting van het nieuwe 'International Nuclear Information System' (INIS) dat in 1973 operationeel was geworden.[31] De Europese Commissie stelde later het INIS-bestand ook voor retrospectief onderzoek ter beschikking. Hierop werd deze vraag toen herhaald, echter met onbevredigende resultaten. Ten slotte werd ook het INIS-bestand, zoals dit bij de IAEA in Wenen is opgeslagen, nog voor deze vraag doorzocht; ook daarbij was de relevantie laag, maar merkwaardig genoeg werden bij beide INIS-onderzoekingen slechts weinig overeenkomstige titels gevonden. Beide profielopstellingen moeten dus verre van ideaal geweest zijn.

CO-TNO, afd. Documentatie De afdeling Documentatie van de Centrale Organisatie TNO is speciaal belast met documentatie op het gebied van octrooiliteratuur. Octrooien zijn wel in Chemical Abstracts, FSTA en enkele CAB-tijdschriften opgenomen, maar worden overigens in secundaire landbouwkundige informatie nogal verwaarloosd. Via de CO-TNO werd een aantal onderzoekingen gedaan met de Central Patent Index van Derwent Ltd. Dezelfde onderwerpen werden tevens door bemiddeling van de CO-TNO verwerkt bij de Internationale Dokumentationsgesellschaft für Chemie (IDC) in Frankfurt. Het IDC-systeem bevat behalve octrooien ook artikelen uit de organisch-chemische literatuur. Het bijzondere van het systeem is dat gezocht kan worden met behulp van een combinatie van fragmenten van chemische structuurformules en zgn. rolindicatoren, die de functie van chemische stoffen aangeven.

Dit literatuuronderzoek leverde in het algemeen veel materiaal op dat niet elders werd gevonden ('uniek' was), maar een bezwaar was dat de gevonden referenties in het algemeen vrij oud waren. Beide systemen zijn arbeidsintensief, omdat er als resultaat allereerst een lijst met referentienummers uitkomt, waarbij dan lokaal de originele teksten moeten worden gezocht, gefotocopieerd en opgestuurd. Het resultaat is dan wel dat de klant direct de beschikking krijgt over zeer uitgebreide informatie. De samenwerking met de

31. ENDS en INIS zijn dus twee aparte bestanden. De organisatiestructuur van INIS met gedecentraliseerde nationale invoer diende als voorbeeld voor AGRIS. De magneetbanden van AGRINDEX worden, evenals INIS, geproduceerd bij de International Atomic Energy Agency (IAEA) in Wenen.

CO-TNO was bijzonder plezierig. Er werd door de staf veel tijd besteed aan het formuleren van de vragen. De totale verwerkingstijd voor de vragen was lang (11 maanden!).

RINGDOC Het RINGDOC-systeem van Derwent Ltd. kent evenals het IDC-systeem de mogelijkheid op fragmenten van chemische structuurformules te zoeken. Philips-Duphar stelde ons in staat met dit systeem kennis te maken. Wederom leverde dit voor de gebruikte profielen een groot aantal unieke titels op. Evenals bij IDC en de Central Patent Index is het tussenresultaat van het literatuuronderzoek een lijst met referentienummers, waarna de documentalist de tekst van de originele referaten op moet zoeken. Deze staat niet op de magneetbanden maar is alleen in gedrukte vorm beschikbaar. Philips-Duphar stelde in dit geval de profielen samen. Het contact was incidenteel en zeer plezierig.

4.4 HET OPSTELLEN VAN DE ZOEKPROFIELEN

Het opstellen van zoekprofielen bleek zowel een kunde als een kunst, die al doende werd geleerd. Met het ideale zoekprofiel moeten alle relevante 'records' die op een bepaalde band staan, worden gevonden maar geen enkele irrelevante 'record'. Deze twee eisen zijn in de praktijk veelal tegenstrijdig. Het informatiegehalte van de 'records', het onderwerp en de behoeften en het temperament (!) van de klant bepalen uiteindelijk de optimale profielopstelling. Ten slotte moet de opsteller van het profiel nog streven naar een zo efficiënt mogelijk profiel: omdat iedere zoekterm extra computerzoektijd met zich meebrengt, moet het aantal termen zo laag mogelijk worden gehouden.

Voordat een eerste opstelling werd gemaakt, werd iedere gebruiker gevraagd een lijst te leveren met ten minste tien titels van reeds bekende publikaties die hij voor zijn onderwerp zeer relevant achtte. Bij het formuleren van het profiel zou ervoor gezorgd moeten worden dat deze titels in ieder geval werden gevonden. Tevens werd hem gevraagd zoveel mogelijk trefwoorden met eventuele synoniemen op te geven die voor zijn onderwerp van belang waren. (Een aanvraagformulier, zoals dat momenteel wordt gebruikt, is toegevoegd als Bijlage 11.)

Bij de eerste opstelling werd er steeds naar gestreefd zoveel mogelijk (misschien) relevante titels te vinden, zonder al te zeer op het relevantiepercentage (precisie) te letten. Op deze manier konden namelijk eenvoudig extra trefwoorden worden gevonden, die in een profiel met grotere precisie wellicht noodzakelijk zijn. Dit is de 'breed naar smal'-strategie. Men kan ook omgekeerd uitgaan van een profiel met hoge precisie, met echter een grote kans dat lang niet alle op de band aanwezige relevante titels worden gevonden. Met die strategie bleek het echter veel moeilijker om het profiel later nog te verbreden, omdat men van de resultaten geen nieuw inzicht kreeg in noodzakelijk toe te voegen trefwoorden. Men is daartoe dan geheel op andere bronnen, zoals woordenlijsten en reeds bestaande bibliografieën, aangewezen. Het belang van zulke bronnen moet toch al niet onderschat worden: het is de taak van de 'profielformulator' zoveel mogelijk van belang zijnde trefwoorden, inclusief synoniemen, verwante termen en woordvormen te vinden en in het profiel op te nemen. Bij verschillende landbouwhuisdieren moet bijvoorbeeld gedacht worden aan de aparte namen voor het mannelijk en vrouwelijk dier, voor jonge dieren, aan afwijkende meervoudsvormen, Engelse en Amerikaanse namen en spellingsvarianten. Een

ander voorbeeld vormen radio-isotopen, waarvoor het ook noodzakelijk is veel verschillende schrijfwijzen op te nemen. Ten slotte moet voor het zoeken naar aminozuren steeds de hele lijst van de individuele aminozuren worden opgenomen! Dit is althans steeds noodzakelijk bij bestanden zonder toegevoegde termen uit een gecontroleerde woordenlijst. Van de verwerkte bestanden gebruikte alleen FSTA een dergelijk systeem. In het algemeen geldt hier: hoe meer intellectuele arbeid bij het maken van het bestand wordt verricht, des te minder werk is er nodig voor het afzoeken van het bestand naar relevante titels.

Het aantal termen per profiel was bijzonder verschillend. Het varieerde van één tot meer dan 250. Het gemiddelde bedroeg voor CAIN bijna 30, voor FSTA iets meer dan 20. Het aantal logische groepen dat werd gebruikt was vrijwel nooit meer dan drie. In het algemeen is het stellen van meer dan drie eisen aan een titel + trefwoorden niet zinvol. In een aantal gevallen werd een vrij ingewikkelde gewichtsstructuur met drempelgewicht gebruikt om verschillende logische combinaties te bewerkstelligen. Met het beschikbaar komen van volledige Booleaanse logica is dat nu niet meer nodig.

Het gebruik van trunkatie vereist inzicht in de mogelijke negatieve gevolgen. Zo leek het gunstig om voor het vinden van alle met pectinen in verband staande termen * *PECT* te gebruiken. De linker trunkatie werd gebruikt om termen als DEPECTINIZATION en PROTOPECTINASE ook direct te vinden. Deze trunkatie leidt echter ook tot het vinden van alle titels met ASPECT*, hetgeen niet de bedoeling was en erg veel irrelevante titels opleverde. Daarom moest naast PECTIN* (alleen rechts getrunkeerd) ook PROTOPECTIN* en DEPECtin* worden opgenomen. Om de risico's van het vinden van onbedoelde woorden bij het trunkeren zoveel mogelijk te verminderen, werd gebruik gemaakt van de tabellen met woordfrequenties van Lockheed Corp. Deze tabellen vermelden voor een aantal bestanden naast elkaar (bijv. CAIN, Biosis, Chemical Abstracts en ISI) het aantal keren dat een bepaalde term voorkomt. De rangschikking is alfabetisch. Uiteraard kan men hierbij alleen de consequenties voor rechter trunkatie zien, niet voor linker trunkatie. Voor Chemical Abstracts bestaat wel een tabel met frequenties van woord-stammen, maar deze werd vanwege de hoge prijs niet aangeschaft.

Voor het opstellen van efficiënte profielen is een goede kennis van de categoriecodes van de diverse bestanden erg belangrijk. Het toepassen van categoriecodes kan soms een aanzienlijke besparing in het aantal in te voeren termen betekenen, vooral als het coderingssysteem hiërarchisch is[32], zodat men door rechter trunkatie direct op grotere groepen van hogere orde kan selecteren. (Bijv. veeteeltkundige onderwerpen beginnen met 25 ...)

De coderingssystemen zijn steeds gericht op het specifieke onderwerp van het bestand en daardoor voor ieder bestand verschillend. Geen van de gebruikte bestanden gebruikt de UDC-codering. De soms zeer summiere omschrijvingen van de betekenis van de diverse codes maakt het directe gebruik van die codes soms gevaarlijk, omdat niet duidelijk is wat men van het gebruik van een code kan verwachten. Men verkrijgt in het algemeen de meeste relevante titels door categoriecodes *naast* overeenkomstige termen op te nemen: wat niet met

32. Jammer genoeg zijn de nummeraanduidingen van de Veterinary Subject headings niet consequent hiërarchisch.

de één wordt gevonden kan met de ander nog gebeuren. Op deze manier gaat echter het voordeel van een kleiner aantal termen en dus een goedkoper profiel verloren. Biosis voegt bijzonder veel coderingen toe, met daarbij nog een hiërarchie naar belangrijkheid in drie trappen ('major', 'secundary' en 'tertiary'). Het gebruik van alle niveaus leidde in het algemeen tot een erg laag relevantiepercentage.

De gewascodes bij AGRINDEX werden als zeer bruikbaar ondervonden - een grote verbetering ten opzichte van CAIN. De taxonomische codes bij BA bleken ook goed bruikbaar, maar zijn niet specifiek op landbouwgewassen gericht; 'grassen' en 'granen' bijvoorbeeld kunnen niet in landbouwkundige zin worden onderscheiden.

De profielen werden gedurende de hele projectperiode vrij regelmatig aangepast. Gemiddeld werd na de verwerking van 3 tot 4 banden ('runs') iets aan het profiel gewijzigd. Doordat gedurende het hele project de geleverde titels steeds beoordeeld terug werden ontvangen, was het ook goed mogelijk voortdurend bij te sturen. Vanwege de grote arbeidsintensiviteit werd hiervan in de operationele fase afgezien. Er wordt nu gevraagd om terugzending van beoordeelde duplicaat-titels voor de eerste drie afleveringen en daarna voor zolang dat noodzakelijk lijkt. In het algemeen kan een profiel na drie 'runs' zodanig zijn bijgestuurd, dat het efficiënt loopt. Het wordt daarna aan de klant overgelaten om eventuele nieuwe of onbevredigde wensen aan Pudoc kenbaar te maken. Het ligt daarnaast in de bedoeling alle gebruikers eenmaal per jaar te vragen of zij met de resultaten van hun attenderingsabonnement nog steeds tevreden zijn en te overleggen wat er eventueel aan zou kunnen worden verbeterd.

Een retrospectief literatuuronderzoek kan een goed uitgangspunt zijn voor het opstellen van een attenderingsprofiel. In veel gevallen zal een attenderingsprofiel echter breder zijn dan het zoekprofiel voor een retrospectief onderzoek, omdat men gemakkelijker maandelijks bijv. 20 titels doorkijkt dan in één keer 960 titels, het resultaat van een retrospectief onderzoek over 4 jaar met hetzelfde profiel. Bij attendering zal men eerder randgebieden meenemen om daardoor de kans op het missen van belangrijke nieuwe publikaties zo klein mogelijk te houden, terwijl eventuele gemiste belangrijke publikaties bij een 'retro' anders wel door het 'sneeuwbalsysteem' bekend zullen worden.

5 Manuele en computer-attendering vergeleken

5.1 INLEIDING

Bij Pudoc bestaat reeds sinds 1954 de mogelijkheid geattendeerd te worden op recent verschenen literatuur in het interessegebied van de opdrachtgever. Hiertoe worden alle publikaties die bij de Centrale Bibliotheek van de Landbouwhogeschool binnenkomen, dagelijks doorgenomen (ca. 200 per dag). Van ieder artikel dat op het terrein van één der opdrachtgevers lijkt te liggen, wordt een titelkaartje getypt, dat aan de klant wordt toegezonden, eventueel vergezeld van een fotocopie van het origineel. Pudoc houdt zelf ook kopieën van alle kaartjes, die systematisch worden opgeborgen. Eind 1975 liepen er 39 attenderingsopdrachten.

Het werk stelt hoge intellectuele eisen: de literatuuronderzoeker moet de inhoud van wetenschappelijke publikaties op zeer uiteenlopend gebied kunnen onderscheiden naar hun relevantie voor de verschillende opdrachtgevers, die allen hun specifieke eisen hebben. Het is zeer arbeidsintensief: 50 verschillende onderwerpen wordt ongeveer als het maximum beschouwd dat één persoon kan behandelen. Men kan echter moeilijk continu met dit soort werk bezig zijn; bovendien moeten ziektes en vakanties kunnen worden opgevangen. De huidige 39 opdrachten vereisen twee literatuuronderzoekers en één administratieve kracht (totaal 41 manuren per week voor de literatuuronderzoekers plus 35 uur per week voor administratief personeel).

Gezien de huidige beperkingen bij het aantrekken van nieuw personeel mag het zo goed als uitgesloten worden geacht dat de bemanning van deze dienst zou kunnen worden uitgebreid, temeer daar het niet goed mogelijk is de (zeer hoge) werkelijke kosten door te berekenen.

Aangezien de manuele attenderingsdienst steeds redelijk dicht bij haar geschatte maximum-capaciteit van 50 opdrachten heeft geopereerd, heeft Pudoc nooit veel reclame gemaakt voor deze dienst. Daardoor bestaat er echter ook weinig inzicht in de potentiële vraag naar deze dienstverlening.

Eén van de centrale vragen waarop het ALADIN-project een antwoord zou moeten geven, is de mate waarin de computer zou kunnen worden ingeschakeld om het volume van de attenderingsdienst te kunnen opvoeren. Het was direct duidelijk dat er voor de computer nauwelijks kwantitatieve limieten zijn voor het aantal klanten dat kan worden bediend. De vraag was dan ook hoofdzakelijk of een computerdienst een produkt kan leveren dat min of meer gelijkwaardig is aan dat wat de manuele attenderingsdienst[33] levert. Het gaat daarbij hoofzake-

33. Onder de 'manuele attenderingsdienst' of kortweg 'manuele attendering' wordt in het vervolg steeds de manuele attenderingsdienst van Pudoc verstaan.

lijk om twee vragen:
1. In welke mate worden met behulp van computersystemen en manuele attendering dezelfde titels gevonden? (zie 5.2)
2. Hoeveel vertraging treedt er op bij het attenderen met de computer in vergelijking met manuele attendering? (zie 5.3)

Daarnaast spelen nog andere kwesties een rol, in het bijzonder de verkrijgbaarheid van het origineel (5.4.3.1), de kosten (7.5), de beschikbaarheid van vakkundige literatuuronderzoekers, het belang van snelle attendering, de irritatie die opgewekt kan worden door het ontvangen van niet-relevante titels bij computerattendering (5.4.3.2), en de aard van het uitgangsmateriaal (5.4.1 en 5.4.2).

Het is wellicht goed zich bij deze vergelijking te realiseren dat de twee systemen niet alleen verschillen in het middel waarmee gezocht wordt, nl. het menselijk intellect tegenover de computer, maar ook in het materiaal waarin gezocht wordt, nl. primaire documenten tegenover secundaire bronnen, zoals referaattijdschriften en titelbibliografieën, die van een magneetband worden gelezen. Een betere vergelijking zou dus worden verkregen door twee vergelijkingen: a) manueel zoeken in primaire documenten tegenover manueel zoeken in secundaire bronnen en b) manueel zoeken in secundaire bronnen tegenover het zoeken met behulp van een computersysteem in deze bronnen.

Hoewel voor Pudoc de alternatieven manueel/primair en computer/secundair als reëel gelden, ligt deze situatie bij vele andere instellingen en personen vaak anders: men beschikt zelden over zoveel primair materiaal als de Centrale Bibliotheek van de LH, maar gebruikt wel vaak secundaire bronnen voor attendering op nieuw verschenen literatuur.

Om op de oorspronkelijke vragen een antwoord te kunnen geven werden de resultaten van een aantal lopende manuele attenderingsopdrachten vergeleken met de resultaten van computerverwerking van magneetbanden met bibliografische bestanden voor dezelfde literatuur-interesseprofielen. De resultaten zijn uiteraard erg afhankelijk van het onderwerp en van de gebruikte bestanden op magneetband. In het totaal werden de resultaten van 25 literatuur-interesseprofielen vergeleken. De onderwerpen en de gebruikte magneetbanden staan vermeld in Bijlage 3. Deze 25 profielen werden geput uit de bestaande abonnementen bij de manuele attenderingsdienst. Van de bij de aanvang beschikbare 34 onderwerpen werden een aantal niet voor de vergelijking geschikt geacht vanwege selectiecriteria die niet door het computerzoekprogramma konden worden gehanteerd, of omdat van te voren vast stond dat het relevante uitgangsmateriaal niet in voor de computer leesbare vorm beschikbaar was. (Zie hiervoor verder 5.4.1 en 5.4.3.3.)

5.2 DE MATE WAARIN MET COMPUTER-ATTENDERING DEZELFDE TITELS GEVONDEN WORDEN ALS MET MANUELE ATTENDERING

Bij het analyseren van de overlapping van twee systemen moet onderscheid worden gemaakt tussen verschillen in het *uitgangsmateriaal* en verschillen in de *vindbaarheid* bij beide systemen. Een titel kan alleen in een systeem worden gevonden als hij er ook in aanwezig is, maar door gebreken in het zoeksysteem zullen niet alle aanwezige relevante titels werkelijk worden gevonden. Deze gebreken zullen bij manuele attendering veel kleiner zijn dan bij computer-attendering, doordat bij de eerste steeds het gehele oorspronkelijke

document kan worden geraadpleegd om te bepalen of (iets uit) de inhoud relevant is voor de klant. Bij computer-attendering is de hele tekst in het algemeen niet in machine leesbare vorm beschikbaar. In het beste geval staat er een informatief referaat op de magneetband, in het ongunstigste geval bevat de band alleen de bibliografische gegevens, waarbij in feite alleen de titel kan worden gebruikt om te selecteren. Daartussen bestaan verschillende vormen van min of meer verfijnde classificaties en toegevoegde trefwoorden, al dan niet afkomstig van een gecontroleerde en systematische woordenlijst.

Het uitgangsmateriaal voor beide systemen is geheel verschillend: de manuele attendering is gebonden aan de collectie van de Bibliotheek van de Landbouwhogeschool, de computer-attendering aan de gebruikte bestanden en het daarin beschikbare materiaal. De manuele attendering krijgt daarbij zo goed als al het nieuwe materiaal van deze bibliotheek te zien met uitzondering van de meeste monografieën; hierop wordt de attenderingsdienst geattendeerd door middel van cataloguskaartjes, waardoor inhoudelijke analyse van deze documenten zelden gebeurt.

Handbibliotheken ontvangen regelmatig publikaties die niet in de Centrale Catalogus worden opgenomen en die daardoor ook de manuele attenderingsdienst niet passeren. Dit betreft vooral geschenken van medewerkers aan de Handbibliotheken, waaronder veel overdrukken. Ten slotte zijn er een aantal tijdschriften die direct aan Handbibliotheken geadresseerd zijn; ook deze passeren de manuele attenderingsdienst niet, evenmin als de weekbladen en vertiendaagse bladen voor de Handbibliotheken, omdat de manuele attendering daarvoor teveel oponthoud veroorzaakt.

De computer-attendering gebruikt nooit alle beschikbare bestanden, doch alleen die waarvan verwacht mag worden dat ze een 'redelijke' opbrengst aan relevante titels zouden leveren, mede gezien de kosten die het verwerken van ieder extra bestand met zich meebrengt. De selectie van de te gebruiken bestanden geschiedt meestal op grond van praktijkervaring en niet zozeer experimenteel. Het is dus van groot belang dat degene die de te gebruiken bestanden selecteert goed op de hoogte is van de aard van al het beschikbare uitgangsmateriaal. Het blijft zo altijd mogelijk dat een bepaalde titel wel in voor de computer leesbare vorm beschikbaar is, maar dat het bestand waarin hij zich bevindt niet wordt doorzocht In de praktijk van het ALADIN-project werd achteraf ook wel een enkele maal geconstateerd dat een bepaald bestand niet gebruikt was, dat beter wel gebruikt had kunnen worden. Met name Biological Abstracts bleek in meer gevallen voor de landbouw relevant (en nieuw) materiaal te bevatten dan oorspronkelijk werd aangenomen.

Voor het vergelijkend onderzoek werden alleen titels gebruikt van publikaties die in 1973 verschenen en die tot eind 1974 werden gevonden. De resultaten zijn samengevat in Fig. 7 en Bijlage 4. Gemiddeld over 25 profielen blijkt door de manuele attenderingsdienst 30% van alle ontdekte relevante[34] titels te worden gevonden en in de met de computer doorzochte bestanden 80%; 10% van alle titels werd door beide systemen gevonden. (Als de cijfers niet gemiddeld worden over het totaal, maar per profiel, zijn deze getallen 35% en 76%.) Dit impliceert dat van de 100 titels gemiddeld 10 zowel door de manuele als door de

34. Inclusief 'misschien relevante titels', waarbij de relevantie moeilijk aan de titel was te beoordelen.

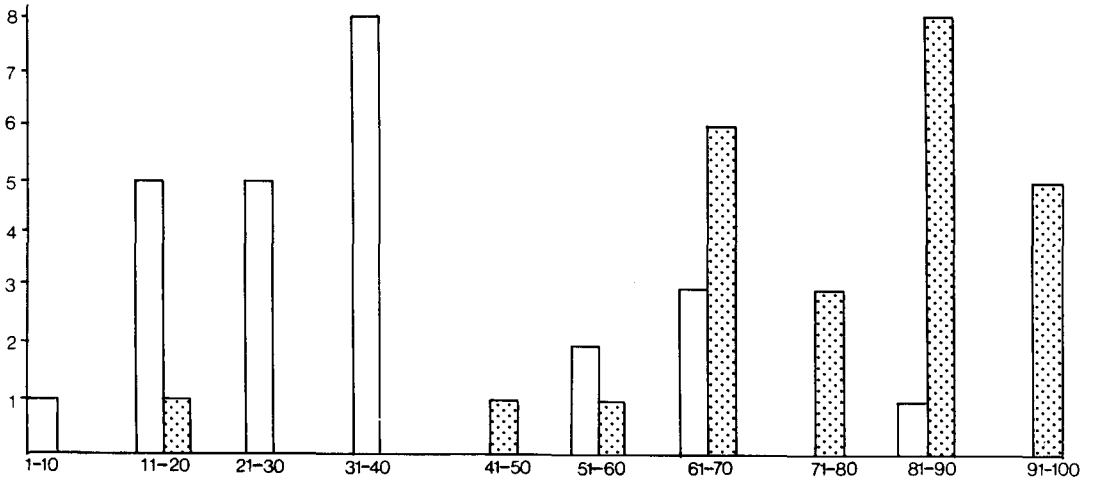


Fig.7. Frequentieverdeling van de dekkingspercentages voor computer-attendering en manuele attendering inclusief misschien-relevante titels.
□ = manuele attendering, ⊡ = computer-attendering.

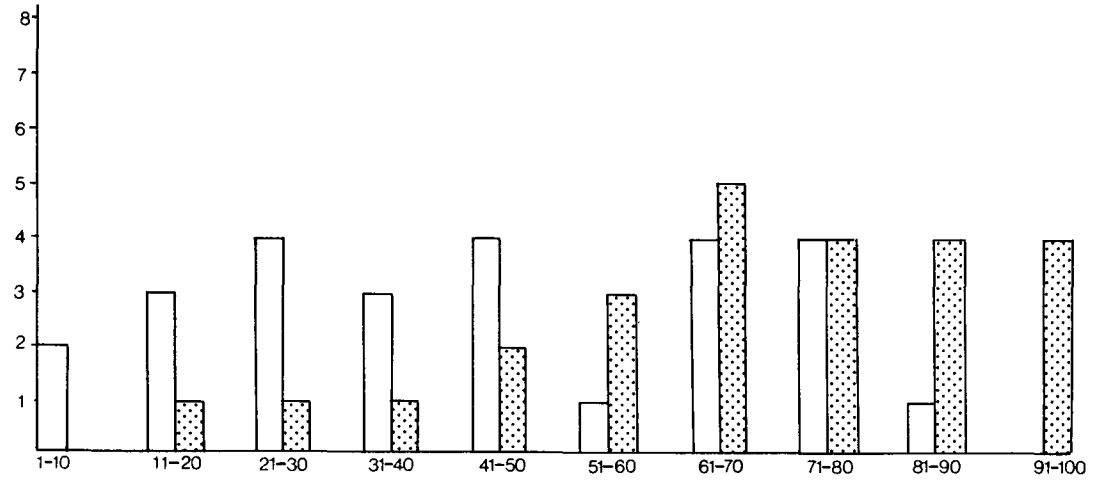


Fig.8. Frequentieverdeling van de dekkingspercentages voor computer-attendering en manuele attendering exclusief misschien-relevante titels.
□ = manuele attendering, ⊡ = computer-attendering.

computer-attendering worden gevonden, 20 alleen door de manuele attendering en 70 alleen door de computer-attendering.

Zoals uit de frequentieverdelingen in Fig.7 blijkt, zijn de verschillen tussen de diverse onderwerpen nogal groot: bij 6 van de 25 profielen werd door manuele attendering meer dan de helft van het totaal aantal titels gevonden. Dat waren de volgende onderwerpen: weefselcultuur (profielnummer 11 uit Bijlage 3, spinazie (12), economie van de zetmeelproduktie (19), organisatie van het landbouwkundig onderzoek (21), invloed van ozon op de plantengroei (22) en invloed van SO_2 op de plantengroei (23). Voor deze onderwerpen levert het materiaal dat via de Centrale Bibliotheek van de LH binnenkomt, dus relatief veel titels op. De dekking is echter nergens bijzonder groot te noemen, op het onderwerp 'organisatie van het landbouwkundig onderzoek' na, waar de dekking 81% bedroeg. Veel titels hadden hier op de typisch Nederlandse situatie betrekking, terwijl het moeilijk was met de computer alleen aan de hand van titels relevant materiaal uit CAIN te laten selecteren. (Er was bij dit profiel geen overlapping tussen CAIN en de manuele attendering.)

Is dus in het algemeen de dekking door de manuele attendering niet uitermate groot, computer-attendering alléén geeft echter ook vrijwel nooit alle relevante titels. In het beste geval leverde computer-attendering 94% dekking op - 'coccidiostatica' (profielnummer 2) en 'fallout' (18) - met daarop volgend 'menselijke ziekten veroorzaakt door aardappelconsumptie' (6) en 'effect van bestraling op DNA' (10) met elk 93%, maar in alle gevallen heeft de manuele attendering toch een aantal titels gevonden die met de computer niet gevonden werden.

Indien de 'misschien relevante' titels allemaal uiteindelijk toch irrelevant blijken te zijn, verschuift het beeld natuurlijk wel ten gunste van de manuele attendering. (Daarbij komt de categorie 'misschien relevant' immers niet voor.) De gegevens staan vermeld in Fig. 8 en Bijlage 5. Het aantal profielen waarbij meer dan 50% van alle titels werd gevonden, is dan 10 i.p.v. 6 voor manuele attendering en 20 i.p.v. 23 voor de computer. Waren er met de 'misschien relevante' titels 3 profielen waarbij de manuele attendering meer opleverde dan de computer-attendering (profielnrs. 19, 21 en 23), zonder de 'misschien relevante' titels zijn dat er nog 4 meer, nl. de nummers 12, 15, 22 en 24.

De keuze tussen manuele en computer-attendering, als er al gekozen kán worden, mag natuurlijk niet bepaald worden door de *gemiddelde* uitkomsten. Van geval tot geval zal moeten worden bekeken wat van de ene en van de andere methode mag worden verwacht en hoe zwaar de voor- en nadelen van beide systemen voor een specifiek geval wegen. Aan het einde van deze paragraaf kan alleen worden geconcludeerd dat computer-attendering in de meerderheid van de gevallen meer relevante titels opleverde dan de manuele attendering; met iedere methode werden echter steeds een aantal titels gevonden, die met de andere methode niet werden ontdekt.

5.3 TIJDSVERSCHILLEN IN SIGNALERING TUSSEN DE MANUELE EN COMPUTER-ATTENDERING EN COMPUTER-DIENSTEN ONDERLING

De snelheid waarmee door manuele attendering titels ter beschikking van de klant komen, is steeds aangeprezen als één van de voordelen van dit systeem. De signalering geschiedt immers direct na binnenkomst van de oorspronkelijke publikatie - sneller kan dus

Tabel 2. Gemiddeld tijdsverschil in signalering tussen verschillende bestanden uitgedrukt in maanden (de getallen tussen haakjes betreffen het aantal observaties).

	manueel	CAIN	CAC	Medlars	FSTA	BA
manueel	0	1,9(100)	- 2,1(100)	*	4,8(38)	8,1(39)
CAIN	-1,9(100)	0	- 4,7(100)	1,7(36)	1,7(100)	5,7(100)
CAC	2,1(100)	4,7(100)	0	7,9(41)	6,2(100)	10,6(35)
Medlars	*	-1,7(36)	- 7,9(41)	0	*	*
FSTA	-4,8(38)	-1,7(100)	- 6,2(100)	*	0	*
BA	-8,1(39)	-5,7(100)	-10,6(35)	*	*	0

* Minder dan 20 observaties.

De volgende veronderstellingen werden gemaakt met betrekking tot de vergelijkingsdata:
manueel - datum van signalering
CAIN - één maand na de datum van de band
CAC - één maand na de datum van de band
Medlars - de maand van de band
FSTA - de maand van de band
BA - de datum van verwerking in Stockholm
Deze veronderstellingen komen het best overeen met het moment waarop het resultaat ter beschikking kan zijn.

nauwelijks. Bij signalering uit secundaire bronnen zit er altijd een stap tussen: de produktie en verzending van het referaatorgaan (of de magneetband daarvan). Alleen secundaire tijdschriften die over vóórinformatie betreffende komende primaire publikaties beschikken, kunnen sneller zijn dan de manuele attendering. 'Current Contents' is daarvan een voorbeeld; dit is een titelbibliografie die veelal wordt samengesteld uit fotocopieën van drukproeven van de inhoudspagina's van nog te verschijnen tijdschriftafleveringen.[35] Gebleken is dat ook Chemical Abstracts (CAC) vroegtijdig geïnformeerd moet zijn, daar verschillende publikaties die in 1973 werden gepubliceerd, reeds eind 1972 als referaat in Chemical Abstracts stonden.[36]

De resultaten van de analyse van de verschillen in snelheid tussen de manuele attendering en de diverse computer-diensten zijn samengevat in Tabel 2. Het getal 1,9 achter 'manueel' en onder 'CAIN' geeft aan dat dezelfde titel bij de manuele attendering gemiddeld 1,9 maand eerder werd gevonden dan in CAIN.

Uit Tabel 2 blijkt dat CAC het tegen de verwachting in meestal in snelheid van de manuele attendering wint: een titel werd gemiddeld twee maanden eerder in CAC gevonden dan door de manuele attendering. Behalve door vóórinformatie kan dit ook veroorzaakt zijn doordat de magneetband (per luchtpost) korter onderweg is dan de per zeepost arriverende gedrukte versie. Dit kan wel 1½ maand schelen. De manuele attendering was echter sneller dan alle andere computer-diensten: bijna twee maanden sneller dan CAIN, bijna vijf maanden sneller dan FSTA en acht maanden sneller dan BA. Met Medlars kan geen directe vergelijking worden gemaakt, maar aangenomen mag worden dat dit verschil in de buurt van de vier maanden zal liggen.

35. Men kan natuurlijk ook een manuele attenderingsdienst op Current Contents baseren. Die zal dan ook sneller zijn dan manuele attendering op basis van aanwezige primaire documenten.
36. Mogelijk betreft dit vooral primaire tijdschriften die door de Chemical Society, waartoe ook Chemical Abstracts behoort, zelf worden uitgegeven.

Tabel 3. Frequentieverdelingen (in %) van de tijdsverschillen tussen het moment waarop een bestand een bepaalde titel signaleerde en het moment waarop dezelfde titel door een ander bestand werd gesignaleerd.

Maanden	CAC/ manueel	CAC/ CAIN	CAC/ Medlars	CAC/ FSTA	CAC/ BA
-12	1				
-11					
-10					
- 9					
- 8	1				
- 7		1			
- 6					
- 5					
- 4	1				
- 3	2			2	
- 2	7				
- 1	4	1			
0	10	2	2,5	2	
1	10	3		4	
2	27	6	2,5	3	
3	18	16	5	5	
4	7	18	12	7	3
5	4	23	20	11	3
6	2	10	22	18	
7		9	5	19	9
8		4	2,5	11	11,5
9	1	2	7	8	14
10		2		3	14
11		2	5	3	14
12		1	2,5	1	9
13			2,5	2	9
14	2		2,5	1	
15	3		7		6
16			2,5		6
17					
18					
19					
20					
21					3

Het is opvallend dat BA bijzonder laat is, terwijl er toch niet meer inhoudelijk werk behoeft te worden verricht dan voor CAC, die tegelijk ter beschikking komt met de band waarvan Chemical Abstracts wordt gedrukt. (De 'Chemical Titles'-band, waaraan geen coderingen en trefwoorden zijn toegevoegd, komt wel bijna drie maanden eerder ter beschikking dan de 'Condensates'-band.) Een mogelijke verklaring zou kunnen zijn dat bij de Chemical Society het merendeel van de referaten zelf wordt geschreven, terwijl er bij BIOSIS veel door verspreide medewerkers buitenshuis wordt gedaan.[37]

37. De ervaring bij het diergeneeskundig bureau van de Commonwealth Agricultural Bureaux is echter dat de verwerking van een titel voor het referaatorgaan 'Veterinary Bulletin' niet meer tijd kost dan de verwerking voor de titelbibliografie 'Index Veterinarius'. De verwerking voor beide bladen geschiedt parallel. Alleen als het referaat buitenshuis wordt geschreven, neemt het meer tijd.

Tabel 3. Vervolg.

Maanden	manueel/ CAIN	manueel/ FSTA	manueel/ BA	CAIN/ Medlars	CAIN/ FSTA	CAIN/ BA
-12						
-11						
-10	1					
- 9	1					
- 8			2,5			
- 7	1					
- 6	1				3	
- 5	2					
- 4	1				3	
- 3	2				3	
- 2	1				1	1
- 1				17	4	1
0	5	3		11	17	2
1	19			25	17	1
2	28	8		14	27	6
3	17	32		11	11	9
4	8	26	2,5	17	3	14
5	6	11	10	2	3	22
6	2	5	13		1	14
7	1	8	23		3	10
8	2		10	2		8
9	2		10		1	2
10			10		1	1
11		3	8			2
12						3
13						2
14			2,5			
15			2,5			1
16						
17			2,5		2	
18						
19						1
20			2,5			

De vergelijkingen FSTA/Medlars, FSTA/BA en BA/Medlars komen niet voor, omdat daarvoor minder dan 20 observaties ter beschikking waren.

Bij de tijdsverschillen tussen signalering bij twee dezelfde bestanden treden vrij grote verschillen op. Zo werd zowel één titel in CAC acht maanden later dan in CAIN gevonden als één 12 maanden vroeger. In 50-60% van de gevallen bedroeg de afwijking van het meest voorkomende verschil echter niet meer dan 1,5 maand naar de ene of de andere kant. De frequentieverdelingen staan vermeld in Tabel 3.

Concluderende kan gesteld worden dat de manuele attendering gemiddeld twee maanden sneller is dan de belangrijkste landbouwkundige data base, CAIN, en iets langzamer dan CAC. Erg groot lijkt het tijdsvoordeel van de manuele attendering dus niet, maar het belang dat eraan gehecht wordt zal individueel zeer kunnen verschillen. Voor sommige andere bestanden, vooral BA, is het verschil echter aanmerkelijk groter. Bedacht moet ook nog worden dat bij de manuele attendering het oorspronkelijke document ook direct ter beschikking kan worden gesteld, hetgeen bij attendering d.m.v. secundaire bronnen meestal niet het geval is. Dit kan een aanzienlijke tijdwinst opleveren.

5.4 ANDERE ASPECTEN BIJ DE VERGELIJKING VAN MANUELE EN COMPUTER-ATTENDERING

5.4.1 Aard van het uitgangsmateriaal

De attenderingsresultaten zijn uiteraard zeer afhankelijk van het uitgangsmateriaal, zoals al in 5.1 werd aangeduid. Onderscheid moet daarbij worden gemaakt naar de samenstelling van de uitgangsverzameling en het informatiegehalte van de elementen uit die verzameling.

Bij de manuele attendering mag men ervan uitgaan dat alle relevante titels die zich in de uitgangsverzameling bevinden, ook inderdaad worden gevonden. De samenstelling van die verzameling wordt echter bepaald door het acquisitiebeleid van de Bibliotheek van de LH, en de gevolgde organisatorische procedures.

Wat betreft het laatste werd in 5.2 reeds opgemerkt dat niet al het materiaal dat aan de Handbibliotheken van de Landbouwhogeschool wordt toegevoegd, de manuele attenderingsdienst passeert, terwijl op nieuwe monografieën alleen door middel van een catalogus-kaartje wordt geattendeerd. De huidige situatie is echter het resultaat van het zo goed mogelijk afwegen van diverse belangen.

De vergelijking tussen computerattendering en manuele attendering kan natuurlijk wel aanwijzingen geven voor mogelijke hiaten in het bestand van de Bibliotheek van de LH. Jammer genoeg zijn de resultaten van het hier gerapporteerde onderzoek[38] al weer vrij verouderd, zodat zij nu voor de acquisitie niet meer van zoveel belang zijn. De belangstelling van de klanten van de manuele attenderingsdienst wijzigt zich ook voortdurend, maar daarbij is het mogelijk dat klanten die de dekking van hun onderwerp door deze attenderingsdienst slecht vinden, hun abonnement ook snel weer opzeggen, voordat een veranderde acquisitie iets aan het uitgangsmateriaal heeft kunnen doen.

Bij de computer-attendering wordt het uitgangsmateriaal bepaald door het totaal van de gebruikte bestanden. Hierop wordt in de volgende paragraaf nog nader ingegaan. Bepaalde categorieën publikaties zullen beter in de Bibliotheek van de LH vertegenwoordigd zijn dan in de internationale bibliografieën. Dit betreft o.a. de Nederlandse rapportliteratuur en artikelen in min of meer populaire (Nederlandse) bladen. Bij de profielen 'spinazie' (12) en 'economie van de zetmeelproduktie' (19) waren met name veel van de bij de manuele attendering gevonden titels uit deze bronnen afkomstig.

Het uitgangsmateriaal is natuurlijk ook verschillend voor de diverse magneetbanden; een goede kennis van de aard van dit uitgangsmateriaal, de manier waarop en de mate waarin het toegankelijk is gemaakt, is van groot belang bij de keuze en het gebruik van een computerbestand. De informatiespecialist dient zo goed mogelijk over deze zaken geïnformeerd te zijn, zodat hij zijn klant zo goed mogelijk kan adviseren. In 5.4.2 wordt op dit aspect nog nader ingegaan.

Computerattendering kan in een aantal gevallen ook toegang verlenen tot uitgangsmateriaal dat minder in de collectie van de LH-bibliotheek is vertegenwoordigd. Dit betreft vooral de randgebieden van de landbouwwetenschappen, zoals chemie, biologie, medicijnen, natuurkunde, techniek en sociale wetenschappen. Ook octrooien en lopende onder-

38. De bij dit onderzoek gebruikte onderwerpen op *biochemisch* terrein werden vooral slecht door de manuele attendering gedekt (profielnrs. 2, 3, 6, 10, 18, 20).

zoeksprojecten worden door de manuele attendering niet gedekt, terwijl daar wel computer-attendering op mogelijk is. (Er zou natuurlijk ook manueel geattendeerd kunnen worden op deze onderwerpen, door gebruik te maken van de diverse referaattijdschriften, die veelal wel op de LH-bibliotheek aanwezig zijn.)

Een tweede factor, die althans voor de computer-attendering bij het uitgangsmateriaal een rol speelt, is het informatiegehalte van de titels. Er mag van worden uitgegaan dat het gebruikte computerprogramma zo goed als alle selectiemogelijkheden biedt die de huidige kennis op dit gebied mogelijk maakt. Dat er toch verscheidene relevante titels die zich wèl in het uitgangsmateriaal bevinden, niet worden gevonden, moet dan geweten worden aan het - voor het gebruikte computer-zoekprogramma - te geringe informatiegehalte van die titel. Dit doet zich vooral voor bij zeer vage titels, waaraan verder geen trefwoorden of coderingen zijn toegevoegd. Bij het profiel 'organisatie van het landbouwkundig onderzoek' kwamen bijvoorbeeld veel artikelen met zeer vage titels voor, zoals 'Agricultural Research in New Zealand' en 'Introduction of vegetable research in South East Asia'. Voor het computer-programma zou selecteren op het woord RESEARCH alleen veel te veel irrelevante titels hebben opgeleverd, terwijl de combinatie van RESEARCH met andere trefwoorden snel een te nauwe begrenzing oplevert.

Daarnaast zijn er ook gevallen waarbij zelfs een uitgebreid referaat nog niet duidelijk kan maken of een artikel relevant is voor een bepaald onderwerp. Dit komt bijvoorbeeld voor als de benadering van waaruit het referaat geschreven is, geheel anders is dan de interesse van de klant. Met name bij methodologische onderwerpen kan dit het geval zijn, omdat de methode van het onderzoek bij de titel, de codering en het referaat meestal veel minder op de voorgrond staat dan het object en het resultaat van het onderzoek. Bij manuele attendering zal men veel gemakkelijker op gebruikte onderzoeksmethodes kunnen selecteren.

Als men dus verwacht dat het relevante uitgangsmateriaal van de LH-bibliotheek slecht door computerdiensten wordt gedekt, zal men eerder tot het gebruik van de manuele attendering komen, evenals in gevallen waar selectie met de computer bijzonder moeilijk is, zoals bij methodologie en bij veel onderwerpen beschreven onder pakkende maar nietszeggende titels.

5.4.2 Keuze van computerbestanden

Bij de computerattendering speelt vooral de keuze van het juiste bestand een belangrijke rol. Tijdens dit onderzoek werden veel gegevens verzameld over de overlapping tussen de diverse bestanden voor ieder profiel, maar deze cijfers hebben als gemiddelden weinig betekenis.[39] Alleen voor ieder afzonderlijk onderwerp kan een optimale combinatie van bestanden worden vastgesteld. Deze is natuurlijk geheel afhankelijk van het onderwerp en de beschikbare bestanden. In het ideale geval zou men eerst een profiel enige tijd moeten

39. Naast de 25 profielen, waarbij de uitkomsten met die van de manuele attendering werden vergeleken, werd nog van 26 profielen nagegaan hoeveel in 1973 gepubliceerde titels in meer dan één bestand werden gevonden. Deze gegevens zijn in dit rapport niet opgenomen, omdat uit de uitkomsten geen algemene conclusies kunnen worden getrokken.

Tabel 4. Overlapping voor het profiel 'dierlijke destructieprodukten' (profielnr. 16).

Totaal aantal titels: 371

uniek CAIN	- 118	titels	(38%)
CAIN/man.	- 17	"	
CAIN/CAC	- 7	"	
CAIN/FSTA	- 10	"	
CAIN/man./CAC	- 1	"	
CAIN/man./FSTA	- 2	"	
CAIN/CAC/FSTA	- 9	"	
CAIN/man./cac/FSTA	- 4	"	
totaal CAIN	- 168	"	(45%)
man./CAC	- 9	"	
man./FSTA	- 1	"	
man./CAC/FSTA	- 2	"	
CAC/FSTA	- 5	"	
uniek CAC	- 74	"	
uniek manueel	- 64	"	
uniek FSTA	- 48	"	

'CAIN/man./CAC' betekent: aantal titels dat behalve in CAIN ook bij de manuele attendering en in Chemical Abstracts werd gevonden.

laten lopen op alle beschikbare bestanden, waarvan maar enigszins verwacht kan worden dat ze relevant materiaal bevatten.

Met behulp van het aantal gevonden relevante titels in ieder bestand, de overlapping tussen de bestanden en de abonnementsprijzen kan de meest verantwoorde samenstelling worden bepaald. Een voorbeeld moge dit verduidelijken. Uit de cijfers in Tabel 4 blijkt dat voor het profiel 'dierlijke destructieprodukten' (profielnr. 16) CAIN het grootste aantal titels opleverde, nl. 168 (45% van alle gevonden relevante titels). Manuele attendering leverde 101 titels (27%), CAC 111 (30%) en FSTA 81 (23%).

Beginnend met CAIN leverde CAC 90 extra titels op, manuele attendering voegde aan de combinatie CAIN+CAC nog 65 titels toe, en FSTA voegde 48 titels toe aan de combinatie CAIN+CAC+manueel.

De keuze van de (combinatie van) bestanden hangt in de praktijk af van de waarde die gehecht wordt aan de additionele relevante titels die een abonnement op een extra bestand oplevert en van eventuele budgettaire beperkingen. Uitgaande van de voor 1977 geldende Pudoc-prijzen (zie Bijlage 12) kunnen de gemiddelde en marginale kosten per relevante titel per bestandcombinatie worden berekend. Deze zijn vermeld in Tabel 5.

Als men slechts *f* 500 ter beschikking heeft, zal men altijd een abonnement op CAIN nemen: men krijgt dan immers de meeste waar voor zijn geld. Voor *f* 600 levert de combinatie CAIN+FSTA de meeste titels. Een extra FSTA-titel kost dan *f* 3,57. Indien een extra titel wel *f* 6,52 waard is en men heeft *f* 1076 ter beschikking, dan is CAIN+FSTA+manueel de beste combinatie. Als ten slotte een extra titel wel *f* 8,11 waard geacht wordt en men *f* 1676 ter beschikking heeft, dan kan men een abonnement op alle vier de bestanden nemen.

Tabel 5. Gemiddelde en marginale kosten per titel voor abonnementen op verschillende combinaties van bestanden gebaseerd op de Pudoc-prijzen per 1-1-1977 voor het profiel 'dierlijke destructieprodukten'.

	Abonnements-prijs	Aantal relevante titels	Kosten per titel	Extra titels	Kosten per extra titel
CAIN	f 500	168	f 2,98		
CAC	f 600	111	f 5,41		
Manueel	f 476	101	f 4,72		
FSTA	f 400	81	f 4,94		
CAIN+CAC	f 1100	258	f 4,26	90	f 6,67
CAIN+man.	f 976	244	f 4,00	76	f 6,26
CAIN+FSTA	f 600	224	f 2,60	56	f 3,57
CAIN+FSTA+CAC	f 1200	307	f 3,91	83	f 7,23
CAIN+FSTA+man.	f 1076	297	f 3,62	73	f 6,52
CAIN+FSTA+man.+ CAC	f 1676	371	f 4,52	74	f 8,11

De extra titels zijn die titels die extra worden gevonden als het laatstgenoemde bestand wordt toegevoegd aan dat waarop reeds een abonnement verondersteld wordt.

Indien geen overlappingspercentages bekend zijn, zal veel afhangen van de ervaring van de informatiedienst die de klant moet adviseren, en van de indrukken die een aantal tests kunnen opleveren. Gezien het tijdsverschil zullen tests echter weinig inzicht in de overlapping kunnen geven.

5.4.3 Verdere voor- en nadelen van beide systemen

5.4.3.1 Verkrijgbaarheid van de originele documenten

Bij manuele attendering, gebaseerd op originele documenten, is men er steeds zeker van dat het originele document ter plaatse aanwezig is. Bij de manuele attenderingsdienst van Pudoc is het mogelijk direct fotocopieën van de gesignaleerde documenten te verkrijgen, zodat het origineel *onmiddellijk* ter beschikking wordt gesteld zonder dat het apart behoeft te worden aangevraagd. Bij attendering gebaseerd op - veelal buitenlandse - bibliografische diensten is de lokale beschikbaarheid van het originele document niet gegarandeerd. Gebleken is dat de gebruikers van de computer-attenderingsdienst vaak op interessante titels stuitten, waarvan het origineel voor hen moeilijk te verkrijgen was. 'Moeilijk' betekent hier zowel 'er moest veel moeite voor worden gedaan' als 'het duurde erg lang'. Het eerste aspect duidt vooral op een aanvankelijk gebrek aan ervaring bij het bibliotheekpersoneel in het aanvragen van documenten van elders. Met name moeten de verantwoordelijke personen goed kunnen beoordelen welke aanvraag het beste door welke bibliotheek kan worden gehonoreerd. Zij moeten daarbij ook goed op de hoogte zijn van de mogelijkheden die diverse buitenlandse bibliotheken bieden en de administratieve procedures die daarvoor zijn vereist. Het rondsturen van aanvragen langs Nederlandse bibliotheken heeft lang niet altijd zin en is zeer tijdrovend. Aanvragen voortgekomen uit het ALADIN-

project hebben er mede toe geleid dat de procedures bij de LH-bibliotheek voor aanvragen uit het buitenland aanzienlijk zijn verbeterd.

Bij een aanvraag van elders, noodzakelijk wanneer een publikatie niet in Wageningen aanwezig is, moet men soms veel geduld oefenen. Het verkrijgen van een publikatie uit de V.S. kan wel twee maanden duren. Microfiches per luchtpost kunnen wel veel sneller gaan, maar de acceptatie van microfiches door gebruikers gaat nog maar langzaam. Microfiches bieden ook de mogelijkheid in Nederland zelf bepaalde collecties documenten ter plaatse ter beschikking te krijgen (bijv. de collecties van ERIC en NTIS[40]).

Behalve bij ERIC en NTIS wordt ook bij CAIN de verkrijgbaarheid van het origineel gegarandeerd. Hierbij staat het bibliotheeknummer van de NAL op ieder titelkaartje vermeld. De LH-bibliotheek beschikt over aanvraagkaarten voor de NAL. Bij AGRINDEX garandeert het nationale invoercentrum de verkrijgbaarheid van het origineel. Er bestaat echter nog nauwelijks ervaring met het opvragen van publikaties bij deze centra. Bij de CAB zijn in het algemeen kopieën van de originele documenten verkrijgbaar via de diverse CAB-bureaus.

Voor een aantal andere bestanden wordt de verkrijgbaarheid van de originele teksten echter niet gegarandeerd. Dit is met name het geval bij Chemical Abstracts en Biological Abstracts. Als hierin gevonden titels niet in Nederland ter beschikking zijn, kan de British Library Lending Division vaak uitkomst bieden.

Sommige gebruikers blijven liever onbekend met het bestaan van bepaalde literatuur dan dat zij wèl van het bestaan op de hoogte zijn maar de literatuur moeilijk kunnen verkrijgen of niet kunnen lezen. Deze houding mag men het systeem niet aanrekenen.

5.4.3.2 Precisie

Een ander nadeel van de computerattendering kan zijn dat men nogal wat kaartjes ontvangt die irrelevante titels bevatten. Dit wordt vooral veroorzaakt door het feit dat de computer zoekt naar het vóórkomen van opgegeven woordcombinaties, waarbij de onderlinge samenhang (betekenis) van de in één titel voorkomende woorden voor de computer geen rol speelt. Daarnaast zouden sommige profielen onhandelbaar lange trefwoordenlijsten vereisen als het profiel zó werd geschreven dat zo goed als alle goede combinaties specifiek worden opgenomen en de slechte uitgesloten. Economische overwegingen leiden er dan toe dat het aantal termen wordt beperkt en een zekere hoeveelheid irrelevante titels ('ruis') op de koop toe wordt genomen (zonder dat daarmee het aantal gevonden relevante titels behoeft te verminderen).

Als vuistregel werd in het begin 50% relevantie als onderste grens gehanteerd. Later werd echter het begrip 'nuisance value' meer gehanteerd, waarmee tot uitdrukking wordt gebracht dat bij de vraag of het resultaat acceptabel is het *absolute aantal kaartjes dat weggegooid moet worden* een wellicht even grote rol speelt als het percentage relevante titels. De indruk bestaat dat het weggooien van acht kaartjes bij het vinden van twee goede vaak wel acceptabel is, maar het weggooien van 80 kaartjes bij 20 goede niet meer. Het hangt ook erg van de klant en de aard van zijn profiel af of hij grote aantallen en/of een hoog percentage irrelevante titels voor lief neemt om zo min mogelijk goede titels

40. ERIC = Educational Resources Information Center (V.S.); NTIS = National Technical Information Service (V.S.). Beide zijn een soort rapportencentrale. De laatste bevat het overgrote deel van alle rapporten die zijn ontstaan als resultaat van een opdracht van de Amerikaanse overheid.

te missen. Bij de manuele attendering kan men er vrijwel steeds verzekerd van zijn dat zo goed als alle titels relevant zijn. Is dit niet (meer) het geval, dan moet door een gesprek met de literatuuronderzoeker het profiel van het onderwerp worden aangepast.

5.4.3.3 Speciale eisen aan de opdracht

Bij manuele attendering is in feite ieder selectiecriterium mogelijk. Bij computer-attendering kan alleen geselecteerd worden op de informatie die op de magneetband staat; daarbinnen zijn echter verscheidene criteria in de praktijk niet goed bruikbaar. Zo is selectie op een beperkt tijdschriftenbestand wel mogelijk, maar economisch moeilijk te verantwoorden als de lijst met tijdschriften erg lang wordt (bijvoorbeeld met het doel alle 'onbelangrijke' tijdschriften uit te sluiten). Selectie op taal van het origineel is met de computer wel mogelijk, maar werd bij het project niet toegepast, omdat het doorzoeken van een extra veld bij gering gebruik relatief duur was. Door een nieuwe kostenberekening geldt deze beperking nu echter niet meer en wordt er voor twee profielen op taal gezocht. Ook het zoeken op geografische gebieden is met de computer vaak moeilijk als er geen hiërarchische geografische codering is. Als bijvoorbeeld literatuur die betrekking heeft op ontwikkelingsgebieden moet worden uitgesloten, is het ondoenlijk een lijst met alle namen van ontwikkelingslanden op te nemen. (AGRINDEX heeft gelukkig wel een coderingssysteem voor dit soort categorieën.) Soms is het selecteren op belangrijkheid van het artikel een wens van de klant, waaraan met de computer onmogelijk kan worden voldaan, terwijl zoiets vaak wèl aan het inzicht van de literatuuronderzoeker kan worden overgelaten. Hiervoor is het absoluut noodzakelijk te beschikken over het origineel. Er lopen bij de manuele attenderingsdienst enkele profielen over zeer brede onderwerpen, waarover de klant alleen regelmatig enkele belangrijke artikelen wil ontvangen, de selectie naar belangrijkheid overlatend aan de specialist.[41] Ten slotte is het bij de manuele attendering mogelijk alle artikelen met lange literatuurlijsten (bijvoorbeeld meer dan 100 referenties) te laten selecteren. Lang niet alle computer-bestanden bevatten echter informatie over het aantal referenties.

5.5 CONCLUSIES

Na een periode van twee jaar, waarin de resultaten van de manuele literatuurattendering en computer-attendering werden vergeleken, kunnen de volgende conclusies worden getrokken.

1. Voor het merendeel van de attenderingsonderwerpen kan manuele attendering door computer-attendering worden vervangen in die zin dat de dekking van de literatuur er niet door

41. Ook met het opsporingsprogramma van IWIS-FLIRT kan een zekere mate van selectie op belangrijkheid worden toegepast. Men kan bijvoorbeeld alle trefwoorden gewichten geven, een hoger gewicht gebruikend naarmate het woord belangrijker wordt geacht. De titels komen er dan uit in volgorde van totaal gewicht, waarbij kan worden afgekapt na een gesteld drempelgewicht. Men kan ook afkappen na een bepaald aantal titels. Van deze mogelijkheid is echter in de praktijk nooit gebruik gemaakt, omdat er een groot element van willekeur in zit.

vermindert; in de meeste gevallen levert computer-attendering zelfs meer relevante titels op dan manuele attendering.

2. De beste dekking van de literatuur wordt verkregen door een combinatie van beide systemen, omdat met ieder systeem steeds een aantal titels zal worden gevonden die met het andere systeem worden gemist.

3. Manuele attendering is gemiddeld wat sneller dan computer-attendering, maar het verschil bedraagt meestal slechts een paar maanden.

4. De verkrijgbaarheid van de originele teksten is bij manuele attendering steeds verzekerd. De toezending kan automatisch geschieden samen met de attenderingskaartjes. Bij computer-attendering moeten de (kopieën van de) originele documenten steeds apart worden aangevraagd en is het soms moeilijk deze te verkrijgen.

5. Er zijn een aantal vakgebieden en soorten informatie die door manuele attendering, gebaseerd op het bij de Bibliotheek van de Landbouwhogeschool binnenkomende primaire materiaal, slecht of niet worden gedekt, maar door computer-attendering goed kunnen worden verzorgd (randgebieden en hulpwetenschappen voor de landbouw, octrooien en beschrijvingen van lopend onderzoek).

6. Een aantal soorten onderwerpen en selectiecriteria zijn voor manuele attendering wèl geschikt , maar voor computer-attendering slecht of in het geheel niet. Voor deze gevallen biedt manuele attendering grote voordelen boven computer-attendering. Het betreft hier o.a.

- selectie uit semi-wetenschappelijke literatuur,
- selectie uit een beperkte tijdschriftenlijst,
- selectie op grotere geografische eenheden,
- selectie op aspecten van de inhoud die niet in de titel of in het referaat zijn vermeld,
- selectie beperkt tot 'belangrijke' artikelen.

7.Bij computer-attendering komen steeds een aantal irrelevante titels mee, terwijl bij manuele attendering bijna alle titels relevant zijn. Voor sommige klanten is een groot aantal en/of percentage irrelevante titels een bezwaar.

6 Waardering door de gebruikers

De leden van de experimentele gebruikerskring werden geselecteerd uit de wetenschappelijke staf van landbouwkundige onderzoeksinstellingen. Er is daarbij wel getracht uit zoveel mogelijk geledingen van het landbouwkundige onderzoek mensen bij het project te betrekken, maar er is niet naar een bewuste stratificatie van de steekproefruimte gestreefd. Van de in totaal 68 deelnemers waren er 29 werkzaam bij instituten van de Directie Landbouwkundig Onderzoek, 23 bij de Landbouwhogeschool, 6 bij de industrie en 10 elders. Zeven van hen hielden zich met enige vorm van dierkundig onderzoek bezig, 15 met plantkundig onderzoek en 24 met technologie. Ten slotte waren er nog 4 bodemkundige en 4 economische onderwerpen, 2 op het gebied van de voeding (van de mens) en ook 2 gericht op de landbouw in een bepaald gebied. Een lijst van profielonderwerpen is toegevoegd als Bijlage 6. De instellingen resp. LH-vakgroepen waartoe de deelnemers behoorden, staan vermeld in Bijlage 7.

De deelname aan het project was gratis, waarbij echter als tegenprestatie beoordeling van de resultaten werd gevraagd. De gebruikers is op verschillende manieren gevraagd de resultaten van de dienstverlening te evalueren. Dit gebeurde allereerst continu door de beoordeling van de relevantie en nieuwswaarde van de geleverde titels (6.1), vervolgens door een enquête betreffende de bereidheid tot continuering van het abonnement tegen betaling. De meeste directe uitspraak over het nut van de attenderingsdienst voor de gebruikers werd echter geleverd door de mate waarin zij tot abonnering tegen betaling overgingen (6.2). Onder de opzeggers werd ten slotte nog een enquête gehouden over de redenen van beeindiging (6.3). In sectie 6.3 komen ook een aantal meer algemene beperkingen van computer-attendering aan de orde.

6.1 BEOORDELING VAN DE RELEVANTIE EN DE NIEUWSWAARDE VAN DE GEVONDEN TITELS

De belangrijkste tegenprestatie die van de deelnemers aan het experiment werd gevraagd, was een regelmatige evaluatie van de toegestuurde titels. De deelnemers ontvingen daartoe steeds alle titels in tweevoud, waarbij één set met een evaluatie-aanduiding moest worden teruggestuurd. De medewerking van de deelnemers is wat dit betreft bijzonder groot geweest: zo goed als alle titels werden beoordeeld terug ontvangen.

Er werd gevraagd twee onafhankelijke criteria te beoordelen: enerzijds de *relevantie* en anderzijds de *nieuwswaarde*. Bij het relevantiecriterium ging het om de vraag *of* de titel voor het onderzoek of het interessegebied van de gebruiker relevant (ter zake) was, niet *waarom* een gevonden titel eventueel irrelevant was. Gevraagd werd onderscheid te maken tussen 'zeer relevant' (++), 'relevant' (+), en 'randgebied' of 'relevantie aan de titel alleen niet te beoordelen' (±) (zie Bijlage 8).

Bij de relevantiebeoordeling deed zich het probleem voor van titels in moeilijk toe-

gankelijke talen. Sommige gebruikers hadden de neiging relevant lijkende titels van artikelen in een taal die zijzelf niet konden lezen en ook niet gemakkelijk (snel en kosteloos) vertaald konden krijgen als 'irrelevant' aan te merken. Dit is onjuist omdat de titel op zichzelf relevant leek en het dus correct was dat hij werd gevonden, tenzij moeilijk toegankelijke talen als groep zouden worden uitgesloten. Gelukkig vonden zo goed als alle gebruikers dat laatste echter een te radicale beperking. Aanbevolen werd om in voorkomende gevallen hiervoor de ± aanduiding te gebruiken.[42]

Voor wat betreft 'nieuwswaarde' werd gevraagd alle relevante titels die de gebruiker nog niet eerder had gezien, met een 'N' te merken. Ervaring leerde, dat sommige gebruikers géén N toekenden als het een titel betrof uit een tijdschrift dat ze normaliter onder ogen behoorden te krijgen, hoewel ze het op het moment dat het attenderingskaartje arriveerde, nog niet hadden gezien. Temeer omdat de nieuwswaarde direct samenhangt met de snelheid van de attendering, werd daarom later speciaal verzocht in zulke gevallen ook een N toe te kennen.

6.1.1 Het belang van de relevantie-parameter

Oorspronkelijk werd de uitkomst van de relevantiebeoordeling als één der belangrijkste parameters voor de waardering van het systeem gezien. Het grote gewicht dat ook in de literatuur aan deze parameter gehecht werd, is echter recent nogal gerelativeerd. Een lage relevantie, waarbij dus iedere zending titels veel niet-relevant materiaal bevat, zou voor de gebruiker inderdaad hinderlijk kunnen zijn en duiden op een slechte profielopstelling, maar dat is niet noodzakelijk. Een hoog percentage irrelevante titels (ruis) is voor veel gebruikers pas hinderlijk als het absolute aantal daarvan ook groot wordt; het wordt vaak niet als hinderlijk ondervonden als 7 van de 10 titels moeten worden weggegooid, maar wel als er 70 van de 100 irrelevant blijken te zijn (het 'nuisance effect', dat reeds in 5.4.3.2 werd besproken). Een laag relevantiepercentage kan ook onvermijdelijk zijn als de gebruiker er grote waarde aan hecht in ieder geval zo min mogelijk relevante titels te missen. Een optimaal zoekprofiel moet ertoe leiden dat er zo min mogelijk relevante titels worden gemist, terwijl irrelevante titels zoveel mogelijk vermeden worden. Deze twee principes zijn echter vaak strijdig. Uiteindelijk hangt het dan van de gebruiker af hoe nadelig het voor hem is eventueel relevante titels te missen.

Als vuistregel werd door ons uiteindelijk het volgende principe gebruikt: Zolang het relevantiepercentage hoger is dan 50%, is het niet opportuun het profiel bij te schaven, tenzij de gebruiker liever minder ruis wil; bij lagere relevantie moet nader worden onderzocht of een andere profielopstelling mogelijk en wenselijk is. Bij een in absolute zin

---

42. Achteraf bezien was dit wellicht geen juist advies. De relevantie mag in het geheel niet van de taal afhangen. De achterliggende gedachte bij het advies was meer een beoordeling op uiteindelijke 'follow-up' dan op relevantie. 'Follow-up' wordt nl. ook wel als relevantiecriterium gebruikt, waarbij dan onderscheid kan worden gemaakt naar de actie die men naar aanleiding van de gevonden titel onderneemt: 'wil het artikel zo snel mogelijk lezen', 'vraag (fotocopie van) het artikel aan om eens te lezen als ik tijd heb', 'bewaar de titel om misschien het artikel eens aan te vragen als ik me intensief met dat specifieke onderwerp ga bezighouden', etc.

geringe opbrengst is een brede opstelling met wellicht veel ruis te verkiezen boven een enge formulering die nog minder relevante titels zou opleveren.

Ook nu de evaluatiestudie voorbij is, wordt aan nieuwe gebruikers nog steeds gevraagd de titels van de eerste drie zendingen op hun relevantie te beoordelen, zodat aan de hand daarvan het profiel eventueel kan worden bijgestuurd.

Ten slotte nog een opmerking over *rangschikking van de uitvoer naar aflopende relevantie*. Als het absolute aantal titels een probleem gaat vormen en het niet van zo groot belang is dat zoveel mogelijk relevante titels worden gevonden, zou men kunnen trachten de uitvoer te rangschikken naar belangrijkheid van de titels en dan niet meer dan een van tevoren bepaald aantal titels laten afdrukken. Ook zou men alleen die titels kunnen laten uitvoeren die aan een bepaald criterium van belangrijkheid voldoen (zie ook 5.4.3.3).

Indien de belangrijkheid van een titel kan worden weergegeven als de som van de gewichten van bepaalde trefwoorden, zijn beide mogelijkheden in principe met het FLIRT-programma te verwezenlijken: de uitvoer vindt nl. automatisch in gewichtsvolgorde plaats; men kan daarbij voor het eerste geval een maximum aan het aantal uit te voeren titels stellen en voor het tweede geval een drempelgewicht aangeven; titels met een gewicht lager dan de drempel worden dan niet afgedrukt (ook al zouden ze eventueel wel logisch relevant zijn). In de praktijk van het ALADIN-project werd slechts sporadisch met deze mogelijkheden geëxperimenteerd, hoofdzakelijk omdat een grote uitvoer van relevante titels nooit tot klachten heeft geleid.

6.1.2 De resultaten van de relevantiebeoordeling

De volgende aantallen titels werden beoordeeld op relevantie resp. nieuwswaarde (tussen haakjes het aantal profielen waarvan de beoordeling werd geanalyseerd).[43]

	relevantie	nieuwswaarde
CAIN	12 425 (59)	9 787 (53)
FSTA	4 224 (16)	3 668 (13)
CAC	13 776 (22)	8 780 (16)
BA	3 566 (8)	3 566 (8)
Medlars	1 172 (6)	1 172 (6)

De *relevantiebeoordeling* van deze titels is samengevat in Tabel 6. Als grove benadering aannemend dat 50% van de *misschien* relevante titels uiteindelijk relevant wordt bevonden, resulteert een relevantiepercentage van 53 voor CAIN, 52 voor FSTA, 37 voor CAC, 36 voor BA en 60 voor Medlars.[44]

---

43. Hier werden alleen de analyseresultaten vermeld van die profielen waarbij de onderscheiding tussen 'zeer relevant' en 'relevant' ook daadwerkelijk werd toegepast. Een aantal gebruikers vond deze onderscheiding te moeilijk en gebruikte daarom de '++' categorie niet. Voor CAIN betrof dit 10 profielen. De 25 profielen die met de resultaten van de manuele attendering werden vergeleken, werden door de projectleider op relevantie beoordeeld, maar ook zonder de '++' en '+' onderscheiding. De dan nog resterende discrepantie met de vermelde aantallen verwerkte profielen in Tabel 6 is te verklaren uit een te gering aantal beschikbare beoordelingen van die profielen.

Indien na de eerste runs aanzienlijke profielwijzigingen werden doorgevoerd om het profiel te verbeteren, werden de beoordelingen van deze eerste runs niet bij de analyse betrokken.

44. Indien ook alle titels in de analyse worden betrokken waarvoor geen onderscheid tussen '++' en '+' werd gemaakt, weken de uitkomsten weinig af van de reeds vermelde, nl.: 52 voor CAIN, 55 voor FSTA, 38 voor CAC, 39 voor BA en 57 voor Medlars.

Tabel 6. Relevantiebeoordeling in percentages van alle gevonden titels.

	++	+	±	-	±/++,+,±	++/++,+
CAIN	14	28	20	38	0,30	0,33
FSTA	6	35	21	38	0,28	0,15
CAC	5	23	17	55	0,35	0,18
BA	16	13	14	57	0,25	0,55
Medlars	13	39	16	36	0,22	0,25

De berekeningen werden gemaakt door de gemiddelden van ieder profiel te middelen. Als alle titels als één verzameling worden beschouwd (waarbij de profielen met een hoge titelopbrengst relatief overgewaardeerd worden), komen de relevantiepercentages meestal iets lager uit: profielen met een relatief grote opbrengst aan titels hadden dus blijkbaar een wat lager relevantiepercentage dan de profielen met een kleiner aantal titels.

Tabel 7. Frequentieverdeling voor de relevantiepercentages 'relevant' en 'misschien relevant' samen. (Om de getallen vergelijkbaar te maken zijn ze tot percentages omgerekend.)

Bestand	0-25%	26-50%	51-75%	76-100%
CAIN	3	34	34	29
FSTA	0	23	45	32
CAC	12	47	26	15
BA	20	50	20	10
Medlars	0	22	56	22

Het is opvallend dat het percentage 'irrelevant' bij CAC relatief hoog is, zulks ondanks het feit dat er bij CAC nogal wat trefwoorden aan de titel worden toegevoegd. Latere ervaringen met het bestand van de Commonwealth Agricultural Bureaux hebben geleerd dat het toevoegen van veel extra trefwoorden de relevantie negatief kan beïnvloeden, omdat bepaalde trefwoorden al worden gebruikt als de tekst er slechts zijdelings verband mee houdt.[45] Dit is ook heel duidelijk het geval voor de codes die bij BA worden toegevoegd. Hier is echter weer een verfijning aangebracht door de codes naar drie niveaus van belangrijkheid te onderscheiden, zodat men op een bepaalde code kan selecteren, naar 'major importance' of naar 'secondary importance', etc.

Zoals uit Tabel 7 blijkt kan eigenlijk aan de gemiddelden in absolute zin niet zoveel waarde worden gehecht, gezien de grote verschillen tussen de individuele profielen. Of het relevantiepercentage voor een bepaald profiel als 'hoog' of 'laag' moet worden beoordeeld, hangt in hoge mate af van het gebruik van de resultaten door de klant, en van de mogelijkheden die het computer-programma t.o.v. het bestand biedt. De vergelijkingen tussen de relevantiepercentages van de diverse bestanden zijn ook slechts van beperkte waarde omdat de verzamelingen profielen verschillend waren. In feite zou men de relevantie van bijv. CAC ten opzichte van CAIN moeten vergelijken voor precies hetzelfde profielen-

45. Een groot aantal toegevoegde trefwoorden kan ook tot ongewenste treffers leiden als bij de profielopstelling een gewichtstructuur met drempelgewicht wordt toegepast. Als aan een titel termen zijn toegevoegd die een alternatief vormen voor reeds in de titel voorkomende woorden, en zowel het titelwoord als het toegevoegde alternatief in het profiel met een bepaald gewicht voorkomen, dan werkt dit gewichtsverhogend, hetgeen tot onbedoelde overschrijding van de drempel kan leiden.

Tabel 8. Gemiddeld aantal titels per jaar per profiel.

	++	+	±	-	Totaal
CAIN	22	44	35	59	160
FSTA	10	66	42	70	184
CAC	10	84	48	199	355
BA	30	95	86	515	742
Medlars	22	84	21	52	178

De verschillen tussen de totalen van de eerste vier kolommen en de vijfde kolom zijn ontstaan door afrondingen.

Deze getallen zijn berekend door het gemiddeld aantal hits per run te vermenigvuldigen met het aantal runs per jaar. Dit laatste aantal bedraagt voor CAIN, FSTA en Medlars 12 (éénmaal per maand), voor Chemical Abstracts 24 en voor BA 36.

bestand, terwijl nu alle CAIN-profielen geplaatst worden tegenover alle CAC-profielen, etc.

Om na te gaan hoe de resultaten van de gebruikte bestanden zich onderling verhouden wat betreft het aandeel van de zeer relevante titels in het totaal aantal relevante titels werden de percentages '++' titels van de som van de '++' en '+' titels berekend. BA kwam daarbij als hoogste op 55%. Gezien het geringe aantal profielen waarop dit getal is gebaseerd en de grote spreiding tussen de individuele uitkomsten, kan aan dit getal echter weinig waarde worden gehecht. CAIN kwam op 33%, Medlars op 25%, CAC op 18% en FSTA op 15%. Het lage percentage voor CAC kan misschien verklaard worden uit de omstandigheid dat de meeste onderwerpen hier in absolute zin bijzonder veel relevante literatuur opleverden, waarbij toch maar enkele titels er als zéér relevant uitspringen. Voor het lage percentage bij FSTA is niet direct een verklaring voorhanden.

Het percentage dat de '±' titels uitmaken van alle relevante en misschien relevante titels, kan een indruk geven hoe moeilijk het was de titels van diverse bestanden te beoordelen. Toegevoegde trefwoorden zouden het bijvoorbeeld gemakkelijker kunnen maken om tussen 'relevant' en 'irrelevant' te kiezen. Dit wordt door de uitkomsten echter niet bevestigd: CAC 35%, CAIN 30%, FSTA 28%, BA 25% en Medlars 22%. Alleen CAIN heeft géén toegevoegde trefwoorden, de andere bestanden wel. Met betrekking tot CAC zou het hoge percentage ook kunnen wijzen op een relatief groot aantal titels uit randgebieden.

Geconcludeerd kan worden dat het CAIN-bestand zich nergens negatief onderscheidt van de andere bestanden, ondanks het relatief lage informatiegehalte van de records van dit bestand; ze bevatten immers nauwelijks meer dan titels. Negatieve effecten van dit lage informatiegehalte zouden wellicht kunnen blijken uit een zgn. recall-onderzoek, waarbij wordt nagegaan welk percentage van alle relevante titels die de CAIN-band bevat, gevonden worden met behulp van een computer-zoekprogramma. Zo'n onderzoek lag echter buiten de opzet van het ALADIN-project. Een indicatie hiervoor is hoogstens het aantal relevante titels dat men per jaar van een bepaald bestand ontvangt. Deze cijfers zijn samengevat in Tabel 8.

In het totaal levert BA 211 'relevante' en 'misschien relevante' titels per jaar op, CAC 142, Medlars 127, FSTA 118 en CAIN 101. CAIN komt hier dus wel achteraan. Het lage relevantiepercentage voor CAC blijkt geenszins te impliceren dat men in absolute zin weinig relevante titels van dit bestand zou ontvangen. Het tegendeel blijkt hier het geval te zijn. De resultaten lijken erop te wijzen dat over typisch landbouwkundige en voedings-

technologische onderwerpen minder wordt gepubliceerd dan over biologische en chemische onderwerpen binnen de landbouwkundige interessesfeer.

Jammer genoeg werd niet geanalyseerd hoeveel relevante titels gemiddeld per onderwerp in alle bestanden samen werden gevonden. De uitkomsten van de telling voor de profielen die met de manuele attendering werden vergeleken kan hiervoor een indicatie geven: uit Bijlage 4 kan worden afgeleid dat er over één jaar voor 25 onderwerpen door de computer 3744 relevante titels werden gevonden, d.i. 150 per onderwerp. Exclusief de 'misschien relevante' titels zouden het er 108 zijn.

Overigens zijn de vergelijkingen tussen de relevantiepercentages van diverse bestanden ook daarom slechts van beperkte waarde omdat de populaties verschillend zijn. In feite zou men de relevantie van bijv. CAC ten opzichte van CAIN moeten vergelijken voor precies hetzelfde profielenbestand, terwijl nu alle CAIN-profielen geplaatst worden tegenover alle CAC-profielen, etc.

Ook al zijn de uitkomsten van de gemiddelde relevantiepercentages dan wellicht niet van dat belang dat er oorspronkelijk aan werd gehecht, de beoordelingen zijn steeds van erg veel waarde geweest bij het verbeteren van de afzonderlijke profielen. Alleen met behulp van grote aantallen relevante (èn irrelevante) titels kan uiteindelijk een goed lopend profiel worden samengesteld.

6.1.3 De resultaten van de nieuwswaardebeoordeling

Voor de vergelijking van de nieuwswaardebeoordeling werd per bestand het totaal aantal als nieuw beoordeelde titels gedeeld op het totaal aantal relevante titels. De uitkomsten van de berekeningen staan vermeld in Tabel 9.

Tabel 9. Percentage nieuwe titels van alle relevante of misschien relevante titels.

Bestand	n/++	n/++,+	n/++,+,±
CAIN	53	60	59
FSTA	25	58	58
CAC	67	71	67
BA	56	62	58
Medlars	36	46	50

'n/++' geeft aan het percentage nieuwe titels van alle zeer relevante titels; 'n/++,+' idem van alle relevante titels en 'n/++,+,±' idem van alle relevante en misschien relevante titels.

Uit deze tabel blijkt dat bij CAIN, CAC en BA gemiddeld iets meer dan de helft van het aantal zeer relevante titels nieuw was. Voor FSTA was dit echter slechts 25%. Dit lage percentage kan veroorzaakt zijn door het feit dat de titelbeschrijvingen d.m.v. FSTA betrekkelijk laat ter beschikking komen, zodat in ieder geval het belangrijkste materiaal al langs andere weg ter kennis van de gebruiker is gekomen. Voor BA, waar de vertraging

Tabel 10. Gemiddeld aantal relevante nieuwe titels per profiel per jaar.

Bestand	++n	+n	±n	Totaal
CAIN	11	26	17	54
FSTA	2	35	26	63
CAC	5	58	31	94
BA	19	72	54	145
Medlars	12	54	11	77

in feite nog groter is, blijkt dit argument echter niet op te gaan.[46] Voor alle relevante titels samen ligt het percenatge nieuwe titels in het algemeen wat hoger, nl. voor vier bestanden rond 60%, voor FSTA op 46%. Inclusief de misschien relevante titels wijzigt dit beeld zich niet sterk meer, behalve dat de positie van FSTA zich relatief verbetert. Het totale aantal relevante nieuwe titels dat ieder bestand per jaar leverde, is vermeld in Tabel 10.

Als alleen de zeer relevante nieuwe titels werkelijk van belang zouden zijn (bijvoorbeeld als dat de enige titels zouden zijn waarvan het origineel zou worden opgevraagd naar aanleiding van de attendering), dan zou BA ook het beste uit de bus komen. Mogelijk bevat dit bestand meer titels die landbouwkundigen minder snel reeds uit andere bron onder ogen krijgen, terwijl ze toch van groot belang voor hen blijken te zijn! Bedacht moet echter worden dat de cijfers voor BA maar op een beperkt aantal (8) profielen zijn gebaseerd. Het lage aantal 'zeer relevante' nieuwe titels voor FSTA en CAC is opvallend. Voor FSTA zou dit kunnen komen door de relatieve ouderdom van de titels, voor CAC kan dat argument echter niet gelden. Voor alle nieuwe relevante titels samen is de rangorde voor de vijf bestanden dezelfde als die voor de nieuwe en reeds bekende samen.

Het is moeilijk een waarde-oordeel uit te spreken over deze resultaten. Men kan de opvatting verdedigen dat de gebruikers al vrij veel materiaal uit andere bron kenden voordat ze er door het computer-systeem op geattendeerd werden; anderzijds was toch meer dan de helft van de door dit systeem geleverde relevante en misschien relevante titels voor de gebruikers nieuw. Men mag weliswaar niet zeggen dat zij deze titels zonder dit systeem dan ook nooit onder ogen zouden hebben gekregen, maar de conclusie lijkt toch gerechtvaardigd dat hun een aanzienlijk aantal relevante titels ontgaan zou zijn zonder attendering. Voor de 'zeer relevante' titels ligt het percentage weliswaar lager, maar het is toch meestal nog meer dan de helft. De hypothese dat de wetenschappelijke onderzoeker het merendeel van de belangrijkste literatuur zonder speciale attendering toch wel in handen krijgt vindt geen steun in de uitkomsten van dit onderzoek. Als hij al meer dan 50% van de belangrijkste literatuur in handen krijgt, zal dat veelal langer duren zónder attendering dan mèt attendering. Hierbij moet overigens wel worden opgemerkt dat de relevante

46. Het is hierbij ook niet ondenkbaar dat de categorie gebruikers die in FSTA is geïnteresseerd, zelf sneller en meer primaire bronnen raadpleegt dan de categorie waarvoor BA speciaal van belang is. In het algemeen is de uitkomst van de nieuwswaardebeoordeling natuurlijk afhankelijk van de bestaande gewoonten van de gebruiker m.b.t. zijn informatieverzameling.

literatuur die de onderzoeker zelf vindt langs andere weg dan via de attenderingsdienst, buiten dit onderzoek valt. Dit kan in sommige gevallen een belangrijk deel van zijn totale literatuurvoorziening zijn (de categorie van gemiste relevante titels).

6.2 CONTINUERING NADAT BETALING WERD GEVRAAGD

De subsidie van NOBIN voor het verwerken van magneetbanden liep af per 1 april 1975 (uitgezonderd de verwerking van CABVET en AGRINDEX, die tot eind 1975 mocht worden gecontinueerd). In 1973 en 1974 konden 100 profielen zonder financiële consequenties voor de gebruiker worden verwerkt; voor 1975 moest echter een andere financiering worden gevonden, welke althans ten dele uit bijdragen van de gebruikers moest bestaan. Alle deelnemers in 1974 kregen dan ook de vraag voorgelegd of zij hun abonnement in 1975 wilden continueren tegen de daarvoor vastgestelde prijzen. De resultaten van dit verzoek zullen in de volgende sectie nader worden geanalyseerd.

Om enige indruk te verkrijgen van de mening van gebruikers over voortzetting tegen betaling en de invloed van de hoogte van de prijs op de bereidheid tot voortzetting, werd in maart 1974 een enquête gehouden onder de gebruikers. Hierin werd gevraagd hoeveel waarde gebruikers aan een attenderingsabonnement hechten, uitgedrukt in guldens. Het daartoe gebruikte enquêteformulier is toegevoegd als Bijlage 9. Er werd daarin ook gevraagd of men ook in staat zou zijn de zelf aangegeven abonnementsprijs voor de tweede helft van 1974 daadwerkelijk te voldoen. Op deze enquête werden 14 antwoorden ontvangen; 12 hiervan vonden een abonnement in ieder geval ƒ 50,- waard, 8 ook ƒ 100,- en 5 ook ƒ 200,-. Eén wilde er tot ƒ 300,- gaan, terwijl er tenslotte 2 onder voorbehoud hun waardering wel tot ƒ 500,- wilden verhogen. Slechts 4 abonnees gaven aan ook daadwerkelijk voor de aangegeven waarde een abonnement te kunnen nemen, gegeven hun budgetaire omstandigheden; 4 deelden mee dit niet te kunnen. De andere abonnees gaven hierop geen antwoord.

6.2.1 Continueringspercentages voor de verschillende bestanden

Voor CAIN geven de volgende cijfers de verandering in het aantal rekeningeenheden[47] nadat betaling werd gevraagd:

aantal rekeningeenheden band 74/12		84
aantal rekeningeenheden band 75/7, doorlopend van 1974		52
waarvan betalend	49	
waarvan niet betalend (voor intern Pudoc-gebruik)	3	
aantal beeindigde rekeningeenheden		32
waarvan onderzoek beëindigd	5	
waarvan 'jaarlijks retrospectief onderzoek zinvoller'	3	
waarvan om andere reden	12	
waarvan vergelijking manuele attendering	12	

47. Per rekeningeenheid wordt eenmaal het abonnementsgeld in rekening gebracht. Een rekeningeenheid kan meer dan één profiel omvatten, indien het uit technische overwegingen beter is een onderwerp te splitsen. In een paar gevallen zijn ook een aantal zeer kleine profielen tot één rekeningeenheid samengevoegd.

Tabel 11. Percentage profielen dat tegen betaling werd voortgezet van alle profielen waarvoor betaling werd gevraagd.

Bestand	Continuerings-percentage	(100% =)*
CAIN (incl. vergelijking manuele attendering)	60	81
CAIN (excl. vergelijking manuele attendering)	67	60
FSTA	54	28
CAC	45	31

* Dit is het absoluut aantal rekeningeenheden waarvoor betaling werd gevraagd

Exclusief de drie Pudoc-profielen waren er dus 81 rekeningeenheden, waarvan er uiteindelijk 49 betalend abonnee zijn geworden, d.i. 60%.

De gebruikers die al betalend abonnee bij de manuele attenderingsdienst waren, vormden uiteraard een enigszins andere groep dan zij die speciaal voor de computer-attendering werden geworven. De eerste groep kreeg de resultaten van de computer-attendering immers geheel ongevraagd thuis gestuurd. Deze groep komt overeen met 21 rekeningeenheden, waarvan er uiteindelijk 9 tot betaling leidden. Bij 2 hiervan werd het manuele attenderingsabonnement beeindigd. In 1975 zegden van de andere 12 abonnees nog eens 8 hun abonnement op en continueerden 4 alleen de manuele attendering. Van de 13 die hun attenderingsabonnement in enigerlei vorm continueerden waren er dus 7 die een abonnement op beide namen, 2 alleen op de computer- en 4 alleen op de manuele attendering.

Voor FSTA bedroeg het aantal opzeggers 13 tegen 15 blijvers, d.i. een betalingspercentage van 54. Op CAC werden per ultimo 1974 38 profielen voor 29 onderwerpen gedraaid. Hiervan waren er 5 voor klanten die niet tot het Ministerie van Landbouw en Visserij behoorden; hiervoor was bij voortzetting geen verdere bemiddeling van Pudoc nodig; zij werden direct naar NOCI doorverwezen. Er liepen 2 profielen alleen nog voor Pudoc zelf voor de vergelijking met de manuele attendering. Er resteerden dus 31 profielen waarvoor betaling gevraagd werd; 14 hiervan continueerden hun abonnement, d.i. 45%. Deze cijfers zijn in Tabel 11 nog eens samengevat.

6.2.2 De continueringspercentages als tevredenheidsindicator

De continueringspercentages kunnen niet zonder meer gebruikt worden als maatstaf voor de tevredenheid van de gebruiker met het geleverde produkt. De hoogte van de kosten spelen hierbij natuurlijk ook een belangrijke rol, waarbij men nog onderscheid moet maken tussen de *waarde van de geleverde informatie en de beschikbaarheid van fondsen*. Er is herhaaldelijk gebleken dat budgettaire belemmeringen betaling onmogelijk maakten, terwijl men objectief gezien het geleverde produkt op zichzelf wel de gevraagde prijs waard vond.

Naast budgettaire inflexibiliteit speelt natuurlijk ook de prijselasticiteit een rol: in het algemeen zal er minder van een bepaald produkt worden gekocht naarmate het duurder is. Men mag dus aannemen dat het continueringspercentage hoger zou zijn geweest bij een lagere prijs en vice versa. Over de invloed van dit principe werd slechts weinig informatie gekregen, al gaf de reeds behandelde kleine enquête wel enig inzicht: boven de f 100,- per jaar was de belangstelling al snel erg gering. Boven de f 500,- was er geen enkele belangstelling. In de praktijk bleek een prijs van f 250,- toch tot een redelijk aantal

klanten te leiden, terwijl een lager bedrag in verband met de kosten en gebrek aan centrale financiering niet haalbaar werd geacht. De gehanteerde principes voor prijsstelling worden nader in hoofdstuk 7 besproken.

6.3 REDENEN VOOR BEEINDIGING

Aan alle personen die hun abonnement niet continueerden (totaal 19), werd een vragenformulier toegestuurd om na te gaan wat de redenen voor het opzeggen waren. Van 14 personen werd een antwoord ontvangen. Het enquêteformulier is opgenomen als Bijlage 10.

Voor 7 van deze 14 waren 'onbevredigende resultaten' de belangrijkste reden voor opzegging, voor 3 was het project intussen beeindigd, voor 2 voldeed de bestaande informatievoorziening reeds goed, terwijl daaraan geen directe kosten verbonden waren; één had er geen geld voor beschikbaar, en één had niet zoveel geld over voor deze dienstverlening, omdat snelle attendering niet zo opportuun werd geacht voor de huidige fase van het onderzoek.

De 'onbevredigende resultaten' hadden zeer verschillende achtergronden. Sommige onderwerpen waren zò specialistisch dat er per maand slechts zeer weinig relevante titels konden worden verwacht. In het algemeen waren de profielen voor die gevallen vrij breed opgesteld om vooral niets te missen, maar dit bevredigde deze gebruikers toch niet. Voor deze gevallen is vermoedelijk een minder frequente attendering beter. Het retrospectieve on-line systeem kan bijvoorbeeld ieder kwartaal geraadpleegd worden, waarbij alleen de in het laatste kwartaal toegevoegde titels worden doorzocht. Alleen als aan snelle attendering zeer grote waarde wordt toegekend, lijkt een maandabonnement voor een zeer beperkt onderwerp zinvol. De kosten per gevonden titel worden immers erg hoog.

In andere gevallen werd de gebruiker te zeer overstroomd met materiaal dat weliswaar veelal relevant was, maar dat hij onmogelijk kon verwerken. Extra selectiecriteria zouden nodig geweest zijn om tot een bevredigend resultaat te komen: alleen gerenommeerde auteurs, alleen overzichtsartikelen, alleen belangrijke tijdschriften, geen moeilijk toegankelijke bronnen, etc. Wellicht had aan de verbetering van de profielopstelling in deze zin meer aandacht moeten worden besteed, maar het is de vraag of het de moeite waard zou zijn geweest. Manuele attendering zou in deze gevallen gemakkelijker met de speciale wensen van de gebruiker rekening kunnen houden.

Bij de gegeven keuze-antwoorden werden het meest frequent de twee mogelijkheden aangekruist die duiden op moeilijke toegankelijkheid van relevante titels: 'veel relevante titels betroffen materiaal dat moeilijk verkrijgbaar bleek' (5) en 'veel titels betroffen artikelen in talen die ik niet kan lezen (Russisch, Japans)' (5). (Dit waren echter niet noodzakelijkerwijs de belangrijkste redenen voor opzegging; men kon net zoveel antwoorden aanstrepen als men wilde. Het totaal aantal verwerkte enquêteformulieren voor deze vraag bedroeg 11.) De moeilijke toegankelijkheid van sommige literatuur kan kennelijk irriterend werken. Moeilijk toegankelijke talen kunnen uiteraard worden uitgesloten[48], zij het dat

48. Bij de kostenformule die IWIS-TNO tot aan 1976 hanteerde was het doorzoeken van een extra informatieveld op de band bijzonder duur. Daarom werd van het doorzoeken van het taal- en het auteursveld destijds ook geen gebruik gemaakt. Nu de oorspronkelijke formule is verlaten en het taalveld met zijn zeer geringe lengte relatief goedkoop kan worden doorzocht, wordt nu wèl op taal geselecteerd.

zulks zoveel mogelijk wordt ontraden: men sluit zich op die manier immers bewust af voor potentieel relevante informatie. Eerder moeten vertaalfaciliteiten worden bevorderd.

Ten slotte werd enkele malen opgemerkt dat een op het eerste gezicht relevante titel uiteindelijk toch voor het onderzoek geen belangrijke informatie bevatte. Het informatiegehalte van de titel was hier dus onvoldoende. Een referaat kan de beoordeling veelal vergemakkelijken, zij het dat alleen de lezing van het origineel uiteindelijk uitsluitsel kan geven over de waarde van het document. Hopelijk zal het CAB/CAIN-vergelijkingsproject enig kwantitatief inzicht geven in het belang van referaten bij de attenderingsselectie. Aangezien het afdrukken van referaten kostenverhogend werkt, zal ook moeten worden nagegaan hoeveel waarde, uitgedrukt in geld, de klant aan een referaat hecht. Klanten die een duidelijke selectie op relevantie van de inhoud wensen voor hun specifieke onderzoek, lijken echter uiteindelijk aangewezen te zijn op manuele attendering, zij het dat het eindoordeel over het belang van een artikel natuurlijk altijd bij de gebruiker zelf ligt.

Een argument van geheel andere aard, dat negatief werkte op het nemen van een attenderingsabonnement, was dat de meeste titels ook langs andere weg werden verkregen. Dit argument duidt deels op een misvatting betreffende de functie van de attendering: deze wil immers niet zozeer alleen maar extra materiaal aandragen, dat niet langs andere weg zou kunnen worden gevonden, als wel huidige vormen van literatuursignalering ten dele vervangen. Als wordt geconstateerd dat men op bestaande manieren ook goed aan zijn literatuur komt, wil dat nog niet zeggen dat die manieren ook efficiënter zijn.

In het geval dat er reeds een attenderingsdienst bestaat op het zeer specifieke gebied van de studie van de betrokken onderzoeker, zal zo'n medium altijd wel goedkoper zijn dan een individueel computer-attenderingsabonnement en is het niet meer dan logisch dat men zich daarop verlaat. (Voorbeelden zijn: 'Intestinal Absorption' en 'Gas Chromatography Abstracts', waarbij de desbetreffende klant voor zijn attenderingsprofiel geïnteresseerd was in alle titels over chromatografie, behalve gaschromatografie!)

De kosten van de andere weg, waarlangs de titels reeds worden verkregen, spelen natuurlijk ook een rol. Voor de individuele onderzoeker is een interne attenderingsdienst of een intern circulatiesysteem meestal een centrale voorziening, die niet op zijn onderzoeksbudget drukt. In dat geval zal hij niet snel bereid zijn in plaats van de bestaande, gratis, dienstverlening uit zijn eigen budget geld vrij te maken voor externe computerattendering.

Ten slotte moet nog eens worden benadrukt dat de computerattendering geen optimale oplossing voor alle wensen met betrekking tot regelmatige literatuursignalering kan en wil zijn. Ten eerste zal het regelmatig lezen van een aantal belangrijke vaktijdschriften altijd noodzakelijk blijven, naast wellicht tijdschriften als 'Nature' en de 'New Scientist' om in algemene zin met ontwikkelingen in de wetenschap op de hoogte te blijven. Daarnaast zal men ook steeds artikelen blijven tegenkomen die de attenderingsdienst om welke reden dan ook mist (zie hiervoor hoofdstuk 5). Voor iedere individuele situatie zal daarom steeds moeten worden nagegaan of computer-attendering een efficiënt hulpmiddel kan zijn. Voorlichting over de mogelijkheden en beperkingen van de attenderingsdienst is hiervoor erg belangrijk; het moet voor de gebruiker duidelijk zijn wat hij van deze dienstverlening mag verwachten en wat niet.

7 Kosten en baten van een operationeel systeem

7.1 INLEIDING

Om een gefundeerde beslissing te kunnen nemen over het al dan niet continueren van een computer-attenderingsdienst ten behoeve van de landbouwwetenschap is een gedetailleerd inzicht in de kosten noodzakelijk. Hiertegenover zal men dan òf de potentiële opbrengsten moeten stellen bij diverse tarieven òf de waarde in geld uitgedrukt die men aan zo'n dienstverlening wil toekennen. In dit hoofdstuk zullen allereerst de kostenelementen worden besproken. Daarna zal nader worden ingegaan op de principes voor het bepalen van de prijs en de nu geldende tarieven.

De historische kosten van het project zijn in het kader van dit rapport van veel minder belang dan een raming van de kosten in de toekomst. De totale kosten van het project werden gedetailleerd vermeld in het aan NOBIN uitgebrachte 'Financieel eindrapport ALADIN-project', gedateerd 18 mei 1976.

De kosten die aan het ALADIN-project zijn gespendeerd - een totaalbedrag van f 497 122,77[49] - moeten voor een deel aangemerkt worden als onderzoekskosten, die niet direct door latere opbrengsten behoeven te worden terugverdiend. Het is bijzonder moeilijk om de zuivere onderzoekskosten te scheiden van de investeringen en de aanloopkosten, die uit zuiver bedrijfseconomisch oogpunt wel afgeschreven zouden moeten worden. Zo zouden de investeringen in de ontwikkeling van de computer-programmatuur wel in aanmerking kunnen komen om terugverdiend te worden. In het totaal hebben deze kosten f 90 132 bedragen. Als deze kosten in de komende vijf jaar zouden moeten worden afgeschreven, zou daarvoor een bedrag van f 25 per verwerkingseenheid per jaar een goede schatting zijn.[50]

De kosten gemaakt voor het gratis verstrekken van attenderingsabonnementen zouden als reclame- en aanloopkosten kunnen worden aangemerkt en zouden bij een zuiver bedrijfseconomische opstelling ook moeten worden afgeschreven. Het is moeilijk aan te geven welk gedeelte van de totale projectkosten hiervoor in aanmerking komt. Maximaal is dit het totaal minus de reeds genoemde kosten voor computer-programmatuur, ofwel f 400 000. Het onderzoeksaspect van het project mag echter toch wel op f 25 000 aan personele kosten worden

---

49. Eindsubsidiebedrag van NOBIN (na aftrek van f 18 995,74 aan inkomsten) f 423 424,62
    Aparte NOBIN-subsidie voor uitbreiding van het FLIRT-programma met Booleaanse logica f 6 666,00
    Voor rekening van Pudoc (Fase I) f 67 032,15
    Totale kosten van het ALADIN-project (exclusief de kosten voor publikatie van het eindrapport) f 497 122,77

50. Uitgaande van een lineaire groei van de huidige 350 verwerkingseenheden tot 1000 verwerkingseenheden in 1981, moet over 3166 verwerkingseenheden worden afgeschreven. Een verwerkingseenheid is daarbij de verwerking van één profiel met één bestand.

gewaardeerd, zodat f 375 000 zou resteren. Afgeschreven over vijf jaar en uitgaande van dezelfde veronderstellingen als in de vorige alinea betekent dit f 115 per verwerkingseenheid per jaar. De kosten van het onderzoeksaspect zijn in feite natuurlijk groter dan f 25 000. Per slot van rekening werd van iedere deelnemer die een 'gratis' abonnement kreeg, een voortdurende evaluatie van de resultaten als tegenprestatie verlangd. Om alle kosten van het project in vijf jaar terug te verdienen zou dus f 140 per verwerkingseenheid per jaar nodig zijn.

7.2 DE HUIDIGE PRODUKTIEKOSTEN

De kosten verbonden aan de computer-attenderingsdienst zoals die momenteel operationeel is, kunnen op verschillende wijzen worden ingedeeld. Een praktische indeling lijkt die naar kosten gemaakt door IWIS-TNO, en door Pudoc direct gemaakte kosten, voornamelijk voor aanschaf van banden en voor personeel. Bij beide bronnen kan dan een onderscheid worden gemaakt naar die kosten die direct verbonden zijn met de verwerking van één band, kosten niet aan één band maar wel aan één profiel gebonden (opstellen van het profiel, kosten van het gebruik van het profiel-modificatieprogramma, administratiekosten) en kosten die niet aan een bepaalde run of een bepaald profiel kunnen worden toegerekend. Vele van deze kostensoorten zijn echter moeilijk apart te schatten. Daarom worden uiteindelijk alleen computerkosten (wèl en nièt bestand-gebonden), abonnementskosten en personeelskosten onderscheiden.

7.2.1 Computer- en abonnementskosten per profiel

De computerkosten voor bandverwerking per profiel staan vermeld in Tabel 12.

Tabel 12. Computer-verwerkingskosten in guldens per profiel per jaar. Voor 1977 en 1978 zijn schattingen opgenomen.

Jaar	Bestandsafhankelijke kosten[1]				Andere kosten[2]
	CAIN	CAB[3]	FSTA	AGRINDEX	
1974	180 (100)		63 (29)		122
1975	173 (92)[4]	64 (35)	62 (24)	103 (99)[4]	90
1976	205 (88)[4]	453 (74)	96 (23)	142 (53)[4]	129
1977	150 (150)	420 (125)	80 (30)	150 (100)	115
1978	125 (200)	375 (175)	60 (50)	150 (150)	105

1. Tussen haakjes het aantal profielen dat gemiddeld in ieder jaar werd verwerkt, resp. wordt verwacht.
2. Wijziging profielenbestand, personeelskosten, papier.
3. In 1975 alleen CABVET (Index Veterinarius en Veterinary Bulletin).
4. Deze cijfers zijn gebaseerd op de eerste twee kwartalen van 1976.
Vanwege de produktie voor het CAIN/AGRINDEX-project werden uiteindelijk veel meer runs verwerkt: voor CAIN 2421 (per jaar gemiddeld 186) en voor AGRINDEX 1909 (per jaar gemiddeld 159). De kosten per profiel per jaar kwamen voor CAIN zo uit op f 132 en voor AGRINDEX op f 100.

Oospronkelijk werden de computerkosten voor het verwerken van één band berekend volgens een formule waarin het aantal titels op de band, het aantal trefwoorden in het zoekprofiel en het aantal doorzochte velden de belangrijkste rol speelden. Voor de drukkosten was een empirische constante opgenomen.[51] De kosten volgens deze formule werden in rekening gebracht, onafhankelijk van de werkelijk gemaakte computerkosten. De formule was empirisch opgesteld; de uitkomsten ervan liepen redelijk parallel met de werkelijke kosten. De eenvoudige voorspelbaarheid van de kosten voor de gebruiker was een groot voordeel van deze regeling.

Bij de verwerking van vier verschillende banden met grote verschillen in hoeveelheid doorzochte informatie per record en grote verschillen in het aantal kaartjes nodig voor het afdrukken van één titel (meestal één voor AGRINDEX, CAIN en FSTA, gemiddeld drie voor CAB, zolang de referaten ook werden afgedrukt), bleek één formule niet langer een goede voorspelling van de werkelijke kosten te kunnen geven. Momenteel worden direct de werkelijke computerkosten volgens de algemene gebruikstarieven van IWIS-TNO in rekening gebracht. In 1977 zal vergeleken worden welke verschillen de formule-uitkomsten te zien geven met de werkelijke kosten.

De grote afwijkingen tussen 1975 en 1976 hebben verschillende oorzaken. Allereerst werd in de oorspronkelijke formule geen rekening gehouden met de extra kosten verbonden aan het afdrukken van extra kopieën. Deze kunnen worden geraamd op gemiddeld *f* 2 per profiel per run. Voor CAIN werd gemiddeld 1,02 kopie per profiel extra afgedrukt in 1976, voor AGRINDEX 1,18. De extra kosten die hier het gevolg van zijn komen dan dus per run op *f* 2,04 resp. *f* 2,35.

De kosten van CAIN zijn daarnaast hoger geworden omdat de banden meer titels bevatten en door de verandering van de technische specificatie van de band, die ongunstig is voor de verwerking door de IWIS-computer. De kosten van AGRINDEX zijn ook gestegen vanwege het grotere aantal titels per band (gemiddeld 4036 in 1975 en 6117 in 1976).

De jaarlijkse abonnementskosten voor de magneetbanden waren in 1975:

CAIN : *f* 1378,62

FSTA : *f* 3094,85

AGRIS: gratis[52]

CAB : *f* 29 032,43 (1976) CABVET alléén: *f* 4148,43 (1975

De verschillen in verwerkingskosten tussen CAIN en AGRINDEX zijn hoofdzakelijk te wijten aan het grotere aantal records op de CAIN-band en duurdere omzetting van het CAIN-format. De verwerking van de CAB-band is vooral duur door de referaten. Ook al worden deze niet doorzocht, ze moeten toch worden omgezet om ze af te drukken. Door een efficiëntere programmering wordt hier echter voor 1977 een belangrijke kostenverlaging verwacht.

De verwerking van de FSTA-band is relatief duur vanwege het geringe aantal profielen, omdat de vaste kosten van het programma de profiel-gebonden kosten overheersen. Bij toename van het aantal profielen wordt de verwerking snel goedkoper.

Aangezien de CAB-band zowel duur is bij de verwerking als in aanschaf, steken de kosten zoals gesommeerd in Tabel 14 zeer ongunstig af bij die van CAIN en AGRINDEX.

De magneetband-abonnementskosten per profiel zijn uiteraard sterk afhankelijk van

---

51. Deze formule luidde, in iets vereenvoudigde vorm: $K = D.(5.10^{-2} + 3.10^{-5}.T + 4.10^{-3}.Z)$ waarin K = kosten in guldens, D = aantal doorzochte documenten, T = aantal zoektermen, en Z = aantal doorzochte velden per document.

52. Ook al zijn er aan de AGRINDEX-band geen directe kosten verbonden, toch staat daar indirect wel de Nederlandse bijdrage aan de AGRIS-invoer tegenover. De personeelskosten die hiermee gemoeid zijn, kunnen op *f* 70 000 worden geraamd.

Tabel 13. Magneetband-abonnementskosten in guldens per profiel per jaar.

	CAIN	CAB	FSTA
Bij huidig aantal profielen	16 (88)	377 (77)	141 (22)
Raming 1977	9 (150	240 (125)	103 (30)
Raming 1978	7 (200)	170 (175)	62 (50)
Alleen CABVET		106 (39)	

Tussen haakjes het aantal profielen dat voor de berekening werd verondersteld.
AGRINDEX is gratis.

Tabel 14. Computer- plus abonnementskosten in guldens per profiel per jaar.

	CAIN	CAB	FSTA	AGRINDEX
1976	350 (88)	959 (74)	366 (23)	271 (53)
1977	274 (150)	775 (125)	296 (30)	265 (100)
1978	237 (200)	650 (175)	227 (50)	255 (150)

Tussen haakjes het aantal profielen dat voor de berekening werd verondersteld.
AGRINDEX is gratis.

het aantal profielen. Deze kosten worden in Tabel 13 vermeld.

Als de gegeven kosten worden opgeteld, ontstaat het beeld van de computer- plus magneetband-abonnementskosten per profiel per jaar (Tabel 14).. Het is uitermate moeilijk om onafhankelijk van een indicatie van de vraag-differentiatie een uitspraak te doen over de meerwaarde van het CAB-resultaat ten opzichte van CAIN en AGRINDEX. (Is een abonnement op CAB inderdaad (ƒ 775 - ƒ 274 =) ƒ 501 meer waard dan een abonnement op CAIN?) De meerwaarde zou gelegen moeten zijn in het referaat, in een wellicht hogere precisie en relevantie, en in de betere selectiviteit en dekkingsgraad van het bestand.[53]

7.2.2 Personele kosten

De schatting van de personele kosten is een hachelijke zaak. Als ervaringsregel wordt wel gehanteerd dat één academicus vier profielen per dag kan formuleren. Bij een uurtarief van ƒ 51 en een werkdag van 8 uur zou dit ƒ 102 per profielopstelling zijn. Aannemende dat een profiel drie jaar gehandhaafd blijft, zouden deze kosten ƒ 34 per jaar bedragen.

De verzending van 80 profieluitkomsten van één band vergt momenteel vier uur administratieve arbeid. Bij een uurtarief van ƒ 17 is dat ƒ 68 per run of ƒ 10,20 per profiel (verwerkingseenheid)[54] per jaar.

53. Als de referaten helemaal niet worden gebruikt, zou dit een aanzienlijke besparing opleveren. De abonnementskosten zijn dan ook lager. Pas als de uitkomsten van het CAB/CAIN-vergelijkingsproject bekend zijn, zal hier nader over worden beslist.
54. Het verschil tussen een profielopstelling en een profiel als verwerkingseenheid is het volgende: één profielopstelling kan gelden voor ieder bestand waarop het profiel wordt verwerkt. Een verwerkingseenheid geldt echter voor één bestand. Momenteel wordt gemiddeld één profielopstelling verwerkt op 2,11 bestanden.

Dit wat betreft de duidelijk bekende en toerekenbare kosten. Behalve deze eist echter het beheer van de dienst nog een aanzienlijke personele inspanning. Dit betreft het contact met IWIS-TNO, overleg met de klanten, algemene voorlichting, bijhouden van administratieve systemen, voeren van de financiële administratie, nationaal en internationaal overleg alsmede oriëntatie op nieuwe ontwikkelingen. Momenteel wordt de computer-attenderingsdienst bemand door een academicus voor hele dagen en een administratieve kracht voor halve dagen. De financiële administratie en de verzending worden door de Afdeling Personeel en Algemene Zaken van Pudoc verzorgd.

Volgens de normbedragen voor de projectadministratie 1976 moet voor hoger personeel ƒ 86 358 per jaar worden berekend en voor 'lager personeel' ƒ 29 304 per jaar of ƒ 14 652 per half jaar. De totale personeelskosten voor de computer-attendering zouden dan ƒ 101 010 bedragen. Dit bedrag zou kunnen worden verdeeld over het aantal profielen dat wordt beheerd. Bij 100 profielen bedragen de personele kosten dan ongeveer ƒ 1000 per profiel per jaar, lineair afnemend tot ƒ 200 bij 500 profielen. Dit laatste aantal kan het maximum geacht worden dat één academicus plus een halve administratieve kracht kan beheren. Daarna is een vergroting van de personele bezetting nodig. De noodzakelijke personele bezetting is echter meer afhankelijk van de *groei* van het aantal profielen dan van het aantal lopende abonnementen. Het formuleren van *nieuwe* profielen is immers het meest tijdrovende onderdeel van het werk. Bij snelle groei is dus relatief veel personeel nodig, terwijl bij stabilisatie van het aantal profielen theoretisch een reductie van de personeelsbezetting mogelijk zou moeten zijn.

7.2.3 Totale kosten computer-attendering

Om althans een indruk te krijgen van de totale kosten van een computer-attenderingsdienst zullen de cijfers uit de vorige paragrafen nog worden gecombineerd. Ten slotte is dan nog een schatting van de verwachte vraag nodig om tot een raming van de totale financiële implicatie van deze vorm van informatieverschaffing te komen.

De cijfers uit 7.2 kunnen alleen gesommeerd worden als bepaalde veronderstellingen worden gedaan betreffende het aantal banden dat per profiel wordt verwerkt en welke banden dat zullen zijn. (Hoe meer CAB hoe duurder!) Momenteel worden gemiddeld 2,11 bestanden per profiel verwerkt. De verhouding is daarbij CAIN : CAB : FSTA : AGRINDEX = 10 : 8,75 : 2,5 : 6,0. Voor de verdere berekeningen is een verhouding van 10 : 8,75 : 2,5 : 10 aangenomen. Dit weerspiegelt de te verwachten gelijkwaardigheid van CAIN en AGRINDEX beter.

In Tabel 15 zijn de totale kosten voor 100 tot 500 profielen met deze bestandssamenstelling weergegeven.[55] Voor 1977 wordt het totaal aantal profielopstellingen op 175 ge-

55. Deze bedragen zijn nog exclusief porto- en eventuele andere bijkomende kosten, voor zover die niet begrepen zijn in de reeds in de personeelskosten verdisconteerde overhead. De portokosten kunnen worden geraamd op ƒ 50 per jaar per profiel. Ook is geen rekening gehouden met attenderingsabonnementen op andere magneetbanden dan de vier die nu in eigen beheer worden verwerkt. Voor 1977 worden de uitgaven aan abonnementen elders, met name bij NOCI voor Chemical Abstracts (33 profielen) en de Biological Information Service (GB) voor Biological Abstracts (13 profielen) op ƒ 50 000 geraamd. Bij gelijkblijvende verhoudingen zou bij 500 'eigen' profielen het bedrag van ƒ 389 000 verhoogd moeten worden met ƒ 72 000 voor elders verwerkte profielen tot ƒ 461 000.

Tabel 15. Totale kosten computer-attendering in duizenden guldens per jaar (tussen haakjes de kosten per profiel).

Aantal profielen	Personeel	Abonnement	Computer	Totaal
100	101(1,01)	35(0,35)	91(0,91)	227(2,27)
200	101(0,51)	35(0,18)	145(0,72)	281(1,40)
300	101(0,33)	35(0,12)	187(0,62)	323(1,08)
400	101(0,25)	35(0,09)	219(0,55)	355(0,89
500	101(0,20)	35(0,07)	253(0,51)	389(0,78)

De computerkosten zijn samengesteld uit een gewogen gemiddelde van de verwerkingskosten per bestand plus de niet van het bestand afhankelijke computerkosten, vermenigvuldigd met de factor 2,11 voor het aantal verwerkte bestanden per profiel. Van alle computerkosten is aangenomen dat zij geleidelijk per eenheid afnemen naarmate het aantal profielen stijgt.

raamd. De totale kosten van de computer-attenderingsdienst zouden dan dus op *f* 270 000 mogen worden geraamd (*f* 1543 per profiel).

Of de genoemde bedragen vergelijking met in het buitenland aangeboden diensten kunnen doorstaan, is moeilijk na te gaan omdat er momenteel noch op CAB noch op AGRINDEX elders een attenderingsdienst voor Nederlandse gebruikers beschikbaar is. Binnenkort is de verwerking van CAIN-, CAB- en AGRINDEX-profielen wellicht in Frankrijk mogelijk. De taalbarrière zal echter de noodzakelijke intensieve contacten met het verwerkingscentrum daar bemoeilijken. CAIN zou wellicht in Zweden kunnen worden verwerkt. De huidige prijs bedraagt daar *f* 146 per profiel, hetgeen aanzienlijk lager is dan onze kosten en ook lager dan onze abonnementsprijs. Het is echter nog maar de vraag of men in Zweden bereid zou zijn de grote subsidie die op deze dienstverlening wordt gegeven ook aan Nederlandse gebruikers te geven, indien wij een zeer groot aantal profielen zouden aanbieden. FSTA zou in Denemarken kunnen worden verwerkt. Voor Deense gebruikers bedraagt de abonnementsprijs daar momenteel *f* 214, maar er werd reeds bij vermeld dat voor buitenlanders een andere, nog niet vastgestelde prijs zou moeten gelden. (*f* 214 is overigens ook lager dan onze kosten en ook lager dan onze abonnementsprijs.) Over de ervaringen met buitenlandse verwerkingsinstellingen werd in 4.3 bericht. In het algemeen zijn de lange communicatiekanalen bezwaarlijk, zeker in de aanloopfase van een nieuw profiel, wanneer snel moet worden bijgestuurd.

Ten slotte doet zich hier de vraag voor met welke aantallen profielen in feite moet worden gerekend. Over de bovengrens van het aantal abonnementen zijn althans enkele cijfers bekend. Volgens het PA consultants rapport[56] zouden er in 1980 in Nederland 1640 landbouwkundige gebruikers zijn van informatiediensten die interactief computerbestanden doorzoeken. Potentieel kunnen zij ook als attenderingsklanten worden beschouwd, misschien met enige reductie. Een ander uitgangspunt is het aantal mensen dat momenteel bij het Ministerie van Landbouw en Visserij werkt in een wetenschappelijke rang (inclusief LH).

56. PA International Management Consultants Ltd., 1974. Forecast of users of on-line retrieval services for scientific and technical information in Europe 1976-1985. Prepared for the Commission of the European Communities.

Dit zijn er ongeveer 1050. Vermoedelijk doen hiervan zo'n 800 direct wetenschappelijk onderzoek. Inclusief de potentiële gebruikers buiten het ministerie lijkt dan een totaal van 1250 potentiële klanten realistischer. Er zal echter nog heel wat tijd overheen gaan voordat het overgrote deel van alle *potentiële* gebruikers *werkelijke* gebruikers zal zijn geworden, aangezien het uit handen geven van een stuk werk dat traditioneel door de onderzoeker zelf werd verricht, een ingreep in een bestaand patroon betekent.

Het is ook nog de vraag of iedere wetenschappelijke onderzoeker een apart attenderingsabonnement nodig heeft. In steeds toenemende mate vindt onderzoek in team-verband plaats, waarbij vaak één literatuur-interesseprofiel kan worden gebruikt voor het hele team. Voor een aantal onderwerpen kan bij verschillende onderzoekers dezelfde belangstelling bestaan ook al werken ze los van elkaar. Dit laatste is bijvoorbeeld het geval met het profiel 'weefselcultuur', waarvan de attenderingskaartjes momenteel aan acht verschillende onderzoekers worden toegestuurd. In het totaal worden van vijf profielen de resultaten aan meer dan één persoon toegestuurd. Bij een grote nadruk op groepsprofielen loopt men natuurlijk wel het gevaar dat het resultaat te algemeen - en te uitgebreid - wordt en de individuele onderzoekers er zelf nog veel in moeten selecteren.

7.3 KOSTEN MANUELE ATTENDERING

De personeelskosten voor de manuele attendering kunnen benaderd worden met de volgende formules:
kosten voor wetenschappelijk personeel: $(27 + 0{,}4\ n) \times 51 \times 52$
kosten voor administratief personeel : $n \times 17 \times 52$
waarin n het aantal profielen is.

Hierbij is uitgegaan van de waarneming dat het wetenschappelijk personeel voor 30 profielen 39 uur per week nodig heeft, waarvan tweederde vast is, d.w.z. besteed moet worden onafhankelijk van het aantal profielen. Voor het resterende gedeelte is een lineaire afhankelijkheid van het aantal profielen verondersteld. De administratieve verwerking vereist per profiel één uur per week. De factoren '51' en '17' geven de geldende uurtarieven weer. '52' is het aantal weken in een jaar.

Voor 30 profielen resulteert dan een afgerond totaalbedrag per jaar van ƒ 130 000, d.i. ƒ 4333 per profiel per jaar, voor 40 profielen ƒ 150 000, d.i. ƒ 3750 per profiel per jaar, en voor 50 profielen ƒ 170 000, d.i. ƒ 3400 per profiel per jaar.

Het uitgangsmateriaal wordt momenteel gratis door de Bibliotheek van de Landbouwhogeschool aan Pudoc voor de attendering ter beschikking gesteld. Voor de huidige situatie zijn de kosten hiervan dus niet relevant. Het zou overigens bijzonder moeilijk zijn om voor de totale kosten van nieuwe aanschaffingen en lopende abonnementen van de Bibliotheek een aparte berekening voor attenderings- en bibliotheekgebruik te maken.

Uit een vergelijking van de kosten voor manuele en computer-attendering blijkt duidelijk dat ook reeds onder de voor 1977 verwachte omstandigheden de computer-attendering goedkoper is dan manuele attendering: de totale kosten voor een gemiddeld abonnement met het computersysteem kunnen bij 175 profielen op ƒ 1550 geraamd worden, tegenover minimaal ƒ 3400 bij de manuele attendering. Voor ƒ 1550 ontvangt men dan gemiddeld 127 relevante titels per jaar, ofwel ƒ 12 per titel, en bij manuele attendering 56 titels per jaar, of

ƒ 61 per titel.[57] Manuele attendering zou dan vijf maal zo duur zijn per gevonden relevante titel als computer-attendering. De absolute waarde van deze cijfers is natuurlijk maar zeer betrekkelijk, maar een indruk van de orde van grootte geven zij wel.

7.4 ALGEMENE OVERWEGINGEN BIJ FINANCIERING EN PRIJSSTELLING

Over de financiering van wetenschappelijke informatievoorziening wordt momenteel in verschillende kringen intensief gediscussieerd. De twee uitersten waartussen de mogelijke oplossingen moeten worden gezocht zijn daarbij: a) directe financiering uit de openbare middelen als een openbare voorziening en b) betaling van alle kosten door de gebruiker. Het eerste principe sluit vooral aan bij de bibliotheektraditie, het laatste is noodzakelijk als de informatie op commerciële basis wordt verschaft. Het wordt tevens verdedigd vanuit een oogpunt van doorzichtigheid van de totale kostenstructuur van onderzoeksprojecten bij de overheid: alle kosten die direct aan een bepaald project toegerekend kunnen worden, moeten ook door dat project worden gedragen.

In Nederland vindt op dit gebied momenteel vruchtbaar overleg plaats in de 'Werkgroep Tarificatie' van NOBIN. Binnen deze werkgroep zijn zowel instellingen vertegenwoordigd die door overheidssubsidie hun diensten geheel gratis aan het publiek kunnen aanbieden als ook instellingen die al hun kosten moeten dekken. De werkgroep heeft voorlopig als principe aanvaard dat de gebruiker de door hem veroorzaakte extra kosten ook zelf moet dragen. In principe zou de overheid de vaste kosten voor haar rekening moeten nemen. Dit laatste wordt vooral verdedigd vanuit het algemene belang van wetenschappelijke informatievoorziening. Daarnaast speelt het praktische argument een rol dat instellingen die nu eenmaal geld tot hun beschikking hebben om diensten gratis of althans niet-kostendekkend aan te bieden, moeilijk te bewegen zijn een hoog, kostendekkend tarief te vragen. Hiermee wordt het opbouwen van een klantenkring immers ernstig bemoeilijkt, zeker in een aanvangsstadium. Een belangrijke rol speelt hierbij ook dat bij de projectfinanciering in de overheidssfeer het doorberekenen van *alle* projectkosten door overheidsdiensten aan elkaar in de praktijk nog steeds niet wordt toegepast. De huidige scheiding tussen personele en materiële budgetten bemoeilijkt de doorzichtigheid van de zuivere projectkosten zeer, terwijl het ook moeilijk is om een gedeelte van het personele budget een herbestemming te geven voor het kopen van diensten van derden, die uit het materiële budget zouden moeten worden betaald. Dit is vooral noodzakelijk wanneer er bij de materiële onderzoeksbudgetten geen rekening wordt gehouden met de kosten voor informatiediensten van derden, hetgeen veelal het geval is.

Een ander argument voor een relatief lage prijs voor deze informatiediensten is de

---

57. De aantallen titels zijn ontleend aan Bijlage 4 en 5, waarbij aangenomen is dat de helft van het aantal misschien relevante titels inderdaad relevant is. De berekening is dan als volgt:
manueel : (939 + 270) / 25 = 56
computer: (470 + 2113 + {(3274 - 2113) / 2)} / 25 = 127.
Bij deze laatste berekening zijn echter de extra titels die CAB en AGRINDEX nog zouden kunnen opleveren, niet meegerekend, terwijl anderzijds de kosten en titelopbrengst van Chemical Abstracts, Biological Abstracts en Medlars niet in de berekening zijn betrokken.

onbekendheid bij de potentiële gebruiker met de beschikbaarheid van externe literatuurattendering, en bovendien de algemene onderwaardering van literatuurinformatie als zodanig bij wetenschappelijke onderzoekers. Toegegeven moet worden dat het zeer moeilijk is aan externe (literatuur)informatie een geldswaarde toe te kennen, maar dit neemt niet weg dat momenteel veel onderzoekers nog weinig informatie-bewust lijken (zie verder 8.3.1 en 8.4.1).

Om bij de huidige lage subjectieve waardering van informatie en het gebrek aan budgettaire voorzieningen voor externe diensten op dit gebied, deze diensten toch ingang te doen vinden, lijkt een aantrekkelijke introductieprijs noodzakelijk. Als de veronderstelling juist is dat door het gebruik van deze diensten de waardering voor informatie zal stijgen, zou uiteindelijk de bereidheid moeten groeien deze waardering ook in geld uit te drukken. Hiervoor is dan echter ook op het niveau waar de budgettaire voorzieningen worden getroffen, een verandering van instelling nodig ten aanzien van de waarde van externe informatiediensten.

7.5 DE TARIEVEN VOOR LITERATUURATTENDERING BIJ PUDOC

De vaststelling van de prijzen waartegen Pudoc literatuurattenderings-abonnementen met behulp van een computersysteem zou gaan aanbieden, is betrekkelijk pragmatisch tot stand gekomen. Het uitgangspunt werd daarbij gevormd door het tarief voor de reeds bestaande manuele attenderingsdienst. In 1975 bedroeg dit *f* 250 per jaar plus *f* 1,25 per geleverde titel. Gemiddeld werd per jaar *f* 500 betaald. Het werd redelijk geacht voor de attendering m.b.v. het computersysteem, althans in het begin, eenzelfde bedrag te vragen. De klant zou op die manier tenminste alleen op grond van de kwaliteit van het produkt kunnen kiezen, zonder dat het prijsverschil daarbij een rol behoefde te spelen.

De vraag òf, en zo ja in hoeverre, de kosten*structuur* en het *volume* van het geleverde produkt in de prijs tot uitdrukking moesten worden gebracht, werd echter uiteindelijk voor beide diensten verschillend opgelost. Voor de manuele attendering werd gekozen voor handhaving van het systeem van een vast bedrag per jaar plus een bedrag per geleverde titel. De argumentatie hiervoor was vooral gelegen in de gebleken bereidheid van de klant om voor méér titels ook méér te betalen. Dat de klant zo vooraf moeilijk kan weten hoeveel zijn abonnement voor het komende jaar gaat kosten wordt dan op de koop toegenomen. Aangezien een belangrijk deel van de totale kosten wordt bepaald door het aantal door te nemen documenten en niet door het aantal gevonden titels, werd wel besloten het vaste jaarlijkse abonnementsbedrag in de komende jaren sneller te laten stijgen dan het bedrag per gevonden titel. Voor 1977 is het tarief dan ook op *f* 350 per jaar plus *f* 1,25 per titel vastgesteld.

Bij de attendering m.b.v. het computersysteem hangt de structuur van de variabele kosten minder samen met het *volume* van het eindprodukt dan bij de manuele attendering. Het aantal zoektermen in het zoekprofiel speelt hier een belangrijke rol bij de kosten, terwijl het aantal af te drukken titels in de kostenformule nauwelijks van belang is. Het werd niet redelijk geacht om de variabele kosten van de profielopstelling aan de klant door te berekenen: meestal kan het deze immers niet worden 'aangerekend' dat het voor zijn onderwerp nodig is een groot aantal zoektermen te gebruiken. Dit is veelal ook erg afhan-

kelijk van het gebruikte bestand.[58] Op basis van deze argumenten werd in 1975 besloten één vast tarief te vragen, onafhankelijk van de vraag-afhankelijke kosten. Voor de klant heeft dit tevens het voordeel dat hij bij het aangaan van het abonnement weet hoeveel het hem voor dat jaar gaat kosten. Wel behoudt Pudoc daarbij zelf het laatste woord over de profielopstelling: de klant moet geen misbruik kunnen maken van het feit dat extra zoektermen hem niet in rekening worden gebracht. De informatiespecialist die het profiel formuleert bepaalt in laatste instantie of het zin heeft bepaalde marginaal geachte termen nog op te nemen of niet.

Het tarief werd voor CAIN op f 500 vastgesteld, gelijk aan het gemiddelde voor de manuele attendering, en voor FSTA op f 350, gezien het veel kleiner bestand. Op deze basis werd overeenstemming bereikt met het Ministerie van Landbouw en Visserij, in die zin dat het verschil tussen de materiële kosten en de opbrengsten van de attenderingsdienst direct aan Pudoc zouden worden gesubsidieerd. Hierbij werd nog overwogen dat ten minste de materiële kosten moesten worden terugverdiend. Voor CAIN alléén kunnen deze bij 100 profielen op f 325 per profiel worden geschat. Officieel moeten voor diensten aan derden ook de personele kosten in rekening worden gebracht, hetgeen dan nog eens f 1010 extra per profiel zou betekenen. Een totale prijs van f 1335 moet echter irreëel geacht worden, ook tegen de achtergrond van de voor manuele attendering gehanteerde prijs. Deze laatste kan ook niet zonder meer drastisch worden verhoogd, ook al houdt hij geen enkel verband met de werkelijke totale kosten. Daarnaast wilde men speciaal binnen het Ministerie van Landbouw en Visserij de drempel voor het gebruik van deze nieuwe dienstverlening graag laag houden; daartoe werd het niet noodzakelijk geacht de personele kosten gemaakt binnen de ene dienst van het ministerie door te berekenen aan andere diensten van datzelfde ministerie. Daarom werd een regeling ontworpen waarbij alle gebruikers die onder het Ministerie van Landbouw en Visserij ressorteren, 50% reductie op de genoemde tarieven zouden krijgen.

Toen er in 1976 drie algemene landbouwkundige bestanden ter beschikking waren, deed zich het probleem voor in hoeverre er voor ieder bestand apart zou moeten worden betaald.[59] Indien aan iedereen alle drie bestanden zouden worden aangeboden, zou dit tot aanzienlijke prijsverhoging per abonnement moeten leiden, als de directe subsidie tenminste niet aanzienlijk zou worden verhoogd.

Aangezien de klant door een abonnement op een extra bestand toch ook meer resultaat kon verwachten werd het redelijk geacht de extra kosten die het verwerken van een extra bestand met zich meebrengt, ook in de tariefstructuur tot uitdrukking te brengen. Een probleem deed zich daarbij voor ten aanzien van de bepaling welk gedeelte van de prijs niet afhankelijk van het bestand zou moeten zijn en welk gedeelte extra per bestand in rekening zou moeten worden gebracht. In 1976 werd als basisbedrag f 150 genomen, waar voor CAIN en CAB dan f 350 bij kwam, voor AGRINDEX f 175 en voor FSTA f 200. Voor 1977 werd deze verhouding echter in belangrijke mate gewijzigd: het basisbedrag werd f 300 en het extra bedrag voor CAIN f 200, voor AGRINDEX f 175, voor CAB f 350 en voor FSTA f 100.

58. Als voorbeeld moge nogmaals het profiel 'eetbare paddestoelen' dienen, waarvoor in CAIN 43 termen nodig waren en in FSTA slechts één (zie 3.3).
59. Voor alle bestanden die niet in eigen beheer werden verwerkt, werd in 1975 reeds een vast bedrag, overeenkomend met de *gemiddelde* kosten, aan de klant doorberekend.

Bij deze opstelling wordt beter tot uitdrukking gebracht dat ook de personele inspanning nauwelijks afhankelijk is van het bestand. Ook wordt nu met de hogere kosten voor CAB rekening gehouden. (Het formulier met de tarieven voor literatuuronderzoek, zoals die nu bij Pudoc gelden, is toegevoegd als Bijlage 12.)

Bij de in 7.2.3 gehanteerde verhouding in het gebruik van de bestanden (CAIN : CAB : FSTA : AGRINDEX = 10 : 8,75 : 2,5 : 10) en een gemiddeld aantal bestanden per profiel van 2,11 bedraagt de gemiddelde prijs per abonnement ƒ 850. Deze prijs zou volgens de gegevens van Tabel 15 bij ongeveer 450 abonnementen geheel kostendekkend moeten zijn. De abonnements- en computerkosten worden reeds bij 300 abonnementen gedekt.

Deze tariefstructuur is nog niet geheel in overeenstemming met de aanbeveling van de Tarificatie-werkgroep van NOBIN. Hierin werd, althans voor retrospectief onderzoek, geconcludeerd dat de gevraagde prijs zo min mogelijk van het aantal gebruikte bestanden en de individuele kosten daarvan afhankelijk zou moeten zijn. Het argument daarvoor is vooral dat de waarde van het geleverde eindprodukt moeilijk is te relateren aan de aan het bestand gebonden kosten. Bij een grote keuze aan bestanden en zeer variabele verwerkingskosten zullen kosten-gebonden prijsformules voor de gebruiker zeer ondoorzichtig worden en een voorspelling over het door de klant te betalen eindbedrag zeer moeilijk maken. Voor de vier door Pudoc zelf verwerkte bestanden lijkt de nu gehanteerde prijsstructuur nog vrij doorzichtig: de klant weet tevoren wat zijn jaarabonnement gaat kosten en kan er in het algemeen op rekenen dat een extra bestand hem ook méér relevante titels zal opleveren. (De kosten per extra nieuwe titel nemen echter bij een abonnement op een extra bestand wel toe.) Voor de extra informatie die met CAB wordt geleverd in de vorm van referaten moet ook extra worden betaald, hetgeen redelijk wordt geacht.

Als er voor iedere attendering eenzelfde eenheidstarief zou gelden, zou dit ook moeten inhouden dat daarvoor een optimaal produkt wordt geleverd, *waarbij Pudoc zou moeten bepalen wanneer het produkt optimaal kan worden genoemd*, mede gegeven de marginale kostenstructuur. Is dit voor een retrospectieve literatuurvraag al moeilijk, voor een attenderingsabonnement lijkt dit welhaast onmogelijk, omdat er te zeer een wissel op de toekomst moet worden getrokken: doordat het resultaat in de tijd verdeeld ter beschikking komt, kan men wel aan het einde van een periode het resultaat evalueren, maar het kan niet meer overgedaan worden. De combinatie van bestanden (en de profielopstelling) kunnen alleen door een volgende periode worden veranderd. Aan een éénmalige retrospectieve vraag kan net zo lang gewerkt worden totdat het best haalbare resultaat is bereikt.

Als laatste principe bij de prijsstelling moet nog de aanbeveling van de Tarificatiewerkgroep worden genoemd om voor gelijke diensten gelijke prijzen te vragen. Bij de huidige, nog zo geringe markt wordt prijsconcurrentie schadelijk geacht voor de nationale informatievoorziening. In Nederland worden echter geen bestanden door meer dan één instantie m.b.v. een computer voor attendering verwerkt.

Internationaal gezien liggen de tarieven van Pudoc momenteel vrij hoog. Op dit moment speelt internationale concurrentie echter nog maar een geringe rol.[60] Door internationaal

60. Wegens de lagere prijs werd echter in 1976 de verwerking van de Biosis-profielen voor Pudoc van BMDC in Stockholm verhuisd naar BIS in Engeland.

aantrekkelijke tarieven te vragen zou Pudoc echter wellicht zijn volume kunnen uitbreiden en daarmee de kosten per eenheid drukken, hetgeen ook voor de Nederlandse gebruikers gunstig zou kunnen zijn.

8 Evaluatie

In dit slothoofdstuk zal worden aangegeven tot welke conclusies en aanbevelingen de ervaringen opgedaan met het ALADIN-project leiden.

8.1 DE TECHNISCHE UITVOERBAARHEID

In technische zin kan het project een succes worden genoemd: het is mogelijk gebleken in Nederland een literatuurattenderingsdienst, gebaseerd op computerverwerking van magneetbanden met documentaire informatie, te realiseren. Door de inspanningen van de staf van IWIS-TNO, in samenwerking met de ALADIN-projectleiding en met name ook NOCI, werd een computerprogramma ontwikkeld dat de gewenste functies uitvoert. Tevens heeft IWIS-TNO getoond een regelmatige en betrouwbare dienstverlening te kunnen waarborgen. Er is aldus een operationeel systeem tot stand gekomen voor literatuurattendering ten behoeve van de landbouwwetenschap met behulp van een computersysteem dat technisch geheel naar bevrediging werkt.

Er werd niet specifiek onderzocht in hoeverre het computerprogramma in staat is alle titels die op een bepaalde magneetband voorkomen en die voor een bepaald onderwerp relevant zijn, ook inderdaad op te sporen. De algemene ervaring met het programma duidt er echter op dat een dergelijk onderzoek in de meeste gevallen tot een zeer positief resultaat zal leiden. Dit is ook de ervaring van zulke analyses in het buitenland met dergelijke computerprogramma's.

De belangrijkste beperkingen van literatuurattendering met behulp van een computersysteem en magneetbanden van secundaire literatuurinformatie zijn gelegen in de samenstelling van het gebruikte uitgangsmateriaal (niet alle bestaande relevante titels bevinden zich in de gebruikte titelverzameling) en het voor het computer-selectieprogramma te geringe informatiegehalte van sommige relevante titels, zodat deze, ondanks hun aanwezigheid, niet door het programma worden gevonden.

8.2 DE KEUZE TUSSEN EEN COMPUTERSYSTEEM VOOR LITERATUURATTENDERING EN MANUELE ATTENDERING

Dit probleem werd uitgebreid in Hoofdstuk 5 behandeld. Het ging daarbij om een praktische vergelijking tussen twee vormen van literatuurattendering die door Pudoc worden aangeboden: met behulp van een computersysteem, gebruikmakend van secundaire bronnen en het manueel selecteren uit primaire documenten, met name uit de collectie van de Bibliotheek van de Landbouwhogeschool. De conclusie was dat de computer-attendering in het algemeen méér relevante titels vond dan de manuele attendering, al kan de computer-attendering de manuele attendering nergens geheel vervangen. Als het uitgangsmateriaal hetzelfde zou zijn en de kosten geen rol zouden spelen, dan zou de mens altijd nog beter kunnen

selecteren dan de computer.

In vele gevallen is het uitgangsmateriaal voor de computer-attendering echter breder dan dat voor Pudoc's manuele attendering. Daarnaast is de computer-attendering duidelijk goedkoper en is het volume ervan veel gemakkelijker uit te breiden. Het lijkt dan ook aan te bevelen de manuele attendering vooral te gebruiken voor die speciale gevallen waar manuele attendering een beter resultaat zal geven dan computer-attendering, of waarvoor althans de aanvulling door de manuele attendering van groot belang geacht kan worden. Het totale aantal van deze gevallen is meer dan groot genoeg om het voortbestaan van de manuele attenderingsdienst bij Pudoc te rechtvaardigen.

8.3 DE BEHOEFTE AAN LITERATUURATTENDERING BIJ DE LANDBOUWWETENSCHAP

De behoefte aan literatuurattendering kan worden besproken onafhankelijk van de wijze waarop in de attendering wordt voorzien. Men stuit hierbij op de moeilijkheid dat het in feite niet mogelijk is de behoefte te onderzoeken aan een produkt of dienst waarmee de potentiële consument nog onbekend is. Men zal mensen die nog nauwelijks van het bestaan van waterleiding en stromend water in huis op de hoogte zijn, ook niet vragen naar de behoefte aan een ligbad of een keukengeyser. Een directe vraag naar de behoefte aan literatuurattendering onder een groep die deze dienst niet al uit eigen ervaring kent, lijkt dan ook weinig zinvol. De subjectieve behoefte kan hoogstens worden benaderd vanuit het meer objectieve, verwachte nut van een dergelijke dienstverlening. Al is dit moeilijk kwantitatief uit te drukken, toch is wel een enkele aanduiding te geven.

8.3.1 Het nut van een literatuurattenderingsdienst voor landbouwwetenschappelijk onderzoek[61]

Literatuurattendering kan bijdragen tot de efficiëntie van het wetenschappelijk onderzoek door:

a) zo snel mogelijk relevante onderzoeksresultaten van elders in het onderzoek hier te incorporeren;

b) zo snel mogelijk gebruik te maken van nieuwe methoden en technieken;

c) te verhinderen dat hier onderzoek wordt begonnen dat elders reeds is of wordt verricht.

De waarde van deze bijdragen is nauwelijks in geld uit te drukken. Door informatie over elders reeds verricht onderzoek kan soms jaren eigen onderzoek worden uitgespaard, of kan blijken dat men zijn eigen onderzoeksactiviteiten beter anders kan richten dan oorspronkelijk de bedoeling was.

Tot aan de recente wijzigingen was het bij de rapportage voor de administratie van landbouwkundige onderzoeksprojecten niet noodzakelijk aan te geven in hoeverre men had onderzocht of het voorgenomen onderzoek niet reeds elders was verricht. Uit een steek-

---

61. Veel van de hier gebruikte argumenten voor externe literatuurattendering gelden mutatis mutandis voor retrospectief literatuuronderzoek. Als er dan ook over 'literatuuronderzoek' wordt gesproken, wordt in feite zowel het eenmalige retrospectieve literatuuronderzoek bedoeld als ook de attendering. Attendering is in veel gevallen een logische voortzetting van een retrospectief literatuuronderzoek.

proef uitgevoerd door COLADIN bleek dat er in de meeste gevallen althans geen formeel literatuuronderzoek aan het entameren van een onderzoeksproject vooraf was gegaan. Het lijkt logisch dat aan de financier van een onderzoeksproject enige garanties worden geboden dat het voorgestelde onderzoek niet reeds elders is of wordt verricht. Literatuuronderzoek is daartoe een goed middel. De kosten ervan bedragen slechts een fractie van wat het totale onderzoek uiteindelijk gaat kosten. Als het voorgestelde onderzoek niet nodig blijkt te zijn, heeft men weliswaar een paar honderd gulden uitgegeven maar daardoor vele duizenden guldens uitgespaard.

Het op de hoogte raken en blijven van de relevante literatuur kan de onderzoeker natuurlijk geheel in eigen hand houden. Door althans een belangrijk gedeelte ervan uit te besteden wint hij allereerst tijd voor zijn eigenlijke onderzoek, waaraan hij zijn tijd zinvoller kan besteden omdat hij daarvoor is opgeleid. Instellingen gespecialiseerd in literatuuronderzoek kunnen op hun gebied meestal efficiënter werken dan de nauwelijks daarin getrainde onderzoeker. Deze instellingen kunnen per tijdseenheid zowel méér 'produkt' als een 'beter' produkt leveren: zij zullen veelal meer relevant materiaal vinden dan de onderzoeker zelf. Zij zijn immers beter op de hoogte van de diverse bronnen en kunnen gemakkelijker gebruik maken van moderne hulpmiddelen zoals de computer. Extern literatuuronderzoek is daarom een kwestie van efficiëntie door specialisatie.

Het aantal bij computer-attendering gevonden zeer relevante *nieuwe* titels geeft ook al aan dat de externe attendering belangrijk materiaal lijkt te leveren, dat de onderzoeker (nog) niet zelf had gevonden. Het aantal zeer relevante nieuwe titels per jaar per profiel kan op grond van de resultaten van dit onderzoek worden geschat op ongeveer 20 titels.

Het literatuuronderzoek kan de onderzoeker echter nooit geheel uit handen worden genomen. Hij zal altijd zelf een aantal tijdschriften moeten blijven lezen om in het algemeen op de hoogte te blijven. Ook zal hij de voor hem geselecteerde literatuur zelf moeten lezen. Omdat er bij externe attendering vaak méér relevante titels boven tafel komen dan de onderzoeker zelf zou hebben gevonden, kan de uiteindelijke tijdsbesparing voor de onderzoeker gering zijn: wat hij wint bij het opsporen raakt hij weer kwijt omdat hij meer moet lezen.[62]

Ten slotte moet nog worden gewezen op het grote belang van intensief contact tussen de onderzoeker en de literatuuronderzoeker: in een dialoog tussen deze twee moet een zo goed mogelijk beeld ontstaan van de informatiebehoeften van de onderzoeker, die dan moeten worden vertaald in de mogelijkheden van het literatuuronderzoek. (Zie ook het slot van 6.3.)

8.3.2 Gebruikerservaringen met de attenderingsdienst

De tweede indirecte benadering van de behoefte aan literatuurattendering is de uit-

62. Het lezen van de originele documenten kan de onderzoeker ook nog uit handen worden genomen door het samenvatten en eventueel evalueren van de literatuur ook aan een literatuuronderzoeker over te laten. De onderzoeker behoeft dan alleen kennis te nemen van het samenvattende literatuurrapport.

eindelijke acceptatie van deze dienstverlening. In hoofdstuk 6 werd daarover in detail bericht. Nadat gebruikers gedurende één tot twee jaar gratis een abonnement op de experimentele computerattenderingsdienst hadden gehad, bleek ongeveer 60% bereid te zijn het abonnement tegen betaling voort te zetten. Bij deze mensen woog de behoefte (of het nut) blijkbaar ten minste op tegen het gevraagde financiële offer. Dit percentage stemt zeker niet tot ontevredenheid en wijst op zichzelf op de aanwezigheid van een redelijke markt voor deze dienstverlening. Deze conclusie behoeft echter een aantal kwalificaties.
- Ten eerste speelt de hoogte van de gevraagde bijdrage een rol. Over de prijselasticiteit van de vraag zijn in feite nauwelijks cijfers bekend, al moet vanaf een zeker drempelniveau op een afnemende vraag bij toenemende prijs worden gerekend. Mede gezien de initiële onbekendheid van deze dienstverlening en de noodzakelijke aanpassing van traditionele werkpatronen leek het vooralsnog niet realistisch om de volledige kosten reeds in de prijs door te berekenen. (Zie 7.7.)
- Daarnaast speelt de beschikbaarheid van fondsen een rol. Er bestaan aanwijzingen dat financieel-organisatorische beperkingen momenteel nog een belangrijke rem vormen bij het gebruiken van extern literatuuronderzoek. (Zie verder 8.4.)

Over de invloed die het gebruik van de ALADIN-attenderingsdienst heeft gehad op de gewoonten voor literatuursignalering van de gebruikers, is helaas niets bekend. Het zou de moeite lonen nader te onderzoeken of het gebruik van externe attendering er inderdaad toe leidt dat de onderzoeker zelf minder primaire en secundaire tijdschriften doorneemt en in hoeverre het aanleiding geeft tot het lezen van meer literatuur.

8.3.3 Gebruikerscategorieën

Naast het nut van een literatuurattenderingsdienst voor potentiële gebruikers in het algemeen speelt bij de te verwachten en te stimuleren ontwikkelingen ook een rol voor welke gebruikerscategorieën deze vorm van dienstverlening het meest geschikt is. Ook al werd hiernaar geen formeel onderzoek gedaan, de algemene ervaringen opgedaan tijdens het project geven wel een aantal aanwijzingen.
- De informatie die momenteel in voor de computer leesbare vorm beschikbaar is lijkt hoofdzakelijk van nut voor het *wetenschappelijk onderzoek*. Voor het *management*, de *beleidsvorming* en de *voorlichting* is zij minder geschikt. Men is bij deze gebruikerscategorieën meer gediend met secundair verwerkte informatie die probleem-gericht is. Voor de voorlichting en voor toegepast onderzoek zoals dat in Nederland op de proefstations wordt verricht, zijn de problemen waarop men zich richt meestal zozeer plaatsgebonden dat de literatuur uit andere delen van de wereld niet zeer relevant zal zijn. De informatiebehoefte van het management is voor een groot deel kwantitatief en betreft gegevens die niet, of veel te laat, in de wetenschappelijke literatuur voorkomen. Hiervoor zijn heel andere informatiekanalen nodig.

Alle drie gebruikerscategorieën zijn geïnteresseerd in *tendensen* die zich in de praktijk en in het onderzoek aftekenen. Over deze tendensen zelf wordt echter weinig als wetenschappelijke literatuur gepubliceerd. Eventueel zou er door bestudering van wetenschappelijke literatuur wel iets over verschillende tendensen te zeggen zijn, maar dat is dan werk voor de literatuuronderzoeker (d.i. secundaire verwerking).

Ten slotte bestaat er bij management, beleidsvorming en voorlichting behoefte aan verschillende vormen van overzichtsartikelen, waarin de resultaten van het wetenschappelijk onderzoek vertaald zijn voor de informatiebehoeften van deze gebruikerscategorieën. Literatuuronderzoek is voor het schrijven van deze artikelen natuurlijk noodzakelijk; helaas maken auteurs van overzichtsartikelen tot op heden betrekkelijk weinig gebruik van de diensten van professionele literatuuronderzoekers.[63]

- Bij het onderzoek in de *industriële sector* is de formele waardering van informatie steeds hoger geweest dan bij het onderzoek in de publieke sector. Dit is met name af te leiden uit de pioniersrol die een aantal grote industriële ondernemingen hebben gespeeld bij de toepassing in de praktijk van computersystemen voor informatieselectie. Met name geldt dit in de chemische sector. Bij de landbouw heeft zich dit fenomeen tot nu toe minder uitgesproken voorgedaan. Dit wordt wellicht enerzijds veroorzaakt door de relatieve kleinschaligheid van het landbouwkundig onderzoek in de particuliere sector buiten de chemie, anderzijds door een grote gerichtheid op eigen toegepast onderzoek, waarbij elders opgedane ervaringen een relatief geringe rol spelen. De geringe beschikbaarheid van informatie over deze ervaringen speelt daarbij ook een rol. Daarnaast kan echter de chemici een vanouds grote gerichtheid op externe informatie niet worden ontzegd, zeker in vergelijking met landbouwkundigen, die in het algemeen wat dit betreft een minder goede naam hebben. Hun arbeidsterrein was (en is nog wel) daarbij vaak zo plaatsgebonden dat ervaringen elders opgedaan relatief gemakkelijk als niet ter zake doende opzij werden gelegd. De situatie bij Unilever Research is daarbij wellicht illustratief: terwijl het technologisch en chemisch onderzoek intensief gebruik maakt van de interne attenderingsdienst met behulp van Chemical Abstracts, FSTA en ISI (Science Citation Index), wordt bij het landbouwkundig (gewas-)onderzoek nog slechts in geringe mate gebruik gemaakt van een attenderingsdienst.

Ook kan het feit dat het landbouwkundig onderzoek in het algemeen in overheidsdienst wordt verricht een rol spelen; er is daardoor minder sprake van direct eigen belang dan bij de industrie. Het is uiteraard verheugend te constateren dat het nauw bij de industrie betrokken Instituut voor Bewaring en Verwerking van Landbouwprodukten (IBVL) vrij intensief gebruik maakt van de attenderingsdienst. Het aantal abonnees uit de industriële sector is verder echter gering. In hoeverre een te geringe voorlichting vanuit Pudoc hier mede debet aan is, moet nog nader worden onderzocht.

Tenslotte moet nog worden opgemerkt dat het belang van *snelle* literatuurinformatie d.m.v. een efficiënte attenderingsdienst voor de industrie veelal van groter belang is dan voor het wetenschappelijk onderzoek in de publieke sector.

Voor welk soort onderzoek een literatuurattendering vooral geschikt is valt in zijn algemeenheid moeilijk aan te geven. Dit hangt vooral af van de aandacht die ook elders aan een bepaald onderzoeksgebied wordt of werd gegeven en van de mate waarin het onderzoek kan profiteren van ervaringen elders op aanverwante gebieden. Tevens hangt het nut van

63. Zie hiervoor verder: Study of a pilot project for compiling review articles in agriculture. Final report on the activities resulting from the study contract between the European Economic Community and the Centre for Agricultural Publishing and Documentation (Pudoc). Pudoc, Wageningen, 1976.

literatuurattendering af van de informatie waarover de onderzoeker toch reeds beschikt zonder daar speciale moeite voor te hoeven doen (bijv. het 'invisible college') en zijn eigen houding ten aanzien van het uit handen geven van dit soort werk. Ook de beschikbaarheid van de noodzakelijke informatie speelt natuurlijk een rol: niet van alle onderzoeksgebieden zijn de primaire documenten goed ontsloten; dat de secundaire informatie nog niet beschikbaar is in de voor de computer leesbare vorm, is daarbij nog een extra beperking.

8.4 BELEMMERINGEN VOOR DE EXTERNE LITERATUURATTENDERING

Aannemende dat voor veel landbouwwetenschappelijke onderzoekers literatuurattendering met behulp van de computer van groot nut kan zijn, zal hier nader worden ingegaan op de factoren die het gebruik van een dergelijke dienstverlening vooral belemmeren. Deze liggen enerzijds in de persoonlijke sfeer, namelijk in de moeilijkheid bestaande werkpatronen te veranderen, en anderzijds in de financieel-organisatorische sfeer, waar nieuwe voorzieningen moeten worden getroffen en eventueel bestaande voorzieningen moeten worden omgebouwd.

8.4.1 De houding van de landbouwwetenschappelijke onderzoeker tegenover literatuuronderzoek

In 8.3.3 werd reeds gesignaleerd dat wetenschappelijke onderzoekers in de landbouw zich wat het gebruik van externe (literatuur-)informatie betreft nogal eens negatief onderscheiden van collega's in andere wetenschappen, met name chemici. Er werd in 8.3.1 reeds gewezen op het feit dat bij een aantal onderzoeksgebieden, met name in het toegepaste onderzoek, de problematiek zo situatie-gebonden is dat men weinig baat heeft bij informatie over ervaringen met elders verricht onderzoek. In het algemeen bestaat echter de indruk dat bij het landbouwwetenschappelijk onderzoek in Nederland minder gebruik wordt gemaakt van literatuurinformatie dan wenselijk zou zijn. De resultaten van de 8.3.1 genoemde COLADIN-enquête hebben deze indruk bevestigd. De opleiding tot landbouwkundig ingenieur is tot nog toe met name te kort geschoten in de training in efficiënt literatuurgebruik door aankomende ingenieurs. De instelling van een docentschap voor de wetenschappelijke informatieverzorging bij de Landbouwhogeschool zal hierin hopelijk verbetering brengen. Bij dit onderwijs komen de nieuwe mogelijkheden voor het literatuuronderzoek ook aan de orde.

Als de wetenschappelijk medewerker dan al overtuigd is van het belang van literatuurinformatie, dan moeten voor de introductie van nieuwe methoden voor literatuurattendering veelal bestaande werkpatronen worden doorbroken. Zoals bij iedere verandering van een bepaalde routine roept dit onwillekeurig weerstanden op. In dit geval gaat het er vooral om dat de gebruiker moet kunnen accepteren dat hij een gedeelte van de literatuursignalering die hij tot nu toe zelf deed, of die manueel binnen zijn instelling voor hem werd gedaan, beter door een computersysteem kan laten overnemen. Hij zal in de gelegenheid moeten worden gesteld om de resultaten van de computer-attendering een tijd lang met de op traditio-

nele wijze verkregen resultaten te vergelijken om aldus tot een verantwoorde keuze te komen. Voor ieder geval zal apart moeten worden bekeken of computer-attendering inderdaad een efficiënt hulpmiddel is en, zo ja, in welke mate bestaande werkzaamheden daardoor kunnen komen te vervallen.

Er blijkt van tevreden gebruikers van de computer-attendering helaas nauwelijks enige wervende werking uit te gaan. Hoewel door schriftelijke informatie in feite zo goed als alle landbouwkundige ingenieurs op de hoogte zouden moeten zijn van de nieuwe mogelijkheden, worden nieuwe abonnees bijna alleen verkregen na individuele benadering door Pudoc. In twee gevallen is duidelijk van een centrale beslissing sprake geweest, waarna er voor een groot aantal medewerkers abonnementen werden afgesloten (bij het Instituut voor Bewaring en Verwerking van Landbouwprodukten en bij de Vakgroep Fysische en Kolloïdchemie van de LH, maar verder is tot nu toe van enige verspreiding van deze informatie weinig gebleken).

8.4.2 Financiële en organisatorische belemmeringen

De belemmeringen van financiële en organisatorische aard zijn reeds in 6.2 en 6.3 aan de orde geweest bij de bespreking van de continuering nadat betaling werd gevraagd en in 7.4 bij de algemene overwegingen ten aanzien van financiering en prijsstelling.

Bij de meeste overheidsinstellingen voor landbouwkundig onderzoek is men niet gewend 'grote' bedragen voor individuele externe literatuurinformatie te betalen. De bibliotheekbudgetten van de meeste van deze instellingen zijn, voor de materiële kosten, betrekkelijk gering, zodat het zonder een aanmerkelijke verhoging van die budgetten niet goed mogelijk zal zijn om daaruit attenderingsabonnementen voor een groot deel van de wetenschappelijke staf te betalen. In feite zou het ook beter zijn om de kosten van literatuurinformatie voor een bepaald onderzoeksproject direct ten laste te brengen van dat project; bij de projectfinanciering wordt daarbij echter nu nog geen rekening gehouden, zodat het moeilijk is uit de projectfondsen geld voor literatuurinformatie vrij te maken, zelfs al zou men het er voor over hebben.

Het belang van budgettaire beperkingen werd geïllustreerd door de uitkomsten van de enquête naar de betalingsbereidheid, die in 6.2 werd behandeld. Hier kwam immers naar voren dat van de zes respondenten die de computer-attendering toch ten minste f 50 per jaar waard vonden, er maar drie ook konden toezeggen er daadwerkelijk fondsen voor ter beschikking te hebben. In de praktijk komt het nu herhaaldelijk voor dat alle literatuurinformatie van een beperkt bibliotheekbudget moet worden betaald en niet ten laste van een relatief ruim budget voor bijv. onderzoeksapparatuur kan worden gebracht, ook al is op dat laatste budget nog wel geld beschikbaar. Men zou soms een marginaal aan te schaffen apparaat waar nog geld voor was, best willen opofferen voor een attenderingsabonnement, maar door de interne financiële regelingen is een dergelijke keuze niet mogelijk.

Aangezien volledige projectfinanciering voorlopig nog uitblijft, lijkt het gewenst dat de onderzoeksinstellingen zo snel mogelijk financieel-organisatorische maatregelen nemen, die het mogelijk moeten maken op efficiënte wijze van individuele externe literatuurattendering gebruik te maken. Gelukkig is bij de Landbouwhogeschool intussen een zeer

bevredigende regeling tot stand gekomen. Er is daar namelijk een apart centraal budget gecreëerd voor extern literatuuronderzoek, waar alle vakgroepen gebruik van kunnen maken. De moeilijkheid om een goede verdeelsleutel tussen de vakgroepen te vinden, gegeven de onbekendheid van de behoefte per vakgroep, vormde hiervoor de belangrijkste overweging. De bibliothecaris oefent nu achteraf controle uit op het gebruik van deze gelden. Een literatuurattenderingsabonnement heeft bij deze regeling voor het vakgroep-budget geen consequenties, al is voor het gebruik ervan wel goedkeuring van de vakgroep nodig. Een dergelijke regeling zou natuurlijk een overmatig gebruik van deze dienst in de hand kunnen werken, maar dit is in de praktijk van de eerste twee jaar niet gebleken. Onder de huidige omstandigheden lijkt een dergelijke regeling een bijzonder gunstige oplossing. Hopelijk is zo'n centrale financiering in de naaste toekomst ook bij andere instellingen mogelijk.

Een extra probleem doet zich voor als externe literatuurattendering moet worden afgewogen tegen een reeds bestaand intern systeem. Deze laatste systemen werken veelal met zeer lage zichtbare kosten, omdat zij hoofdzakelijk tot stand komen met behulp van vast personeel. De individuele onderzoeker kan de externe abonnementskosten dan moeilijk verdedigen tegenover de interne dienst die noch de onderzoeker noch zijn instelling iets lijkt te kosten. De abonnementskosten zouden dan moeten opwegen tegen de waarde die de informatie van de externe dienst voor hem heeft *boven* die van de interne dienst, en die zou hij ook nog moeten betalen uit een onderzoeksbudget dat daar meestal geen voorzieningen voor heeft.

In een dergelijke situatie zouden eigenlijk de werkelijke totale kosten, de kwaliteit en de efficiëntie moeten worden vergeleken met die van de externe dienst, waarbij ook de eventuele veranderingen in personeelstaken in de beschouwing moeten worden betrokken. Bedacht moet daarbij ook worden dat abonnementen op de computer-attenderingsdienst vooral bij instellingen met een relatief kleine eigen bibliotheek een aanzienlijke verhoging in het aantal aanvragen voor originele documenten van elders met zich mee zal brengen. Hiervoor is ook enige personele capaciteit nodig. De waardering van de extra informatie die het externe systeem zal leveren, blijft echter ook hier het moeilijkste probleem. Alleen als de voordelen zeer evident zijn en het nieuwe systeem intern een aantal warme pleitbezorgers heeft, zal men er in zo'n geval toe komen traditionele patronen te verlaten. Tegenover een uitbreiding van de materiële kosten staan meestal alleen maar besparingen in de personele sfeer, waardoor de overgang naar externe dienstverlening nog extra bemoeilijkt wordt

In het algemeen geldt voor deze hele problematiek dat er principiële beslissingen nodig zijn, op een hoger niveau dan dat van de individuele onderzoeker, om de mogelijkheden voor daadwerkelijke invoering van externe literatuurattendering op grotere schaal te scheppen

8.5 EEN EIGEN BESTAND MET LITERATUURINFORMATIE

Herhaaldelijk is in het kader van het ALADIN-project gediscussieerd over de vraag of er in Nederland niet een eigen bestand met landbouwwetenschappelijke literatuurinformatie

zou moeten worden opgebouwd. Het kruisbestuivingseffect dat kan ontstaan bij de integratie van invoer en uitvoer kan namelijk van groot nut zijn. Op nationale schaal speelt daarbij een rol dat Nederland op het gebied van bibliografische bestanden op de internationale markt betrekkelijk weinig aan te bieden heeft.[64] Het is bovendien merkwaardig dat een organisatie als Pudoc, die weliswaar de rol van speerpunt speelt bij het gebruik van computersystemen voor het doorzoeken van bibliografische bestanden, de eigen lopende bibliografieën - het Pudoc Bulletin, Landbouwdocumentatie en de Euromarktlijst - niet in een voor de computer leesbare vorm gebruikt.

Voor het ontwikkelen van een eigen produkt op dit gebied kan men drie mogelijkheden onderscheiden: het opzetten van een geheel nieuw bibliografisch bestand (bijv. op een terrein dat nog slecht gedekt wordt); het in een voor de computer leesbare vorm brengen van een bestaande bibliografie; en het samenvoegen van onderdelen van bestaande bestanden naar een bepaald selectiecriterium. Deze drie mogelijkheden zullen achtereenvolgens nader worden behandeld.

8.5.1 Een nieuwe bibliografie op het gebied van de landbouw

Voor het tot stand brengen van een geheel nieuw bibliografisch bestand zijn de mogelijkheden gering. Er is een uitgebreide organisatie met medewerkers van hoog niveau nodig om zo'n onderneming te doen slagen. Sommige onderzoeksinstituten in Nederland zouden echter wel in staat geacht moeten worden om hiervoor de wetenschappelijke inbreng te kunnen leveren, waarbij Pudoc voor de organisatie en de technische uitvoering zou kunnen zorgen. Bij de onderzoeksinstituten ligt de prioriteit voor een dergelijke onderneming echter laag; de kans dat er iets in deze zin gerealiseerd kan worden lijkt gering, gegeven de huidige geringe personele en financiële armslag.

Pudoc onderzoekt intussen wel de mogelijkheid om 'Botanical Titles' weer tot leven te brengen. De opbouw van dit bestand werd na twee jaar gestaakt wegens financiële moeilijkheden. Er zijn reeds 80 000 titels verzameld door een net van internationale correspondenten.

8.5.2 De mogelijkheden voor het in computer-leesbare vorm brengen van bestaande Nederlandse bibliografieën op het gebied van de landbouw

De redenen dat Pudoc's eigen lopende bibliografieën niet als een voor de computer leesbaar bestand worden geproduceerd zijn van financieel-organisatorische aard. Zowel

---

64. Het 'Excerpta Medica'-bestand is ook internationaal gezien wel één van de zeer groten. Door haar zeer commerciële instelling neemt zij in het nationale en internationale overleg een speciale plaats in. De uitwisseling van gegevens en ervaringen met EM is veel geringer dan met de meerderheid der gouvernementele en niet met winst werkende instellingen.
Naast EM worden in Nederland nog gemaakt:
- 'Abstracts on Tropical Agriculture' door het Koninklijk Instituut voor de Tropen, en
- 'Economic Titles' door de Economische Voorlichtingsdienst.
Binnenkort is nog een bestand op het gebied van de hydrologie te verwachten.

Landbouwdocumentatie[65] als Landbouwkundige Aspecten van de Euromarkt[66] worden zuiver als attenderingsorganen gezien, waarbij een eventuele cumulatie nauwelijks van enig belang wordt geacht. Ook al is bij de produktie van dergelijke lopende bibliografieën het gebruik van een computersysteem in technische zin aantrekkelijk, de kosten van een dergelijke produktiewijze blijken vooralsnog aanzienlijk hoger dan die volgens de huidige, traditionele methode. Naarmate de kosten van computerverwerking verder dalen zou deze produktiewijze in de toekomst echter wel eens voordeliger kunnen worden. Mocht Pudoc voor andere doeleinden invoerapparatuur in huis krijgen, dan kan dit de computerverwerking van haar bibliografieën mede stimuleren. Als er eenmaal een voor een computer leesbaar bestand bestaat, kan altijd nog nader worden overwogen of dit ook voor andere doeleinden kan worden gebruikt, bijv. voor het maken van registers.

Voor Pudoc Bulletin[67] ligt de situatie enigszins anders. Van de titels wordt 40% ook voor AGRINDEX beschreven, terwijl het voor de AGRINDEX-invoer in de bedoeling ligt de Nederlandse bijdrage op magneetband te gaan leveren, zulks op verzoek van de Europese Commissie. Om dubbel werk te vermijden lijkt het logisch het relevante deel van de AGRINDEX-invoer direct van de band af voor Pudoc Bulletin te zetten. Of de rest van Pudoc Bulletin dan op dezelfde wijze zal moeten worden geproduceerd zal een kostenvergelijking moeten uitwijzen. In ieder geval is het niet denkbeeldig dat over één of twee jaar een Nederlandse landbouwkundige bibliografie in voor de computer leesbare vorm ter beschikking zal zijn. Met het cumulatieve bestand zullen dan ook gemakkelijk persoons-, onderwerps- en instituutsregisters kunnen worden samengesteld.

8.5.3 Het samenvoegen van bestaande bestanden

Er zijn nu voor het terrein van de landbouwwetenschap drie bibliografische bestanden, die het terrein in al zijn geledingen betrekkelijk goed dekken. Binnen enkele jaren zal het hopelijk niet meer nodig zijn om CAIN naast AGRINDEX te doorzoeken, maar het CAB-bestand zal voorlopig wel naast AGRINDEX blijven bestaan (zie 3.4). Naast deze drie algemene bestanden is het voor vele landbouwkundige onderwerpen noodzakelijk om nog andere bestanden te doorzoeken, zoals Chemical Abstracts, Biological Abstracts en FSTA. Veel van de gevonden relevante titels zullen daarbij in meer dan één bestand voorkomen, waarbij zich de vraag voor doet of deze verdubbeling van werk (zowel bij de invoer als bij de uitvoer) niet vermeden had kunnen worden. Een oplossing zou kunnen worden gevonden in het samenvoegen van (onderdelen van) de relevante bestanden. Bij zo'n samenvoeging (data base-merging) worden alle verwijzingen naar hetzelfde oorspronkelijke document bij elkaar gezet

---

65. 'Landbouwdocumentatie' is een wekelijks verschijnende selectie uit de landbouwkundige literatuur uit landen met een gematigd klimaat. De selectie geschiedt met het oog op de technische, wetenschappelijke en economische belangen van de landbouw in Nederland. Het blad is bestemd voor de 'tussen'- of 'doorgeef'groep (voorlichting en onderwijs) en het landbouwbedrijfsleven.

66. 'Landbouwkundige Aspecten van de Euromarkt' is een maandelijks verschijnende titelbibliografie van publikaties m.b.t. de Euromarkt voorzover van belang voor de landbouw.

67. 'Pudoc Bulletin' is een viermaal per jaar verschijnende titelbibliografie in de Engelse taal van publikaties van resultaten van wetenschappelijk landbouwkundig onderzoek in en buiten Nederland door Nederlandse onderzoekers gepubliceerd.

en wordt alle informatie die alle bestanden samen over die referentie bevatten, in één record samengevat. (Men neemt ook wel alleen alle informatie van een voorkeursbestand op.) Een voorbeeld van een succesvolle bestandssamenvoeging is TOXLINE, waarin informatie van acht verschillende bestanden is opgenomen. Voor vragen op het gebied van de toxicologie is het nu niet meer nodig deze acht bestanden ieder afzonderlijk te doorzoeken, maar is het doorzoeken van één (afgeleid) bestand voldoende.

Er is met de gedachte gespeeld voor specifiek Nederlandse behoeften een aantal afgeleide bestanden samen te stellen, waardoor de gebruiker geen dubbele verwijzingen meer krijgt en met name het terugzoeken goedkoper kan worden naarmate ten minste een aantal jaren wordt gedekt. Het opbouwen en instandhouden van een bestand van enige omvang (ten minste enkele tienduizenden titels of aanzienlijk meer, afhankelijk van de mate van specialisatie) is echter een zeer dure aangelegenheid. Voor gebruik binnen Nederland alleen is dit in het algemeen niet te rechtvaardigen. Pas als er voor zo'n bestand een redelijke internationale markt zou zijn, valt aan verdere ontwikkeling te denken. Het is moeilijk voor te stellen dat een bestand dat afgestemd zou zijn op de specifieke behoefte van het Nederlandse landbouwwetenschappelijk onderzoek zich duidelijk zou onderscheiden van een bestand gericht op de behoeften van collega's in andere landen in de gematigde luchtstreken.

Dat men tot nu toe in Nederland nog niet verder is gekomen, is ook te wijten aan een gebrek aan gekwalificeerde krachten om ideeën in deze richting verder gestalte te geven en eventueel uit te voeren. Met name in het kader van de werkzaamheden op dit gebied van de Europese Commissie moet het idee voor de toekomst echter zeker levend gehouden worden: nu er steeds meer informatiebestanden ter beschikking komen om met behulp van computersystemen te doorzoeken, wordt het steeds moeilijker en duurder om alle relevante bestanden te identificeren en te doorzoeken. De noodzaak tot het samenstellen van afgeleide bestanden zal zich daardoor steeds sterker doen gevoelen.

8.6 TOEKOMSTIGE ONTWIKKELINGEN

Op dit moment beschikt het wetenschappelijk landbouwkundig onderzoek in Nederland over een literatuurattenderingsdienst die in staat is de meerderheid van de wetenschappelijke onderzoekers regelmatig op de hoogte te houden van de nieuwe literatuur op hun vakgebied. De geleverde informatie zal het grootste deel van de gepubliceerde relevante literatuur kunnen dekken. Door het systeem zal uiteraard het zelf literatuur verzamelen door de onderzoeker niet geheel kunnen worden vervangen, maar wel voor een belangrijk deel uit handen kunnen worden genomen. Hiermee kan worden bereikt dat de onderzoeker meer tijd voor zijn eigen onderzoek vrij krijgt en wordt de kans aanzienlijk vergroot dat hij op de hoogte raakt van relevante informatie waarvan hij anders onkundig zou zijn gebleven.

In technische zin is op dit ogenblik eerder consolidatie nodig dan het verder uitbreiden van het systeem. Er zijn nog wel enkele kleine verbeteringen van het computerprogramma mogelijk. Het ware ook te overwegen om Biological Abstracts en Botanical Titles in eigen beheer te gaan verwerken. Een kwantitatieve uitbreiding moet echter in dit stadium een veel grotere prioriteit krijgen dan de verdere uitbreiding van de mogelijkheden.

Wel zou echter moeten worden nagegaan of het gebruik van de huidige computer bij

IWIS-TNO optimaal is. Onderzocht moet worden of mini-computers niet een goedkoper alternatief bieden. De aanschaf van zo'n computer plus het vervaardigen van de hiervoor noodzakelijke programmatuur kan een investering van wel f 200 000 eisen. Indien echter tot continuering van het systeem over een langere periode wordt besloten, kan zo'n investering over een langere termijn worden afgeschreven.

Onderzocht zou ook moeten worden in hoeverre het gebruik van de nieuwe literatuur-informatiediensten het *bibliotheekgebruik* beïnvloedt. Het is nog onvoldoende bekend in welke mate de huidige bibliotheekvoorzieningen efficiënt aan verzoeken om originele teksten kunnen voldoen. Het is denkbaar dat de gebruikers zich bij het aanvragen van publikaties waarop zij zijn geattendeerd laten afschrikken doordat deze op het eerste oog moeilijk verkrijgbaar lijken. Bovendien is het noodzakelijk inzicht te verkrijgen in de vraag naar vertalingen uit moeilijk toegankelijke talen. Deze talen blijken betrekkelijk veel bij de geattendeerde titels voor te komen. De neiging bestond dergelijke titels op grond van de ontoegankelijke taal als irrelevant te beoordelen, terwijl de vertaalde titel op zichzelf wel 'ter zake' was.

De toekomst van de nieuwe vormen van literatuur-informatiediensten is geheel afhankelijk van de mate waarin het onderzoeksapparaat deze diensten ook daadwerkelijk gaat gebruiken. De potentiële vraag lijkt voorlopig nog aanzienlijk groter dan de huidige. In het afgelopen jaar is de vraag duidelijk toegenomen: van 95 profielen in januari 1976 tot 129 eind oktober (een stijging van 36%). Vermoedelijk kan echter pas bij 500 abonnees gezegd worden dat de meerderheid van de landbouwkundige onderzoekers van externe literatuurattendering gebruik maakt. In hoeverre een uitbreiding van het aantal abonnees - en het bestaan van deze dienstverlening in het algemeen - mogelijk is, is een financiële en daarmee een beleidskwestie. Op dit moment dekt de gevraagde abonnementsprijs de totale kosten niet. Alleen als dit wel het geval zou zijn, zou de vraag betreffende de *handhaving* van deze dienstverlening niet behoeven te worden gesteld. Zowel de structuur van de gevraagde abonnementsprijzen als de financiering van de abonnementen zal in de naaste toekomst nader moeten worden bezien.

Ten aanzien van de prijsstructuur blijft de vraag actueel of, en in hoeverre, de prijsverschillen tussen de diverse bestanden de verschillen in kosten voor het verwerken van de bestanden moeten weerspiegelen of dat één prijs voor attendering moet gelden, waarbij Pudoc de optimale bestandssamenstelling vaststelt. Dit probleem werd in 7.5 behandeld.

Ten aanzien van de financiering is nader overleg met de Directie Landbouwkundig Onderzoek (DLO) noodzakelijk. Het feit dat de instituten voor landbouwkundig onderzoek niet over een speciaal budget voor informatievoorziening beschikken, blijkt een belangrijke belemmering voor het gebruik van externe literatuur-informatiediensten te zijn. Het speciale budget voor externe literatuurinformatie bij de Landbouwhogeschool heeft het gebruik van deze informatie in belangrijke mate bevorderd. Of de financiering van literatuur-informatie voor de onder DLO ressorterende instituten (en eventueel de andere onderzoeksinstellingen van het Ministerie van Landbouw en Visserij) via aparte budgettaire voorzieningen per instituut, via een speciale voorziening bij DLO of via de begroting van Pudoc zou moeten lopen, moet nader worden onderzocht.

Verbetering van de financiering van literatuur-informatievoorziening verdient overigens niet alleen de aandacht van DLO, maar ook van het verdere wetenschappelijke onder-

zoek binnen het Ministerie van Landbouw en Visserij en trouwens ook daarbuiten.

8.7 SLOTOPMERKING

Indien men d.m.v. een 'pilot study' in de praktijk wil nagaan of een bepaalde dienst in technische en economische zin verantwoord kan worden verleend is het noodzakelijk rekening te houden met een vrij lange periode tussen de eindrapportage waarop de beslissing over continuering moet worden gebaseerd, en het moment waarop deze beslissing uiteindelijk kan worden genomen. Als het opgebouwde systeem voor de technische uitvoering gedurende die periode wordt stopgezet, zullen er bij het begin van een eventuele operationele fase grote vertragingen ontstaan met daaraan verbonden extra kosten. Anderzijds mag voortzetting tijdens de beslissingsperiode de beslissing niet in de richting van voortzetting dwingen.

9 Samenvatting en conclusie

Het doel van het project 'Automatisering Landbouwkundige Documentatie en Informatieverspreiding in Nederland' (ALADIN) was het evalueren van de mogelijkheden en de bruikbaarheid van en de behoefte aan een literatuurattenderingsdienst met behulp van magneetbanden ten behoeve van wetenschappelijk onderzoek op het gebied van de landbouw.

Het project werd geconcipieerd door de Commissie Landbouwkundige Documentatie en Informatie (COLADIN) van de Nationale Raad Landbouwkundig Onderzoek TNO (NRLO-TNO), gesubsidieerd door de Stichting Nederlands Orgaan voor de Bevordering van de Informatieverzorging (NOBIN) en uitgevoerd door het Centrum voor Landbouwpublikaties en Landbouwdocumentatie (Pudoc). De uitvoering nam een aanvang in 1972. Het project werd eind 1975 afgesloten.

Literatuurattendering met behulp van een computer berust op het regelmatig afzoeken van magneetbanden met bibliografische informatie d.m.v. een computerprogramma dat uit de verzameling referenties op de band die titels selecteert die voor een bepaald onderwerp relevant zijn.

De gebruikte magneetbanden komen in het algemeen overeen met gedrukte versies van lopende bibliografieën. De belangrijkste zes, die voor het project werden gebruikt, waren:
Bibliography of Agriculture,
AGRINDEX,
de referaattijdschriften van de Commonwealth Agricultural Bureaux,
Food Science and Technology Abstracts,
Chemical Abstracts, en
Biological Abstracts.
De eerste vier werden in eigen beheer verwerkt door het Instituut TNO voor Wiskunde, Informatieverwerking en Statistiek (IWIS-TNO), de andere twee door tussenkomst van derden.

Voor ongeveer 100 onderwerpen werden zgn. literatuur-interesseprofielen opgesteld. Dit zijn samenstellingen van trefwoorden, waarmee het computer-zoekprogramma voor het onderwerp relevante titels op de magneetband opspoort. Van de onderwerpen werden er 75 geleverd door wetenschappelijke onderzoekers op het gebied van de landbouw, diergeneeskunde en voedseltechnologie. Deze 'leden van de experimentele gebruikerskring' beoordeelden iedere maand de geleverde titels op hun relevantie en nieuwswaarde. De gebruikers ontvingen gemiddeld 9 relevante titels per maand plus (nog) 4 waarvan de relevantie niet direct te beoordelen was. Ongeveer 60% van deze titels kende de onderzoeker nog niet. Na afloop van de gratis periode zette 67% van hen het attenderingsabonnement tegen betaling voort.

Van 25 onderwerpen werden de resultaten van computer-attendering vergeleken met de uitkomsten van de manuele attenderingsdienst van Pudoc. De laatste is gebaseerd op de originele documenten die de Centrale Bibliotheek van de Landbouwhogeschool bereiken. Het aantal d.m.v. computer-attendering gevonden relevante titels bleek aanmerkelijk groter

dan het aantal dat manueel werd gevonden: 77% van alle titels tegenover 35%. Er bleken echter duidelijk een aantal categorieën vragen te zijn waarbij manuele attendering moeilijk door computer-attendering kon worden vervangen.

Manuele attendering was in het algemeen ook iets sneller dan computer-attendering, hoewel dit verschil meestal niet meer dan een paar maanden bedroeg.

De kosten van het in eigen beheer verwerken van magneetbanden voor literatuurattendering zijn zeer afhankelijk van het aantal literatuur-interesseprofielen dat wordt verwerkt. Op basis van 100 profielen werden de verwerkings- plus de abonnementskosten ruwweg op ƒ 1250 per profiel per jaar geschat; op basis van 500 profielen zouden deze kosten gehalveerd worden. Aan personele kosten komt daar bij 100 profielen ƒ 1000 en bij 500 profielen ƒ 200 per profiel bij. De personeelskosten zijn echter niet zozeer afhankelijk van het aantal verwerkte eenheden, maar vooral van het aantal nieuw op te stellen profielen. Voor 1977 wordt het totaal aantal profielen op 175 geschat, waarvan de kosten per profiel per jaar dan ruwweg ƒ 1500 zijn. Voor manuele attendering moet minimaal op ƒ 3400 worden gerekend. Bij de voor 1977 vastgestelde prijzen kost een gemiddeld abonnement aan de gebruiker ƒ 850. Bij 450 abonnees zou deze prijs geheel kostendekkend zijn. Bij 250 abonnees zouden alleen de materiële kosten worden gedekt. Het Ministerie van Landbouw en Visserij verleent op ieder abonnement voor medewerkers van dit Ministerie een korting van 50% op de geldende prijzen.

De *financiering van het project is na* de beëindiging van de subsidiëring door NOBIN geheel door het Ministerie van Landbouw en Visserij overgenomen en is geïntegreerd in de boekhouding van Pudoc.

Bij *centrale manuele attendering* is verdere uitbreiding van het volume moeilijk, terwijl er bij de computer-attendering nauwelijks technische grenzen zijn. De vergelijking van de resultaten van beide attenderingsdiensten leidt tot de conclusie dat computer-attendering doorgaans betere resultaten geeft dan manuele attendering zoals die bij Pudoc bestaat, al zijn er een aantal speciale onderwerpen en zoekcriteria waarvoor manuele attendering zich moeilijk door de computer laat vervangen. De literatuurattenderingsdienst met behulp van computersystemen lijkt binnen het landbouwwetenschappelijke onderzoek een duidelijke markt te hebben, gezien het hoge percentage gebruikers die hun abonnement continueerden nadat betaling werd gevraagd (60%) en gezien de snelle groei van het aantal abonnementen sindsdien: 70% in twee jaar.

Of de voordelen tegen de kosten opwegen is moeilijk te berekenen. Naast de tijdsbesparing die het de onderzoeker oplevert (hij behoeft immers veel minder tijdschriften regelmatig door te nemen), bestaat de winst vooral ook uit het onder ogen krijgen van literatuur die anders gemist zou zijn of waarvan niet tijdig kennis zou zijn genomen. De besparingen die het tijdig op de hoogte raken van relevante onderzoeksresultaten van anderen op kunnen leveren, zijn nu eenmaal moeilijk in geld uit te drukken maar kunnen groot zijn.

Twee factoren werden onderkend die belemmerend werken op de externe literatuurattendering voor landbouwwetenschappelijke onderzoekers. Allereerst is dat de traditionele werkhouding van de onderzoeker. Om externe literatuurattendering te gaan gebruiken zal een bestaand patroon moeten worden doorbroken, hetgeen waarschijnlijk slechts zeer gelei-

delijk kan gebeuren. Ook zijn nog lang niet alle potentiële gebruikers voldoende op de hoogte van de mogelijkheden van externe literatuurattendering voor hun specifieke situatie. Ten tweede zijn er financieel-organisatorische belemmeringen, die het de potentiële gebruikers moeilijk maken om van deze diensten gebruik te maken. Voor externe literatuurinformatie zijn veelal nog geen speciale voorzieningen getroffen, terwijl besparingen op het personele budget veelal moeilijk te vereffenen zijn met extra uitgaven van het materiële budget. Een centraal budget voor externe literatuurinformatie per instelling lijkt vooralsnog de beste oplossing, zolang in de overheidssfeer niet *alle* project-gebonden kosten voor rekening van het project-budget komen. Bij de Landbouwhogeschool werden bijzonder goede ervaringen met een centraal budget opgedaan.

Ten slotte mag worden gesteld dat het project aan zijn doel heeft beantwoord. Het landbouwkundige onderzoek in Nederland beschikt nu over een technisch goed functionerende literatuurattenderingsdienst, gebaseerd op het verwerken van magneetbanden met behulp van een computersysteem. Deze attenderingsdienst blijkt een zeer waardevolle uitbreiding van de mogelijkheden voor literatuurattendering, zowel wat betreft het volume van de dienstverlening als de eenvoudigere beschikbaarheid van een zeer grote hoeveelheid referenties. In het algemeen zal externe attendering de onderzoeker daarbij zowel tijd- als informatiewinst opleveren.

De nu operationele attenderingsdienst mag zich in een groeiende belangstelling verheugen, ondanks de beperkingen op het gebied van de financiering, de nog geringe bekendheid en de nog nauwelijks doorbroken traditionele werkpatronen op het gebied van de informatieverzorging.

Lijst met gebruikte afkortingen

AGRINDEX: Product van het AGRIS-systeem; de gedrukte maandelijkse bibliografie
AGRIS : Agricultural Research Information System (van FAO)
ALADIN : Automatisering Landbouwkundige Documentatie en Informatieverspreiding in Nederland
BA : Biological Abstracts, geproduceerd door Biosis. (Het gebruikte bestand heet officieel Biosis Previews. Het bevat behalve Biological Abstracts ook de Bio-Research Index.)
BIOSIS : Bio Sciences Information Service of Biological Abstracts
BMDC : Bio Medical Documentation Centre
BoA : Bibliography of Agriculture
CAB : Commonwealth Agricultural Bureaux
CABVET : CAB-Veterinair. (Het diergeneeskundig gedeelte van het CAB-bestand. Dit wordt gevormd door Veterinary Bulletin samen met Index Veterinarius.)
CAC : Chemical Abstracts Condensates
CAIN : Cataloguing and Indexing System (bibliografisch bestand van de NAL)
COLADIN : Commissie Landbouwkundige Documentatie en Informatie (van de NRLO-TNO)
EM : Excerpta Medica
FAO : Food and Agriculture Organisation of the United Nations
FLIRT : Fast Literature Information Retrieval Tool (computerprogramma van IWIS-TNO)
FSTA : Food Science and Technology Abstracts
INIS : International Nuclear Information System
IWIS-TNO: Instituut voor Wiskunde, Informatieverwerking en Statistiek - TNO
LH : Landbouwhogeschool
MED : Zie Medlars
MEDLARS : Medical Literature Analysis and Retrieval System (van de National Library of Medicine - USA)
MvLV : Ministerie van Landbouw en Visserij
NAL : National Agricultural Library (van de USA)
NOBIN : Stichting Nederlands Orgaan voor de Bevordering van de Informatieverzorging
NOCI : Nederlandse Organisatie voor Chemische Informatie
NRLO-TNO: Nationale Raad Landbouwkundig Onderzoek - TNO
Pudoc : Centrum voor Landbouwpublikaties en Landbouwdocumentatie
TNO : Nederlandse Centrale Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek

Verklarende woordenlijst

batch verwerking - Bij batch verwerking van een computerprogramma wordt het totaal van het te verwerken materiaal als 'stapel' aangeboden en verwerkt. Er kan in principe tijdens de verwerking niet worden ingegrepen om het programma te wijzigen aan de hand van op dat moment reeds verkregen resultaten. Tegendeel van interactieve verwerking.

bestand - Verzameling van informatie in computer-leesbare vorm. In het verband van dit rapport in het bijzonder bibliografische informatie, zoals regelmatig verschijnende bibliografieën en referaattijdschriften. Ook wel: data base.

conversie - Omzetting, bij. het veranderen van de volgorde van de elementen die tot één registratie-eenheid (record) behoren.

format - Het samenstel van voorschriften voor een eenduidige beschrijving van gegevens op een informatiedrager (bijvoorbeeld een magneetband). Ook: vormbeschrijving.

hit - Treffer. Bij het afzoeken van een bestand ontstaat een *treffer*, als het record voldoet aan alle eisen die het zoekprofiel stelt. De vereiste combinatie van trefwoorden en codes komt dan inderdaad in dat record voor.

interactieve verwerking - Computerverwerking van gegevens, waarbij telkens kan worden bijgestuurd aan de hand van resultaten verkregen uit deelverwerkingen. Interactieve systemen worden ook wel 'conversationeel' genoemd, omdat de verwerking zich afspeelt in een vraag- en antwoord-'spel' tussen gebruiker en computer.

literatuur-interesseprofiel - In algemene zin een zo nauwkeurig mogelijke bschrijving van een bepaald onderwerp, gericht op het opsporen van literatuur die op dat onderwerp betrekking heeft. In engere zin een opstelling van trefwoorden, eventueel samen met codes, auteursnamen etc. waarmee een computer-retrieval-programma titels selecteert uit een bibliografisch bestand. Een profiel is bij de huidige opzet van Pudoc's attenderingsdienst in principe onafhankelijk van het gebruikte bestand, zodat er maar één profielopstelling nodig is voor alle te doorzoeken bestanden. Eventuele bestand-specifieke codes worden voor de andere bestanden inoperatief gemaakt.

profiel - Zie 'literatuur-interesseprofiel'.

record - Registratie-eenheid. In het verband van dit rapport alle informatie die met betrekking tot één titel in het bestand is opgenomen.

retrieval - Terugzoeken. Een *retrieval* programma is een computerprogramma waarmee records uit een bestand worden opgezocht, die voldoen aan een aantal voor ieder geval apart opgestelde eisen.

retrospectief - Teruggaand in de tijd. Bij *retrospectief* literatuuronderzoek wordt in principe naar alle publikaties gezocht die over een bepaald onderwerp in het verleden zijn verschenen. Zulks in tegenstelling tot *attendering*, waarbij alleen uit nieuwe literatuur wordt geselecteerd.

run - Het verwerken van één magneetband. In het verband van het ALADIN-project betrof dit steeds een magneetband met bibliografische informatie die meestal overeen kwam met de maandelijkse aflevering van het referaattijdschrift.

verwerkingseenheid - Eén profiel verwerkt met één bestand.

zoekprofiel - Zie 'literatuur-interesseprofiel'.

Bijlagen

1. Enquête documentatie- en informatie-verstrekking in Nederland op het gebied van de agrarische en veterinaire wetenschappen 102
2. Verwerkingsdata van de CAIN-, BoA- en FSTA-banden bij IWIS-TNO 1972-1974 114
3. Literatuur-interesseprofielen waarvoor de resultaten van manuele attendering werden vergeleken met die van computer-attendering. Tevens aanduiding van de gebruikte bibliografische bestanden op magneetband 116
4. Overlapping tussen manuele en computer-attendering van relevante artikelen gepubliceerd in 1973 voor de 25 onderwerpen genoemd in Bijlage 3 117
5. Overlapping tussen manuele en computer-attendering exclusief misschien relevante titels. Zie Bijlage 4 118
6. Lijst van profielonderwerpen die voor het ALADIN-project werden verwerkt en beoordeeld door de gebruikers 119
7. Lijst van instellingen resp. vakgroepen waarvan medewerkers hebben deelgenomen aan het ALADIN-project 120
8. Evaluatieformulier dat met iedere output werd meegezonden 121
9. Enquêteformulier betreffende (A) financiële waardering van computer-attendering en (B) betreffende informatie per titel van een artikel 122
10. Enquête onder ex-ALADIN-gebruikers 125
11. Aanvraagformulier voor een literatuuronderzoek 129
12. Tarieven literatuuronderzoek (per 1-1-1977) 133

Bijlage 1. Enquête documentatie- en informatie-verstrekking in Nederland op het gebied van de agrarische en veterinaire wetenschappen.

Gehouden onder auspiciën van de Commissie voor Landbouwdocumentatie en -Informatie van de Nationale Raad Landbouwkundig Onderzoek - TNO.

Een analyse van de resultaten.

Ir. H.C. Molster
Augustus 1974

Inleiding

De enquête werd gehouden in het najaar van 1971. Door gebrek aan personeel heeft de verwerking van de resultaten jammer genoeg erg lang op zich laten wachten. Hoewel zich in de tussentijd zeer snelle ontwikkelingen hebben voorgedaan met betrekking tot de toepassingsmogelijkheden van moderne informatietechnieken bij de documentatie- en informatieverstrekking, is het gebruik hiervan in de agrarische en veterinaire sector pas eigenlijk het laatste jaar enigszins op gang gekomen, zodat de situatie, zoals die in 1971 heerste, nog niet essentieel gewijzigd zal zijn.

Methodiek

Na identificatie van de niet van toepassing zijnde en overigens onbeantwoorde (sub-)vragen, werd het data bestand overgebracht op ponskaarten, waarna een correlatie matrix werd berekend. Deze werd vervolgens gebruikt voor een factoranalyse. In eerste instantie leken de op deze wijze verkregen resultaten voldoende. Kruistabellen werden niet verder opgesteld.

Aan de vragen van het enquêteformulier werden een aantal variabelen toegevoegd: een lijst van 21 meest gebruikte referaattijdschriften (uit het antwoord op vraag 1.1.1), en een aantal categorieën, waarin de responderende instellingen konden worden verdeeld en het aantal wetenschappelijke medewerkers, om een indruk te krijgen van de grootte van de instelling.

Het lag in de bedoeling die cijfers, waarvan verwacht mocht worden dat ze sterk met de grootte van de instelling zouden samenhangen, eerst door dit getal te delen, maar het gebruikte computerprogramma miste deze faciliteit, waardoor het toch verder achterwege bleef.

De enquête werd naar 169 instellingen verstuurd, waarvan werd verondersteld dat deze bibliotheek- en misschien ook documentatiewerk verrichten dat betrekking had op de landbouw in de meest brede zin. Vijf hiervan vielen af, omdat de desbetreffende instelling en geen eigen bibliotheek hadden noch documentatiewerk

verrichtten. 102 antwoorden bleken uiteindelijk bruikbaar voor verwerking. 17 instellingen meenden, dat de enquête niet van toepassing was op hun instelling, zoals afdelingen van de Landbouwhogeschool, die zich hoegenaamd niet met landbouwkundige vraagstukken bezighouden. Van 45 instellingen werd taal noch teken vernomen, ook niet na een herhaald verzoek.

In welke mate de niet-respondenten de resultaten van de enquête scheef zouden kunnen trekken, kan getest worden aan de verdeling van respondenten en niet-respondenten over een aantal categorieën:

Categorie*	verwerkt aantal	%	verzonden aantal	%
Onderwijs	39	38	61	37
Landbouwhogeschool	31	30	53	32
Veterinaire Faculteit	8	8	8	5
Onderzoeksinstituten	27	26	35	21
Bedrijfsleven	11	11	24	15
Proefstations	5	5	16	10
Gezondheidsdiensten	8	8	13	8
Diversen	13	13	18	11
Wetenschappelijk onderzoek	87	85	127	77
Routine onderzoek	29	28	49	30
Voorlichting	25	25	37	23
Totaal	102		164	

In het algemeen blijkt de verschuiving niet al te groot. Onderzoeksinstellingen zijn bij de verwerking wat over-vertegenwoordigd, terwijl het bedrijfsleven en de proefstations onder-vertegenwoordigd zijn.

---

* Één instelling kan in verschillende categorieën voorkomen.

Resultaten per vraag*

Vraag 1

Zo goed als alle respondenten gebruiken secundaire informatiebronnen, alleen niet:

14 Inst. v. Mout, Gerst en Bier TNO
20 Inst. v. Zuivelbereiding en Melkkunde TNO
21 Bedrijfslab. v.d. Landbouw, Friesland
8 Rechts- en Staatw. m.w.g. (LH Afd.)
54 K.N.M.I. afd. III
57 IWIS-TNO, Wageningen

Vraag 1.1

Het aantal externe bronnen waarvan gebruik gemaakt werd was gemiddeld 9. Iets minder voor de Voorlichtings en Routine instellingen, iets meer voor de Instituten.

Er is een positieve correlatie met het aantal externe verzoeken om literatuur-inlichtingen (vraag 3.b). De Instituten gebruiken in het algemeen meer externe bronnen, dan andere instellingen.

Vraag 1.1.1

Van de met name genoemde externe secundaire informatiebronnen werd de volgende frequentietabel samengesteld:

* n = aantal antwoorden waarop de berekening is gebaseerd.
S = standaard deviatie.

Lijst van titels van meer dan een maal genoemde externe secundaire bronnen naar frequentie gerangschikt, en daarbinnen alfabetisch.

Landbouwdocumentatie	39
Chemical Abstracts	20
Veterinary Bulletin	16
Biological Abstracts	15
Current Contents	15
Horticultural Abstracts	14
Annual Review of Biochemistry	12
Field Crop Abstracts	11
Plant Breeding Abstracts	11
Soils and Fertilizers	11
Nutrition Abstracts and Reviews	10
Weed Abstracts	10
Annual Review of Plant Physiology	9
Food Science and Technology Abstracts	9
Herbage Abstracts	9
Tropical Abstracts	9
Analytical Abstracts	7
Animal Breeding Abstracts	7
Documentation East-Europen Agricultural Literature	7
Forestry Abstracts	7
Bibliography of Agriculture	6
Dairy Science Abstracts	6
Journal of the Science of Food and Agriculture	6
Advances in Agronomy	5
Review of Applied Entomology	5
Review of Plant Pathology	5
Annual Review of Microbiology	4
Excerpta Medica	4
Helminthological Abstracts	4
Landwirtschaftliches Zentralblatt	4
Pesticides Documentation Bulletin	4
Pudoc Bulletin	4
Abstracts of Mycology	3
Bacteriological Reviews	3
Bulletin Signalétique	3
Economic Abstracts	3
Environmental Health	3
Review of Applid Mycology	3
Uit de Pluimveepers	3
World Agricultural Economics and Rural Sociological Abstracts	3
Zeitschrift für Pflanzenernährung und Bodenkunde	3
Zentralblatt für Bakteriologie u.s.w.	3
Nuclear Science Abstracts	3
Dissertation Abstracts	2

Lijst van titels van meer dan een maal genoemde externe secundaire bronnen naar frequentie gerangschikt, en daarbinnen alfabetisch. (vervolg)

Advances in Enzymology	2
Advances in Immunology	2
Advances in Veterinary Science	2
Air Pollution Abstracts	2
Annual Review of Pharmacology	2
Annual Review of Physiology	2
BFMIRA-Abstracts	2
Bioresearch Index	2
Botanical Review	2
Bulletin d'Institut International du Froid	2
Das Schriftum der Agrarwirtschaft	2
Entomology Abstracts	2
Excerpta Botanica	2
Fortschritte der Botanik	2
Fruits	2
Genetics Abstracts	2
Health Aspects of Pesticides; Abstract Bulletin	2
Journal of Animal Science	2
Journal of Economic Literature	2
L.E.I. Literature Bulletin	2
Literatuurlijst van het PUDOC	2
Pest Articles and News Summaries	2
Pollution Abstracts	2
Recreatiedocumentatie	2
Statistical Theory and Management Abstracts	2
Tijdschrift foor Diergeneeskunde	2
Water Pollution Abstracts	2

De hoge score van Landbouwdocumentatie is opvallend. De correlatiecoëfficiënten laten zien dat LD vooral gebruikt wordt door het bedrijfsleven en de voorlichting, terwijl het relatief weinig wordt genoemd door het onderwijs in het algemeen en de Veterinaire Faculteit in het bijzonder. Als LD wordt gebruikt is de kans relatief groot, dat ook Soils and Fertilizers, Field Crops Abstracts, Plant Breeding Abstracts, Weed Abstracts en Herbage Abstracts worden gebruikt.
Dus met name de teeltkundige CAB-tijdschriften. Er is ook een positieve correlatie met de tijd, die onderzoekers besteden aan het doorzoeken van primaire bronnen (2.2).

Chemical Abstracts wordt vooral op de Instituten gebruikt en weinig door de voorlichting.

Current Contents wordt vooral door de Vet. Fac. veel gebruikt; ook wat meer dan gemiddeld op de Instituten. Er is ook hier een positieve correlatie met de tijd, die onderzoekers besteden aan het raadplegen van primaire bronnen.
Current Contents komt veel voor naast Chemical Abstracts en vooral de Annual Review of Biochemistry en de Annual Review of Plant Physiology. Het wordt relatief veel gebruikt bij instellingen met een grote wetenschappelijke staf.

Biological Abstracts wordt ook relatief veel op de Instituten gebruikt, weinig bij het onderwijs, de Gezondheidsdiensten, de Vet. Fac. en bij Routine instellingen.
Biological Abstracts wordt relatief veel gebruikt bij instellingen waar veel tijd wordt besteed aan het zoeken in secundaire bronnen t.b.v. derden (1.3).

Over de andere algemenere referaattijdschriften kan nog het volgende worden opgemerkt:

Documentation on East-European Literature werd vooral gebruikt door enkele Instituten:
Tuinbouwplantenteelt LH, Landbouwplantenteelt LH, Proefstation Groenteteelt in de Volle Grond, IVRO, IVT, IVO, CPO.
Dit tijdschrift werd intussen opgeheven.

Het gering gebruik van de Bibliography of Agriculture is opvallend: slechts 6x genoemd. (Centrum Techniek, PUDOC, Bibl. LH, IB Haren, Prov.Dir.Bedr.Ontw. Drente, Unilever)

Van de 10 meest genoemde referaattijdschriften zijn er vijf van de CAB, van de 20 meest genoemde elf.

1.2

72% van de respondenten produceren zelf secundaire informatie.(n = 96)

1.2.1

Hieraan worden gemiddeld per maand 92 manuren besteed.(n = 65 S = 211)
Dit getal ligt hoger voor de Instituten en wat lager voor de onderwijsinstellingen, met name de LH.
Er is een negatieve correlatie met het %-dekking van aanwezige primaire door aanwezige secundaire bronnen (vraag 2.3), hetgeen zou inhouden dat er relatief meer manuren worden besteed aan het creëren van secundaire informatie als deze informatie niet afdoende van elders kan worden betrokken.
De correlatiecoëfficiënt met het aantal verwerkte tijdschriftartikelen is vooral hoog (r = 0.54)

1.2.2

Het gemiddeld aantal verwerkte titels bedraagt per maand voor:

tijdschriftartikelen	126	(n = 56, S = 144)
boeken	10	(n = 46, S = 10)
rapporten	13	(n = 47, S = 23)
diversen	14	(n = 45, S = 22)

Gezien de grote spreiding hebben deze gemiddelden weinig waarde.

1.2.3

Van 65 respondenten zouden er 36 doorgaan met het zelf produceren van secundaire informatie, indien deze op dezelfde wijze geselecteerd en gegroepeerd van elders zou kunnen worden betrokken.
Met name Gezondheidsdiensten zouden er liever mee ophouden, terwijl een meerderheid van de LH-afdelingen zich voor doorgaan uitspreekt. Uit de factoranalyse blijkt dat ook de proefstations liever doorgaan met het zelf produceren van secundaire informatie ook al zou een externe service beschikbaar zijn.
Zij die willen doorgaan met eigen secundaire informatie te produceren, ook als zij die ook van elders kunnen betrekken, zullen zich ook niet door een SDI-service van elders laten beïnvloeden. (vraag 1.3.2)
Zij die op deze wijze willen doorgaan besteden relatief weinig tijd aan het zoeken in secundaire bronnen t.b.v. de eigen instelling (vraag 1.4), maar rel. veel aan het zoeken t.b.v. derden (vraag 1.3), terwijl onderzoekers bij deze instellingen relatief weinig tijd besteden aan het zoeken in primaire bronnen.
Het aantal interne verzoeken on literatuur hier is relatief gering, maar het aantal externe verzoeken relatief groot (vraag 3).
Deze instellingen verwerkten een relatief grote hoeveelheid tijdschriftartikelen en rapporten per maand.

1.3

De tijd besteed aan het zoeken in secundaire bronnen ten behoeve van derden, bedroeg gemiddeld 20 uur per maand (69 respondenten; S = 65).
Deze tijd correleerde positief met het gebruik van halfmechanische hulpmiddelen hierbij (vraag 1.3.1), evenals met de tijd gebruikt door documentalisten voor het zoeken in primaire bronnen (vraag 2.1) en ook zeer sterk met het aantal externe verzoeken om literatuurinlichtingen (vraag 3).
Deze tijd correleerde ook zeer sterk positief met het aantal verwerkte titels voor het eigen documentatiesysteem.
Gelukkig besteden de instellingen, die veel tijd investeren in het bijhouden van een eigen systeem, dus ook veel tijd in het gebruiken van dat systeem voor derden. Dit geldt niet zozeer voor het zoeken t.b.v. de eigen instelling (zie vraag 1.4).

1.3.1

Slechts in 9% van de gevallen (95) werd gebruik gemaakt van half-mechanische hulpmiddelen bij het zoeken van secundaire informatie.
Dit zijn vooral instellingen die een groot aantal tijdschriftartikelen en rapporten verwerken (1.2.2); waar de tijd gebruikt voor het zoeken t.b.v. derden relatief hoog is (1.3); waar de tijd die individuele onderzoekers besteden aan het zoeken in secundaire bronnen relatief laag is (1.4) en waar het % primaire literatuur gedekt door secundaire bronnen, relatief hoog is (2.3). Het betreft hier een relatief groot aantal instellingen uit het bedrijfsleven.

1.3.2

Indien voor maximaal f 1000,-- een SDI-abonnement mogelijk was, dan zouden 29 van de 58 antwoordende instellingen het onder 1.3 vermelde werk blijven verrichten: precies de helft.
Overschakelen op een externe SDI-service werd vooral genoemd door instellingen waar onderzoekers relatief veel tijd besteden aan het raadplegen van primaire bronnen (2.2); waar het aantal interne verzoeken om te raadplegen literatuur groot is (3), terwijl het vaker instellingen betreft waar veel fotocopieën e.d. worden gemaakt voor het verstrekken van primaire literatuur (4.2). Ook de correlatie met het aantal wetenschappelijke werkers is positief.
SDI-services zouden echter weinig van het bestaande werk overnemen bij instellingen waar veel tijdschriften worden verwerkt voor eigen systemen (1.2.2) en waar men ook zou doorgaan met het produceren van eigen secundaire informatie, indien deze in enige andere vorm extern zou kunnen worden verkregen (1.2.3).
Ook betreft het een groter aantal instellingen, die zelf primaire informatie beschikbaar stellen aan gebruikers (4).

Uit de factoranalyse kristalliseert ook de positieve relatie uit met het zelf produceren van secundaire informatie. (De correlatiecoëfficiënt was slechts + 0.13, maar deze kon zijn samengesteld uit tegengestelde partiële correlatiecoëfficiënten. De factorladingen op factor IX waren resp. 0.77 voor "produceert u zelf sec. info.", 0.78 voor "verschaft u zelf prim. info op aanvraag" en 0.48 voor "gaat u met eigen systeem door indien er een SDI beschikbaar is". Een geautomatiseerde SDI-service maakt vooral een kans bij gebruikers van Chemical Abstracts, Analytical Abstracts en Horticultural Abstracts, terwijl gebruikers van Tropical Abstracts er eerder geen behoefte aan hebben.

1.4
De tijd, ten eigen behoeve besteed door individuele onderzoekers aan het zoeken in secundaire bronnen, is gemiddeld 41.6 manuur per maand (n = 76, S = 68). Een hoog aantal correleerde positief met het eventueel overstappen op het gebruik van externe secundaire bronnen, indien mogelijk (1.2.3). Het gebruik van de Annual Review of Biochemistry, Advances in Agronomy, Bibliography of Agriculture, Herbage Abstracts was ook positief met dit aantal manuren gecorreleerd.
Tenslotte was het aantal relatief groter voor de instituten dan voor andere instellingen.

2
Primaire bronnen worden haast overal geraadpleegd, alleen bij Ligtermoet Chemie niet (?)

2.1
Bibliotheek- en Documentatiepersoneel besteedt gemiddeld 20 manuren per maand aan het raadplegen van primaire literatuur (buiten de tijd nodig voor het refereren t.b.v. het eigen systeem). (n = 80, S = 35.7)
Er zijn hier positieve correlaties met instituten en andere onderzoeksinstellingen, terwijl hieraan bij voorlichtingsinstellingen relatief weinig tijd wordt besteed. Een relatief groot aantal uren wordt hieraan besteed bij instellingen die Nutrition Abstracts en Analytical Abstracts gebruiken. Er is ook een positieve correlatie met het aantal manuren besteed aan het zoeken in secundaire bronnen t.b.v. derden (1.3) en met het aantal interne literatuurinlichtingen.

2.2
Individuele onderzoekers besteden gemiddeld 141 manuren per maand per instituut aan het raadplegen van primaire bronnen (n = 82, S = 170).
Er is een negatieve correlatie van deze grootheid met het doorgaan met de produktie van eigen secundaire systemen (vraag 1.2.3), speciaal als er een gemechaniseerde

SDI-service beschikbaar zou zijn (1.3.2).
Landbouwdocumentatie, Annual Review of Biochemistry, Analytical Abstracts, Herbage Abstracts en Current Abstracts worden vooral veel gebruikt op instellingen waar ook dit aantal manuren hoog ligt. Het betreft hier relatief veel instituten en weinig Gezondheidsdiensten en Diversen.

2.3
Gemiddeld wordt 61% van de artikelen in aanwezige primaire tijdschriften gedekt door in de instelling aanwezige secundaire tijdschriften (n = 92; S = 29).
Dit percentage ligt hoger voor instellingen die veel proceedings, advances etc. verwerken. (Eigenlijk tegengesteld aan verwachting.)
Het is eveneens hoger voor instellingen waar een relatief hoog aantal manuren door individuele onderzoekers aan het raadplegen van primaire bronnen wordt besteed (2.2).
Het ligt lager bij instellingen die veel tijd besteden aan het zelf creëren van secundaire systemen (1.2).
(De vraag specificeerde niet of onder de bedoelde secundaire bronnen ook het eventuele eigen systeem moest worden begrepen. Waarschijnlijk hebben de respondenten 'secundaire bronnen' hier als 'externe secundaire bronnen' opgevat.)
Een grote boekenaanwas gaat ook samen met een slechtere dekking.
Proefstations en Voorlichtingsdiensten vinden hun aanwezige primaire literatuur relatief weinig door hun secundaire bronnen gedekt, in tegenstelling tot onderwijsinstellingen.

3a
Het aantal interne verzoeken on inlichtingen betreffende te raadplegen literatuur per maand is gemiddeld 75 (n = 76; S = 178).
Dit aantal is positief gecorreleerd met het aantal gebruikte secundaire bronnen (1.1), het aantal verwerkte boeken en vooral rapporten (1.2.2) (corr. coeff. = 0.80), het aantal manuren dat het bibliotheek- en documentatiepersoneel alsook individuele onderzoekers besteden aan het raadplegen van primaire bronnen (2.1, 2.2), het aantal externe verzoeken (3b) en de produktie van kopieën (4.2).
Het betreft vooral Instituten.
Voor Routine en Voorlichtingsinstellingen is het aantal relatief gering.
Een negatieve correlatie is er ook met het doorgaan met zelf produceren van eigen secundaire systemen (1.2.3, 1.3.2).

3b
Het aantal externe verzoeken bedraagt gemiddeld 28 per maand (n = 74; S = 47).
Hier is dit aantal groter dan het gemiddelde bij een relatief groot aantal gebruikte secundaire bronnen (1.1), een groot aantal verwerkte tijdschrift-

artikelen, boeken en rapporten (1.2.2), de tijd besteed aan het zoeken in secundaire bronnen t.b.v. derden (1.3), de interne verzoeken (3a), het verschaffen van primaire informatie (4) en het gebruik van Soils and Fertilizers, Forestry Abstracts en Dairy Science Abstracts.
Het ligt voor onderwijsinstellingen relatief hoog, met name voor de Landbouwhogeschool.
Voor Gezondheidsdiensten en Routine instellingen ligt het relatief laag.

4
61% van de respondenten (97) verschaft ook de inhoud van primaire documenten op aanvraag.
Dit gebeurt vooral door instellingen die zelf secundaire informatie produceren en met name veel tijdschriften en rapporten verwerken (1.2), die niet op een SDI-service zouden overgaan (1.3.2), die veel externe verzoeken verwerken (3b) en die relatief veel gebruikmaken van Landbouwdocumentatie, Nutrition Abstracts en Forestry Abstracts.
Het betreft vooral de Instituten en weinig onderwijsinstellingen noch 'diverse' instellingen.

4.1
Voor het verstrekken van de onder 4 genoemde primaire informatie wordt door 90% van de instellingen gebruik gemaakt van reproduktie-apparatuur in eigen beheer.
Niet in eigen beheer (of helemaal niet??) gebeurt dit alleen bij de Apotheek van de Faculteit voor Diergeneeskunde, het K.N.M.I. afd. III, het L.E.I. (?) en drie LH-afdelingen.

4.2
Het aantal gereproduceerde pagina's per maand bedraagt gemiddeld 629 (n = 47; S = 1797).
Dit is positief gecorreleerd met het aantal gebruikte secundaire bronnen, het aantal per maand verwerkte 'diversen' (1.2.2), het doorgaan met het samenstellen van een eigen secundair systeem, ook indien er een SDI-service te krijgen zou zijn (1.3.2), het aantal manuren dat onderzoekers steken in het raadplegen van secundaire bronnen (1.4), als ook in het raadplegen van primaire bronnen (2.2), het aantal interne verzoeken om inlichtingen en het gebruik van Field Crop Abstracts, Plant Breeding Abstracts, Review of Applied Entomology, alsmede de Review of Plant Pathology.
De Proefstations reproduceren vooral veel (r = 0.40), de correlaties voor de andere categorieën van instellingen zijn laag.

Categorieën instellingen

De correlaties met de verschillende categorieën van instellingen zijn bij de meeste vragen reeds ter sprake gekomen.

Aantal wetenschappelijke onderzoekers

Volgens de factoranalyse is deze variabele het sterkst gerelateerd met het zelf produceren van secundaire informatie als er een SDI-service beschikbaar zou zijn (1.3.2), en wel in negatieve zin. De grotere instellingen zien dus blijkbaar eerder een belang bij een externe dienstverlening dan de kleinere. Daarnaast is er natuurlijk een sterke positieve relatie met de duidelijk kwantitatieve vragen, waarvan de tijd t.b.v. eigen instelling besteed aan het zoeken in secundaire bronnen (1.4) en de tijd, besteed door onderzoekers aan het raadplegen van primaire bronnen (2.2) het meest in het oog springen.

Bijlage 2. Verwerkingsdata van de CAIN-, BoA- en FSTA-banden bij IWIS-TNO 1972-1974.

Banddatum	CAIN	BoA	FSTA
72- 7	21/9	8/11	9/9
8	21/9	19/12	1/11
9	2/11		1/11
10	6/2		2/2
11	7/2	7/2	5/2
12	6/5		1/5
73- 1	22/5	7/5	23/5
2	8/6	4/6	8/6
3	20/6	12/6	20/6
4	2/7	20/6	2/7
5	12/7	2/7	12/7
6	18/7	12/7	18/7
7	26/7	18/7	24/7
8	2/11	30/10	13/11
9	14/11	8/11	14/11
10	22/11	16/11	22/11
11	21/1/74	7/12	4/12
12	8/3/74	17/1/74	27/2/74
74- 1	29/3		29/3
2	23/4		22/4
3	24/5		24/5
4	4/6		30/5
5	19/7		31/7
6	21/8		21/8
7	25/9		25/9
8	23/10		21/10
9	6/11		6/11

Bijlage 2. Vervolg. Verwerkingsdata van de CAIN-, FSTA-, CAB- en AGRINDEX-banden bij IWIS-TNO 1975-1976.

Banddatum	CAIN	FSTA	CAB	AGRINDEX
75- 1	8/4	8/4	25/4	20/6
2	22/4	22/4	16/5	16/7
3	10/6	10/6	15/7	30/7
4	1/7	1/7	29/7	15/8
5	30/7	30/7	15/8	15/8
6	15/8	15/8	26/8	26/8
7	4/9	29/8	26/8	26/8
8	5/11	17/9	17/9	17/9
9	18/12	18/9	8/10	8/10
10	12/2/76	5/11	17/12	8/10
11	18/3/76	26/11	13/1/76	26/11
12	8/4/76	13/1/76	14/1/76	17/12
76- 1	27/4	27/1	25/2	28/1
2	11/5	18/3	15/4	18/2
3	22/6	6/4	4/5	23/3
4	29/6	11/5	15/6	6/4
5	24/8	8/6	6/7	8/6
6	31/8	22/6	28/7	29/6
7	14/9	6/7	10/8	13/7
8	28/9	10/8	17/8	3/8
9	26/10	31/8	14/9	7/9
10	21/12	28/9	12/10	28/9
11	11/1/77	9/11	9/11	10/11
12	25/1	7/12	7/12	21/12

'CAB' betrof in 1975 alleen CABVET, in 1976 het gehele CAB-bestand.

Bijlage 3. Literatuur-interesseprofielen waarvoor de resultaten van manuele attendering werden vergeleken met die van computer-attendering. Tevens aanduiding van de gebruikte bibliografische bestanden op magneetband.

1. Invloed van fluor uit de lucht op plantengroei	CAIN/CAC/BA
2. Coccidiostatica (bestrijding van een protozoaire ziekte-verwekker, vooral van belang bij kippen	CAIN/CAC/Medlars/EM
3. Nitrovin (Het gebruik en effect van een groeibevorderings-middel bij pluimvee en andere diersoorten)	CAIN/CAC/EM
4. Vliegtuigen in de landbouw	CAIN/Compendex
5. Aardappelprodukten voor consumptie	CAIN/FSTA/CAC
6. Menselijke ziekten veroorzaakt door aardappelconsumptie	CAIN/FSTA/CAC/EM/ Medlars
7. Olie en vet in de aardappelverwerking	CAIN/FSTA/CAC
8. Bewaren van aardappelen	CAIN/FSTA/CAC
9. Eetbare paddestoelen	CAIN/FSTA
10. Effect van bestraling op chromosomen, DNA en RNA	CAIN/CAC/BA
11. Weefselcultuur	CAIN/BA
12. De teelt van spinazie	CAIN
13. De invloed van aardgas op plantengroei alsmede het ontstaan en effect van biologische gasvorming (methaan)	CAIN/CAC
14. Varkensbatterijen	CAIN
15. Economische integratie van dierlijke eenheden	CAIN
16. Dierlijke destructieprodukten (productie, gebruik)	CAIN/CAC/FSTA
17. De invloed van luchtverontreiniging op plantengroei (voor zover niet vallend onder nrs. 1, 22, 23 en 24)	CAIN/CAC
18. De invloed van natuurlijke radioactieve straling en onvrijwillig vrijgekomen straling op levende organismen	CAIN/CAC/Medlars/EM/ INIS
19. Economie van de zetmeelproduktie	CAIN/FSTA
20. Vorming van zetmeel in planten	CAIN
21. Organisatie van het landbouwkundig onderzoek	CAIN
22. Invloed van ozon op plantengroei	CAIN/CAC
23. Invloed van SO_2 op plantengroei	CAIN/CAC
24. Invloed van lood op plantengroei	CAIN/CAC
25. Zonnebloemen	CAIN/FSTA

Bijlage 4. Overlapping tussen manuele en computer-attendering van relevante* artikelen gepubliceerd in 1973 voor de 25 onderwerpen genoemd in Bijlage 3.

profielnr.	alleen m	m + c	alleen c	totaal
1	7 (11)	14 (22)	42 (67)	63 (100)
2	10 (6)	15 (8)	157 (86)	182
3	9 (11)	4 (5)	71 (85)	84
4	42 (26)	14 (9)	106 (65)	162
5	14 (15)	14 (15)	67 (71)	95
6	5 (7)	4 (6)	58 (87)	67
7	7 (26)	3 (11)	17 (63)	27
8	13 (9)	16 (11)	114 (80)	143
9	44 (19)	37 (16)	151 (65)	232
10	28 (7)	35 (9)	347 (84)	410
11	231 (36)	149 (23)	266 (41)	647
12	12 (30)	14 (35)	14 (35)	40
13	9 (33)	0 (0)	18 (67)	27
14	36 (34)	4 (4)	66 (62)	106
15	29 (31)	1 (1)	65 (68)	95
16	60 (16)	41 (11)	270 (73)	371
17	76 (19)	23 (6)	299 (75)	398
18	47 (6)	8 (1)	755 (93)	810
19	26 (52)	0 (0)	24 (48)	50
20	4 (11)	1 (3)	31 (86)	36
21	118 (81)	0 (0)	27 (19)	145
22	21 (35)	16 (27)	23 (38)	60
23	31 (41)	15 (20)	30 (39)	76
24	23 (25)	10 (11)	58 (64)	91
25	37 (14)	32 (12)	197 (74)	266
totaal	939 (20)	470 (10)	3274 (70)	4683 (100)
	(24)**	(11)**	(66)**	

m = aantal titels gevonden door de manuele attendering,
c = idem, gevonden door de computer-attendering,
m + c = idem, gevonden door beide systemen.
Tussen haakjes de percentages, waarbij het totaal aantal gevonden titels op 100% is gesteld.

* Inclusief 'misschien relevante' titels, waarbij de relevantie moeilijk aan de titel was te beoordelen.
** Dit is het gemiddelde percentage berekend door de percentages van de 25 profielen te middelen.

Bijlage 5. Overlapping tussen manuele en computer-attendering exclusief misschien relevante titels. Zie Bijlage 4.

profielnr.	alleen m	m + c	alleen c	totaal
1	7 (23)	14 (47)	9 (30)	30 (100)
2	10 (7)	15 (10)	121 (83)	146
3	9 (17)	4 (8)	40 (75)	53
4	42 (32)	14 (11)	74 (57)	130
5	14 (18)	14 (18)	48 (63)	76
6	5 (10)	4 (8)	43 (83)	52
7	13 (10)	16 (13)	97 (77)	126
8	7 (32)	3 (14)	12 (54)	22
9	44 (20)	37 (17)	135 (63)	216
10	28 (9)	35 (11)	244 (80)	307
11	231 (37)	149 (24)	238 (39)	618
12	12 (34)	14 (40)	9 (26)	35
13	9 (47)	0 (0)	12 (57)	21
14	36 (61)	4 (7)	19 (32)	59
15	29 (50)	1 (1)	28 (48)	58
16	60 (24)	41 (17)	145 (59)	246
17	76 (31)	23 (9)	143 (59)	242
18	47 (8)	8 (1)	526 (91)	581
19	26 (72)	0 (0)	10 (28)	36
20	4 (24)	1 (6)	12 (70)	17
21	118 (89)	0 (0)	14 (11)	132
22	21 (42)	16 (32)	13 (26)	50
23	31 (43)	15 (21)	26 (36)	72
24	23 (56)	10 (24)	8 (20)	41
25	37 (14)	32 (12)	197 (74)	266
	939 (26)	470 (13)	2223 (61)	3632 (100)
	(32)	(14)	(54)	

Tussen haakjes de percentages, waarbij het totaal aantal gevonden titels op 100 is gesteld.

Bijlage 6. Lijst van profielonderwerpen die voor het ALADIN-project werden verwerkt en beoordeeld door de gebruikers.

Geslachtsbeïnvloeding bij planten
Aromacomponenten in voedingsmiddelen
Aminozuren en landbouwhuisdieren
Invloed van de voeding op de samenstelling van het bloed bij landbouwhuisdieren
Eierconsumptie, cholesterol en bloedsamenstelling
Analyse van vruchtensappen
Aminozuurabsorptie in de darm
Perinatale mortaliteit bij schapen
Nieuwe produkten uit stro
Beschikbaarheid van aminozuren
Apicale dominantie bij kweek
Fysische eigenschappen van biologische materialen
Phytochroom
Carboxylaten en ionenopname
Zink en cadmium in de grond
Langzaam werkende meststoffen
Waterbinding bij levensmiddelen
Eiwittoxiciteit
Eiwitconcentraten
Kenya
Selectie-index theorie (landbouwhuisdieren)
Melk/vlees relatie
Eiwitdepositie bij varkens
Genenbanken
Bodem en landschap
Furaanderivaten
Schimmelziekten bij tarwe
Mechanisatie suikerbieten
Stikstof in de microbiologie
Economie van de eiwitvoorziening
Antagonisten van Rhizoctonia solanum
Simulatie
Effect van fall-out en organische chloorverbindingen op dieren
Houtafval
Populieren
Fosfaatbehoefte van landbouwhuisdieren
Bereiding van eiwitconcentraten
Niet-eiwit stikstof als voeding voor herkauwers
Pectolitische enzymen
Proteïneafbraakprodukten
Hemicellulase
Troebelingsstabiliteit in vruchtensappen
Kattekleien
Eigenschappen van aggregaten
Bastvezels
Droging als scheidingsmethode
Poriënstructuur en absorptieverschijnselen
Koffie, thee en tabak (technologie)
Invloed van stikstof op de droge stof produktie
Regulatie lebmaagactiviteit
Ionentransport door de grond naar de wortel
Oliën en eiwitten uit zaden
Smaak van eiwitten
Nitriettoxiciteit bij planten
Voederopname
Invloed van leeftijd en seizoenen op de melkproduktie
Invriezen van berensperma
Septoria
Verwerking van groenvoeders
Fysiologie van dwergtarwe
Infectieziekten van de vrouwelijke geslachtsorganen bij varkens
Gebruik van tuinbonen
Vietnam
Besliskunde
Landbouw als deelberoep
Ziekte van Marek
Ademhalingsziekten van honden
Invloed van de voedingsstatus op fysieke prestaties

Bijlage 7. Lijst van instellingen resp. vakgroepen waarvan medewerkers hebben deelgenomen aan het ALADIN-project.

Centraal Instituut voor Voedingsonderzoek - TNO (CIVO)
Centrum voor Landbouwpublikatie en Landbouwdocumentatie (Pudoc)
Douwe Egberts Research
Instituut voor Bewaring en Verwerking van Landbouwprodukten (IBVL)
Instituut voor Bodemvruchtbaarheid (IB)
Instituut de Haaf (Stichting voor Plantenveredeling - SVP)
Instituut voor Pluimveeonderzoek 'het Spelderholt' (IPS)
Instituut voor Toepassing van Atoomenergie in de Landbouw (ITAL)
Instituut voor Veeteeltkundig Onderzoek 'Schoonoord' (IVO)
Instituut voor Veevoedings Onderzoek 'Hoorn' (IVVO)
Instituut voor de Veredeling van Tuinbouwgewassen (IVT)
Koninklijk Instituut voor de Tropen (KIT)
Landbouwhogeschool - Vakgroep Bodemkunde
- Vakgroep Fysiologie der Dieren
- Vakgroep Fytopathologie
- Vakgroep Landbouwplantenteelt
- Vakgroep Landbouwscheikunde
- Vakgroep Levensmiddelentechnologie
- Vakgroep Microbiologie
- Vakgroep Plantenfysiologisch Onderzoek
- Vakgroep Veeteelt
- Werkgroep Viet-Nam

Ministerie van Landbouw en Visserij, Directie Algemene Zaken, Afd. Economische Verdedigingsvoorbereiding
Ministerie van Landbouw en Visserij, Directie Landbouwkundig Onderzoek
Pedagogische Hogeschool in Bonn (BRD)
Philips-Duphar
Rijksuniversiteit Utrecht, Veterinaire Faculteit, Veterinair Pathologisch Instituut
Rijksuniversiteit Utrecht, Veterinaire Faculteit, Veterinaire Verloskunde
Sprenger Instituut
Staatsbosbeheer
Stichting Industriehout
Werkgroep Mineralen in Krachtvoer in relatie tot Bemesting en Milieu
Werkgroep Voederopname Herkauwers



Centrum voor Landbouwpublikaties en Landbouwdocumentatie

ALADIN PROJEKT

Aan:

Datum:

Bijgaand:

Graag ontvang ik de duplikaat-output zo spoedig mogelijk van u terug, met een aanduiding voor de relevantie en nieuwswaarde:

++ zeer relevant
\+ relevant
± randgebied
relevantie alleen aan de titel niet te beoordelen
– irrelevant
n nog niet eerder gezien

Marijkeweg 17, Wageningen, Postbus 4 (08370) 1 91 46

Bijlage 9. Enquêteformulier betreffende (A) financiële waardering van computer-attendering en (B) betreffende informatie per titel van een artikel.

Centrum voor Landbouwpublikaties en Landbouwdocumentatie
Marijkeweg 17, Postbus 4, Wageningen, Tel. (08370) 19146

pudoc

Aan de leden van de experimentele
gebruikersgroep 'ALADIN'

Nu U reeds gedurende enige tijd door ons op voor U hopelijk relevante literatuur wordt geattendeerd door middel van met behulp van de computer geselecteerde titels, zou ik graag een nadere indruk krijgen van uw waardering van het geleverde materiaal uitgedrukt in een mogelijk financieel offer. Het ligt daarbij in de bedoeling om tegen de helft van dit jaar over te gaan tot het vragen van een daadwerkelijke financiële bijdrage, aangezien het aantal nieuwe aanvragen onze huidige financiële capaciteit te boven dreigt te gaan. Tevens is met het oog op een mogelijk operationeel worden van deze dienstverlening na afloop van de proefperiode in april 1975 psychologisch niet juist om de gebruikers gedurende een zo lange tijd onze diensten gratis aan te bieden, waarna dan plotseling een fors bedrag voor continuering moet worden gevraagd.

Onderstreept moet echter worden, dat het antwoord op de vraag ten aanzien van een financiële waardering c.q. eventueel te brengen offer geenszins inhoudt dat U ook voor dat bedrag zult worden aangeslagen.
De te vragen prijs zal hoofdzakelijk worden bepaald door de kostprijs en de door het Ministerie van Landbouw en Visserij te verlenen subsidie. De vraag is hoofdzakelijk bedoeld als meting van Uw waardering voor de geleverde diensten.

Tevens maak ik van de gelegenheid gebruik om U een aantal vragen te stellen over de inhoud van de output - zodat in de toekomst wellicht beter met Uw wensen rekening kan worden gehouden.

Het A-formulier met de financiële vraag moet ingevuld worden voor iedere data base die U ontvangt. Het B-formulier maar éénmaal.

Hopende op Uw medewerking verblijf ik,

met de meeste hoogachting

Ir. H.C. Molster
ALADIN-projectleider

ALADIN-Enquête maart 1974

A

Data Base: Profiel nr.:

A-1 Een jaarabonnement ALADIN-attendering is mij het volgende bedrag waard:

f 0	ja/nee
f 50	ja/nee
f 100	ja/nee
f 200	ja/nee
f 300	ja/nee
f 400	ja/nee
f 500	ja/nee
f 750	ja/nee
f 1000	ja/nee

A-2 Indien noodzakelijk, zal ik in staat zijn per 1-6-1974 de helft van een dergelijk bedrag voor een halfjaar abonnement ook daadwerkelijk te (doen) betalen.

juist/onjuist

Toelichting:

ALADIN-Enquête maart 1974

B

Profiel nr.:

B-1 Is het van belang dat de titel in de oorspronkelijke taal wordt afgedrukt, of is de vertaling in het Engels door de indexerende instantie voldoende?

Oorspronkelijke taal is voor mij

0 essentieel
0 erg handig maar niet essentieel
0 onbelangrijk

(juiste antwoord aankruisen)

B-2 Is voor U het afdrukken van een abstract - meestal op 2 kaartjes -

0 essentieel
0 erg handig maar niet essentieel
0 onbelangrijk

B-3 Is het voor U van belang dat bij de titel ook de affiliatie van de auteur(s) (hun instituut etc.) wordt afgedrukt?

Het afdrukken van de auteurs affiliatie is voor mij

0 essentieel
0 erg handig maar niet essentieel
0 onbelangrijk

B-4 Is het voor U van belang dat monografieën (boeken, onregelmatige publikaties, niet zijnde tijdschriften reeksen noch tijdschriftartikelen) in de data base zijn opgenomen?

Monografieën zijn voor mij

0 erg belangrijk
0 soms van belang maar niet essentieel
0 onbelangrijk

Verdere opmerkingen:

Naam :

Functie :

Leeftijd :

Naam Instelling:

Type Instelling:
- Universiteit/Hogeschool
- Rijks Onderzoeksinstelling
- Industrie
- Anders

Profiel: Tape: Kosten*:

* In geval van betalend abonnement.

Redenen voor discontinuering

A. ONBEVREDIGENDE RESULTATEN

- teveel titels die niet relevant waren
- te weinig relevante titels
- ik kreeg veel meer titels dan ik kan verwerken
- van veel titels kan ik de relevantie niet direct beoordelen
- veel titels leken op het eerste gezicht relevant, maar bleken dit toch niet te zijn
- veel titels waren wel relevant, maar het materiaal was van geringe kwaliteit
- veel relevante titels betrof materiaal dat moeilijk verkrijgbaar bleek
- veel titels betroffen artikelen in talen die ik niet kan lezen (Russich, Japans)

Andere redenen, opmerkingen:

B. FINANCIELE REDENEN

- geen fondsen voor dit soort uitgaven beschikbaar
- theoretisch wel fondsen beschikbaar, maar die worden beter ergens anders voor gebruikt
- indien de prijs verlaagd werd tot *f* zou ik wel een abonnement nemen

Andre redenen, opmerkingen:

C. GEEN PLAATS VOOR DIT SOORT DIENSTVERLENING

- de onderzoeksfase, waarin ik literatuurattendering nodig had, was voorbij
- de meeste relevante titels kwam ik reeds langs andere weg tegen

  de attendering is geen reden om deze andere wegen minder te gebruiken

  de belangrijkste wegen, waarlangs ik nu op de hoogte kwam van deze titels, zijn:

  - circulerende leesmap
  - primaire tijdschriften regelmatig geraadpleegd in bibliotheek
  - regelmatig doorkijken van referaattijdschriften
  - contact met collega's in Nederland
  - contact met collega's in het buitenland
  - anders:

- Er waren weliswaar allerlei relevante titels bij die ik niet vanzelf langs andere weg zou hebben gekregen, maar de aard van mijn werk maakt het moeilijk om van dit materiaal gebruik te maken.

  Toelichting:

Opmerkingen:

D. ALGEMENE OPMERKINGEN

Welke elementen in uw huidige informatievoorziening beschouwt u als voldoende?

E. ALGEMENE OPMERKINGEN t.a.v. ALADIN

Centrum voor Landbouwpublikaties en Landbouwdocumentatie
Marijkeweg 17, Postbus 4, Wageningen, Tel. (08370) 19146

pudoc

AANVRAAGFORMULIER VOOR EEN LITERATUURONDERZOEK

Opdrachtgever

1. Naam:
   Adres: Tel:
2. Instelling/Vakgroep:
   Adres: Tel:

*Aard van de opdracht**

☐ Retrospectief literatuuronderzoek
☐ Manuele attendering op nieuw verschenen primaire literatuur
☐ Computerattendering op nieuw verschenen gerefereerde literatuur
☐ Literatuurrapport

*Doel van het literatuuronderzoek**

☐ Nieuw onderzoek
☐ Reeds gevorderd onderzoek
☐ Onderwijs
☐ Publikatie/rapport
☐ Ontwikkelingsproject
☐ Beleidsontwikkeling
☐ Andere:.....
.....

Omschrijving van het onderwerp

Zoekdefinities

Relevante trefwoorden:

Uit te sluiten trefwoorden:

Reeds bekende zeer relevante literatuur (titel, auteur, tijdschrift)
U wordt verzocht zoveel mogelijk literatuuropgaven (indien mogelijk 10 of meer) op een afzonderlijk vel bij te voegen.

NAVOLGENDE ALLEEN TE BEANTWOORDEN VOOR RETROSPECTIEF ONDERZOEK

Beperkingen
- tot jaren: 19.. t/m 19 ..
- tot landen:
- tot talen: ned/eng/frans/duits/russisch/overige met eng. summary/geen beperking**
- tot maximaal aantal titels: 50/100/200/300/400/600**
- tot technieken en apparatuur: ja / nee**
- tot: zeer relevante publikaties / volledigheid**
- wilt u informatie over patenten of kwekersrechten? ja / nee**
- wilt u informatie over lopend onderzoek? ja / nee**

Uiterste termijn voor uitvoering van de opdracht
- weken na dagtekening: 1/2/3/4/5/6/8/10**

Reeds geraadpleegde literatuurbronnen welke bij het onderzoek kunnen worden uitgesloten.
Referaattijdschriften:

Primaire tijdschriften:

Alleen voor attendering
Ik wil per keer niet meer titels ontvangen dan 10/25/50/100/250**

* aankruisen wat van toepassing is
** doorstrepen wat niet van toepassing is

DATUM: HANDTEKENING:

TOELICHTING

RETROSPECTIEF LITERATUURONDERZOEK EN ATTENDERING

Het doel van een retrospectief onderzoek is anders van aard dan een abonnement op de attenderingsdienst; in het ideale geval zou men, voor het onderzoek van start gaat, met een retrospectief literatuuronderzoek moeten beginnen om daarna met behulp van de attenderingsdienst op de hoogte te blijven van nieuwe ontwikkelingen op het vakgebied. Retrospectief onderzoek kan eveneens nuttig zijn om informatie op te sporen betreffende speciale aspecten en problemen welke gedurende het onderzoek naar voren komen, en kan nodig zijn om op korte termijn literatuur te verzamelen over een specifiek onderwerp.

De attendering op nieuw verschijnende literatuur is speciaal bedoeld om tijd te besparen op het doornemen van vaktijdschriften; het is uiteraard wel noodzakelijk de belangrijkste primaire tijdschriften zelf te blijven lezen, echter, belangrijke publikaties die elders verschijnen kunnen eenvoudiger worden gevonden door een selectief attenderingssysteem, dat u maandelijks van relevante titels voorziet.

De opzet van een attenderingsprofiel is over het algemeen breder dan dat van een retrospectief onderzoek, omdat de belangstelling tijdens het onderzoek eveneens gericht zal zijn op de randgebieden, terwijl voor het 'terugzoeken in de tijd' het onderwerp meestal specifieker gericht is. Ook kan men over langere tijd meer literatuur verwerken dan in een korte periode. 10 titels per maand zijn gemakkelijker door te nemen dan 600 titels, wat het resultaat zou zijn van een retrospecitief onderzoek over 5 jaar met dezelfde dekking.

Retrospectief literatuuronderzoek wordt uitgevoerd met de best beschikbare middelen, wat impliceert dat zowel bronnen in gedrukte vorm als met computer doorzoekbare bibliografieën worden doorzocht. De recente literatuur wordt over het algemeen eerst met behulp van de computer doorzocht, de oudere literatuur (vóór 1970) zal vaak met de hand moeten worden gezocht omdat deze meestal niet in computerleesbare vorm beschikbaar is. Op veel gebieden zijn echter goede referaattijdschriften voor handen, wat het gebruik van de computer voor literatuuronderzoek niet noodzakelijk maakt.

ATTENDERING:MANUEEL EN PER COMPUTER

Snelheid Aangezien de manuele attenderingsdienst werkt met de (primaire) tijdschriften welke binnenkomen bij de Centrale Bibliotheek van de Landbouwhogeschool, is deze attendering over het algemeen sneller dan de computerattendering, welke wordt uitgevoerd met de magneetbanden van referaattijdschriften. De computerattendering is meestal aanzienlijk sneller dan de gedrukte versie van de referaattijdschriften.

Precisie Wanneer onze informatiespecialist zich een goed beeld heeft gevormd van uw onderzoek, zal hij alleen artikelen selecteren welke werkelijk relevant zijn. Het is mogelijk dat u titels ontvangt welke op het eerste gezicht niet relevant lijken maar, omdat geselecteerd wordt op basis van het gehele artikel, mag u verwachten dat zij toch van belang kunnen zijn voor uw onderzoek.

De computer is meestal minder precies: alleen titels en trefwoorden worden doorzocht, waardoor de computer volledig afhankelijk is van de informatie die daarin verstrekt wordt. Door ongewenste, maar niet te vermijden trefwoordcombinaties zullen er titels geselecteerd worden die niet relevant zijn. Over het algemeen wordt met de profielopstelling gestreefd naar 50% relevantie. Een hoger percentage kan wel worden bereikt, maar men loopt dan het risico een aantal van belang zijnde titels te missen. Een lager relevantiepercentage kan acceptabel zijn wanneer er weinig over het onderwerp wordt gepubliceerd.

Dekking Het is uiteraard onmogelijk om meer literatuur te vinden dan er beschikbaar is in het bestand waaruit de selectie wordt gemaakt. De manuele attendering selecteert uit de tijdschriftabonnementen van de Centrale Bibliotheek van de Landbouwhogeschool, welke een van de grootste landbouwkundige collecties ter wereld bevat. In de randgebieden van de landbouw is deze echter niet kompleet. Anderszijds is het de meest volledige literatuurverzameling van in Nederland gepubliceerde literatuur.

De beste dekking zal worden bereikt door de combinatie van manuele en computer attendering. De laatste geeft in het algemeen een betere dekking dan de eerste, hoewel dat afhankelijk kan zijn van het onderwerp.

Voor elke vraag moet worden gezocht naar de optimale mogelijkheden, welke wij trachten te bereiken in nauwe samenwerking met de opdrachtgever.

Beschikbaarheid van het origineel De originele artikelen van de door de manuele attenderingdienst gevonden titels zijn steeds in de LH-bibliotheek beschikbaar. Tegen betaling van 35 cent per pagina kan van alle uitgezochte artikelen direct een kopie van het origineel worden geleverd. Voor de computer-attendering en het retrospectief onderzoek kan beschikbaarheid in de LH-bibliotheek niet worden gegarandeerd, echter titels uit de Bibliography of Agriculture (CAIN) zijn in ieder geval aanwezig in de NAL-bibliotheek, waarvoor aanvraagkaarten beschikbaar zijn bij de LH-bibliotheek.

OPMERKING

De computer selecteert titels op basis van een trefwoord of combinatie daarvan, zodat titels die niet op duidelijke wijze het artikel karakterizeren, meestal niet door de computer zullen worden gevonden. Dat is een van de redenen waarom zoveel mogelijk synoniemen en verwante zoektermen bij de aanvraag moeten worden toegevoegd. Ook een lijst van reeds gevonden relevante titels is voor ons een van de belangrijkste hulpmiddelen bij het formuleren van een zoekprogramma.

Centrum voor Landbouwpublikaties en Landbouwdocumentatie
Marijkeweg 17, Postbus 4, Wageningen, Tel. (08370) 19146

pudoc

TARIEVEN LITERATUURONDERZOEK (per 1-1-77)

MANUELE ATTENDERING

Kosten *f* 350,-- per jaar plus *f* 1,25 per geleverde titel. Fotokopieën *f* 0,35 per pagina (alleen bij abonnement).

COMPUTERATTENDERING

GROEP I Voor een abonnement op de databases van groep I zijn de basiskosten per jaar *f* 300,--. Daarboven komen de kosten per doorzochte database.

- CAIN (Bibliography of Agriculture)	*f* 200,--
- AGRINDEX	*f* 175,--
- CAB (Commonwealth Agricultural Bureaux)	*f* 350,--
- F.S.T.A. (Food Science & Technology Abstracts)	*f* 100,--
- Dierkunde Sectie uit CAIN of AGRINDEX	*f* 50,--
- Dierkunde Sectie uit CAB	*f* 80,--

GROEP II Voor de databases van groep II worden geen basiskosten berekend en geldt als prijs per jaar (korting *niet* mogelijk):

- CHEMICAL ABSTRACTS (abonnement alleen mogelijk voor instellingen die vallen onder het Ministerie van Landbouw en Visserij. Anderen dienen zich direct te wenden tot de NOCI, Burnierstraat 1, Den Haag)	*f* 400,--
- MEDLARS (Index Medicus)	*f* 225,--
- BIOSIS (Biological Abstracts)	*f* 250,--

Prijzen van andere databases zijn op aanvraag verkrijgbaar. De prijzen gelden per kalenderjaar en zijn exclusief B.T.W.

Gratis kennismaking De eerste drie zendingen van de computerattendering zijn ter kennismaking altijd gratis.

Meerdere kopieën Het is mogelijk van een profiel meerdere kopieën te ontvangen. Indien verschillende deelnemers zich op hetzelfde profiel abonneren wordt ook korting verleend.

Aanvang en beëindiging van het abonnement De opdrachten kunnen ieder moment ingaan en worden ingetrokken, waarbij verrekening naar verleende diensten plaats vindt, met inachtneming van daarvoor geldende bepalingen die per database kunnen verschillen.

Gebruik van meer databases Indien van de databases van groep I meer dan één wordt gebruikt voor hetzelfde profiel, worden de basiskosten slechts éénmaal in rekening gebracht. Een abonnement op zowel CAIN als FSTA kost zodoende niet *f* 900,--, maar *f* 600,--.

De 3 databases welke voor het landbouwkundig onderzoek de meeste waarde hebben zijn: CAIN, AGRINDEX en CAB.

CAIN
CAIN is in gedrukte vorm bekend als de Bibliography of Agriculture. De database bestaat uit de meeste monografieën en artikelen in tijdschriften die ontvangen worden door de National Agricultural Library van het U.S. Dept. of Agriculture in Beltsville, USA. Per jaar komen er ongeveer 140.000 dokumenten bij. De belangrijkste wereldliteratuur op het gebied van de landbouw wordt in het algemeen gedekt. CAIN is een titelbibliografie.

CAB (Commonwealth Agricultural Bureaux)
De CAB database bevat al het materiaal uit de volgende referaattijdschriften:

WAERSA	Soils and Fertilizers
Nutrition Abstracts	Herbage Abstracts
Animal Breeding Abstracts	Fieldcrop Abstracts
Dairy Science Abstracts	Plant Breeding Abstracts
Veterinary Bulletin	Weed Abstracts
Index Veterinarius	Review of applied Entomology (A)
Forestry Abstracts	Review of Plant Pathology
Horticultural Abstracts	Helminthological Abstracts (quarterly)

De CAB-titels hebben bijna allemaal referaten. De intellectuele arbeid kost tijd, waardoor dezelfde titel in CAIN in het algemeen sneller zal verschijnen dan in CAB.

AGRINDEX
AGRINDEX is een "pilot project" voor een wereldwijde landbouwbibliografie van de Food and Agricultural Organization. Als zodanig heeft het echter reeds een belangrijke omvang gekregen, welke die van CAIN en CAB snel benadert.
Voorlopig is de samenstelling van AGRINDEX relatief onevenwichtig, zeker ten opzichte van CAIN en CAB. Dit hangt samen met de gedecentraliseerde invoerorganisatie voor AGRINDEX tegenover de zeer gecentraliseerde invoer bij CAIN; bij CAB vindt de invoer plaats door gespecialiseerde bureaux, die allemaal in het Verenigde Koninkrijk zijn gevestigd. Bij AGRINDEX is ieder land (of regio) verantwoordelijk voor de invoer van alle publikaties die binnen haar gebied worden uitgegeven.

RETROSPECTIEF ONDERZOEK

Voor de behandeling van snel te beantwoorden literatuurvragen worden slechts administratiekosten berekend. Voor meer tijd vergende opdrachten wordt ƒ 250,-- (excl. BTW) berekend. Voor opdrachtgevers, die persoonlijk aanwezig willen zijn bij de behandeling van hun vraag met de computer, kan desgewenst een tarief per minuut worden gehanteerd (voor de meeste databases ƒ 7,50).
De voor het landbouwkundig onderzoek belangrijkste databases welke worden doorzocht met de computer zijn:
- CAIN (Bibliography of Agriculture)
- CAB (zie computerattendering)
- BIOSIS (Biological Abstracts)
- Chemical Abstracts

Het aantal beschikbare databases neemt maandelijks toe en bedraagt nu ongeveer 45.

SPECIAAL TARIEF

Voor instellingen die resorteren onder het Ministerie van Landbouw en Visserij zijn de tarieven 50% lager (behalve voor de databases van groep II bij computerattendering).