

AUTOMATISERING IN BIBLIOTHEEK EN INFORMATIECENTRUM

door P. SCHERER-GOOSSENS
N.C.W.T.D., Koninklijke Bibliotheek Albert I

INLEIDING

De functies, vatbaar voor automatisering in bibliotheek en informatiecentrum, zijn zeer talrijk en zeer verscheiden. Het is niet onze bedoeling, in dit werk, al deze problemen uitvoerig te bespreken. We beperken ons tot het belichten van enkele bijzondere aspecten die het nut van het gebruik van de computer duidelijk illustreren.

Zoals bekend kunnen de bibliotheconomische functies in drie grote groepen ingedeeld worden, nl. :

- 1) de beheersproblemen;
- 2) de catalografie;
- 3) de documentatie- en informatieverstrekking.

Deze groepen worden afzonderlijk in de paragrafen 1, 2 en 3 behandeld.

Voor wat de automatisering van de beheersproblemen betreft zijn er, technisch gezien, geen bijzondere moeilijkheden. Zij vertonen een sterke verwantschap met de klassieke commerciële systemen, zoals de bestelling van goederen, periodieke controles van voorraden, boekhouding, het opmaken van statistieken.

Slechts in paragraaf twee over de catalografie en in paragraaf drie over de documentatie- en informatieverstrekking, treden de specifieke bibliothecaire moeilijkheden, voor de verwerking van niet-numerieke gegevens, op het voorplan.

Wegens de hoge kostprijs van geautomatiseerde systemen, dient hier met grote omzichtigheid tewerk gegaan. Met het oog op uitwisseling van bibliografische informatie op nationaal en internationaal vlak, is het nodig dat algemeen aanvaarde, genormaliseerde systemen uitgebouwd worden. Een dergelijk internationaal project, operationeel sinds begin '70, is het MARC-systeem, waarin de bibliografische beschrijving van een document, onder gestandaardiseerde vorm, op magnetische band wordt gebracht.

Vele belangrijke problemen dienen echter nog te worden bestudeerd waaronder : het opstellen van internationale catalografische regels, de ontwikkeling van een wetenschappelijk gefundeerd universeel indexatiesysteem, efficiënte methodes voor automatische indexatie, enz...

In paragraaf 4 van dit artikel wordt een overzicht gegeven van de tot nog toe rerealiseerde projecten binnen de Koninklijke Bibliotheek :

- 1) de permanente inventaris van de Belgische wetenschappelijke productie;
- 2) de permanente inventaris van de nationale en internationale congressen;
- 3) de programma's voor de catalografische verwerking der aanwinsten van het Quetelet Fonds (Bibliotheek van het ministerie van Economische Zaken).

1. DE BEHEERSPROBLEMEN

Hieronder vallen :

- de aanwinstenadministratie
- de tijdschriftenadministratie
- de uitleendienst
- de diensten voor interbibliothecair leenverkeer
- de fotokopieërdienst
- de boekhouding.

Zoals reeds vermeld in de inleiding, zijn deze functies nauw verwant met klassieke bedrijfsprocessen. De automatisering van deze diensten biedt de lezer, afgezien van het feit dat de boeken vlugger ter beschikking zijn, geen opzienbarende nieuwe mogelijkheden. Zij betekent echter wel een zeer grote winst voor het administratief personeel der bibliotheek. Enerzijds wordt het werk sterk vereenvoudigd, en anderzijds wordt er voortdurend een controle uitgedefend. De computer neemt het routinewerk volledig over en schakelt also het veelvuldig hercopiëren van dezelfde bibliografische gegevens uit. Bovendien controleert hij, per programma, de opeenvolging en de volledigheid van de verschillende ingevoerde gegevens.

Als voorbeeld zullen we hier enkele nieuwe perspectieven voor de aanwinstenadministratie schetsen

Een volledig geïntegreerd systeem

We merken onmiddellijk op dat in een geautomatiseerd systeem het onderscheid tussen de punten 1, 2 en 3 (de beheersproblemen, de catalografie en de documentatie- en informatieverstrekking) niet meer zo scherp omlijnd is als in het manuele systeem. Het einddoel is immers, elke bibliografische informatie slechts éénmaal in de computer in te voeren, zelfs indien deze in verschillende stadia wordt gebruikt. Dergelijke werkmethode wordt een volledig geïntegreerd systeem genoemd. Zo kan bvb. een auteursnaam die reeds van bij de bestelling werd ingevoerd, achtereenvolgens worden gebruikt voor : de bestelling, de aanwinstenlijsten, de catalogussteekkaarten, auteursindexen, bibliografieën en zelfs voor information retrieval.

Nieuwe technische mogelijkheden voor de bibliothecaris

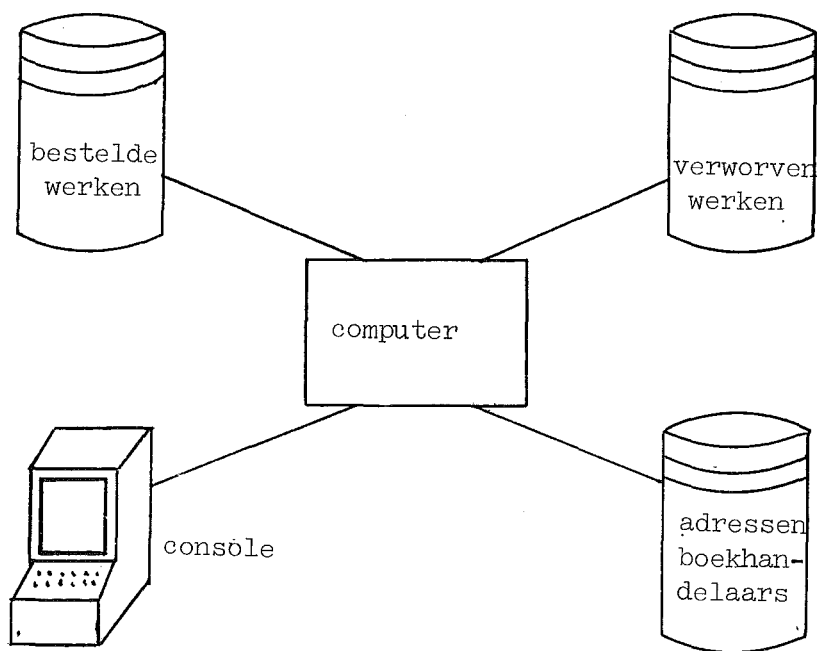
De meest efficiënte werkwijze voor de aanwinstenadministratie ligt in de rechtstreekse conversatie met de computer. De bibliothecaris bedient hierbij een console, die op een van de computer verwijderde plaats kan opgesteld zijn. Een console (cf. fig. 1) bestaat uit een cathodestraalbuis (beeldscherm), die de gewenste gegevens visueel reproduceert, en een klavier, waarlangs de instructies en de nieuwe informatie worden ingevoerd.

De selectie

De selectie gebeurt zoals voorheen, door raadpleging van diverse bronnen, catalogi, bibliografieën, uitgeversprospectussen, enz.

De bibliothecaris meldt de computer dat bij een werk wenst te bestellen. In opeenvolgende stappen verschijnen op het beeldscherm de verschillende elementen vereist voor de bestelling (auteursnaam, titel, uitgever ...).

Langs het klavier worden deze, in zoverre hij erover beschikt, door de bibliothecaris aangevuld. In de eerste plaats zal de computer nagaan of dit werk zich niet reeds bevindt in een der bestanden van de verworven werken of van de bestelde werken (cf. fig. 1).



Indien de computer een gelijkaardige beschrijving vindt, reproduceert zij de voornaamste elementen daarvan, met alle gegevens over de uitgave, op het beeldscherm. De bibliothecaris beslist nu of hij het werk opnieuw aankoopt of niet.

Wordt het werk niet teruggevonden op één der bovenvermelde bestanden of zo de bibliothecaris beslist een tweede exemplaar aan te schaffen, dan kent de computer een definitief identificatienummer aan dit nieuwe document toe. Dit nummer wordt samen met de ingevoerde informatie in het bestellingsbestand opgeslagen. De datum wordt automatisch toegevoegd.

De bestelling

De bestelling gebeurt op vaste tijdstippen (bvb. wekelijks), aan de hand van de gegevens die zich bevinden in het bestand der bestellingen en het bestand der boekhandelaars en wetenschappelijke en commerciële uitgevers (bevattende het adres en een aantal karakteristieken zoals specialiteit, maximale leveringstijd ...) (cf. fig. 1).

Voor elk document wordt per programma, volgens vaste criteria, een keuze gedaan onder de boekhandelaars en wetenschappelijke uitgevers. De bestelbrief wordt automatisch geleverd.

De aanmaningen

Indien het werk niet binnen een vooraf bepaalde periode geleverd wordt, drukt de computer een aanmaningsbrief af.

De ontvangst

Een deel der bewerkingen die hier gebeuren behoren eigenlijk reeds tot de catalografie.

Nu men over het werk zelf beschikt, kunnen de juiste gegevens definitief naar het bestand der verworven werken overgebracht worden. Bij de ontvangst van het document worden de reeds ingevoerde gegevens aan de hand van het identificatienummer opgeroepen. Indien men niet over dit nummer beschikt, volstaat het bvb. enkel de auteur op te geven. De computer tast het bestellingsbestand af en reproduceert alle gevonden werken van deze auteur op het scherm. Alzo kan de bibliothecaris het juiste identificatienummer terugvinden.

De reeds ingevoerde informatie wordt nagezien, aangevuld en verbeterd. Bovendien worden de nieuwe gegevens, vereist voor de aanwinstenadministratie (kostprijs ...), toegevoegd. De computer vult zelf de datum van ontvangst aan.

In fig. 2 volgt een schematische voorstelling van deze werkmethode. Hieruit blijkt onmiddellijk dat, gebruik makend van de computer, veel tijdrovend en weinig interessant werk wordt uitgeschakeld. Bovendien moet men nog in acht nemen dat in een manueel systeem, eerder fouten zullen gemaakt worden dan in een geautomatiseerd systeem (op voorwaarde natuurlijk dat dit helemaal op punt staat). Zo bijvoorbeeld zal de controle die de computer uitvoert bij het nagaan of een werk reeds besteld werd, veel betrouwbaarder zijn dan deze die de bibliothecaris doet. Immers, steekkaarten kunnen ontbreken, op een verkeerde plaats geklasseerd zijn, enz.

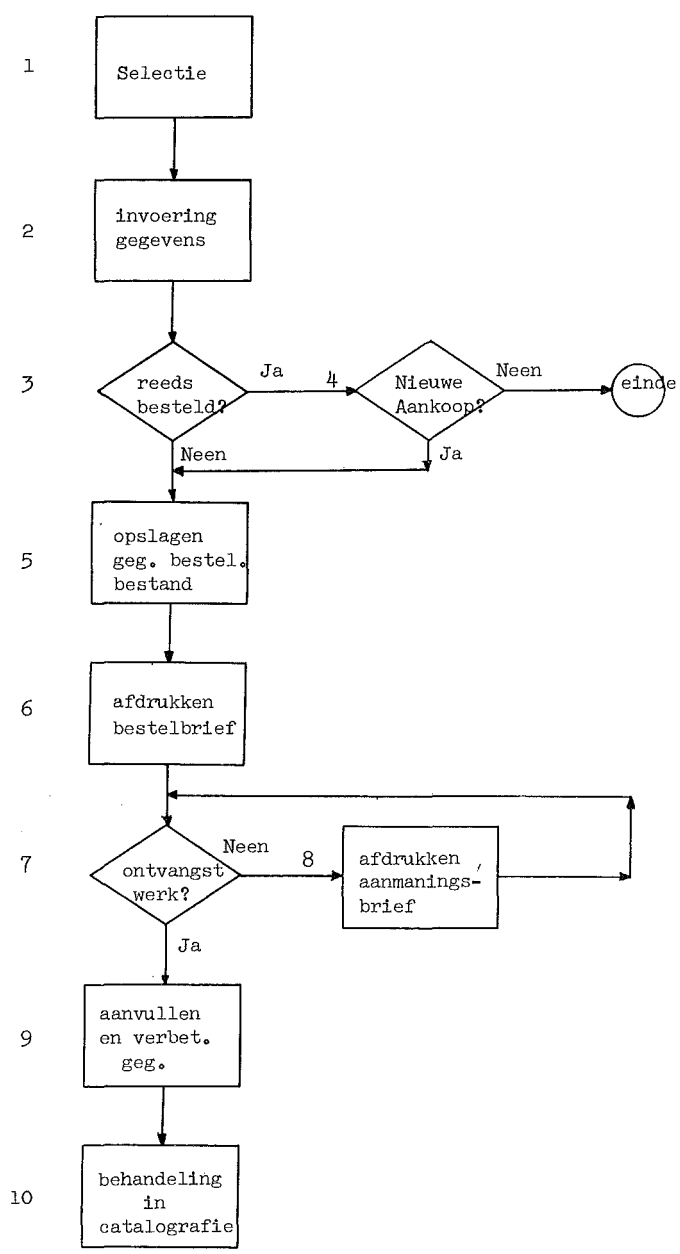


Fig. 2

Tot nog toe hebben vooral Duitse bibliotheken zich toegelegd op de automatisering van beheersproblemen. Bekende voorbeelden hiervan zijn de Duitse Universiteitsbibliotheken (1) waaronder Bochum (2, 3, 4), Regensburg (5, 6, 7), Konstanz (8) en Aken (9).

2. DE CATALOGRAFIE

2.1 *Op het niveau van de bibliotheek*

De bibliografische beschrijving van elk document wordt, op analoge wijze als bij het inbrengen der gegevens bij de aanwinsten, vervolledigd en onmiddellijk verbeterd. Per programma vraagt de computer de ontbrekende elementen in te vullen. Voor elk type van document bestaat een verschillend programma : voor een monografie, een reeks, een tijdschrift, een congresverslag, een rapport ...

Noteer dat, eens de nodige gegevens ingevoerd zijn, alle gewenste steekkaarten, lijsten, bibliografieën, indexen en dergelijke, automatisch kunnen worden samengesteld. Bovendien moet men, voor gelijk welk dezer resultaten, geen enkele verbetering meer aanbrengen. Het zeer onaangenaam werk van het steeds herschrijven en opnieuw verbeteren van dezelfde gegevens wordt hierdoor volledig uitgeschakeld. Enkele interessante projecten worden uitvoeriger beschreven in de publicaties (10, 11 en 12).

Deze voorgespiegelde werkwijze betekent dus reeds een grote vooruitgang voor het werk binnen de bibliotheek.

2.2 *Op inter-bibliothecair niveau*

Gezien de technische mogelijkheden en de economische noodzaak, heeft men het nut ingezien deze nieuwe faciliteiten nog op een hoger vlak te exploiteren. In de U.S.A. is vooreerst de idee van een "shared catalogue" (een gedeelte catalogus) ontstaan. Hieronder verstaat men, een netwerk van bibliotheken die het catalografisch werk, t.t.z. het onder machine-leesbare vorm brengen van het bibliografisch materiaal, onder elkaar verdelen en daarna onderling uitwisselen. Als informatiedrager wordt de magneetband gekozen.

Zo gebeurt, op inter-bibliothecair niveau, de bibliografische beschrijving van elk werk slechts éénmaal en beschikken alle bibliotheken van het netwerk over deze gegevens.

De efficiënte werking van een dergelijk vooruitstrevend project is natuurlijk afhankelijk van zeer uiteenlopende factoren. Voor wat ons interessegebied betreft, stellen zich onmiddellijk twee belangrijke problemen, nl. in verband met de normalisatie der uit te wisselen gegevens en de verdeling van het bibliografisch werk onder de verschillende bibliotheken.

2.2.1 - De normalisatieproblemen

Drie punten dienen onderzocht te worden :

- a) het formaat : nl. de structuur waarin de gegevens onder machineleesbare vorm gebracht worden, de definitie van de op te nemen elementen en de overeenkomende codificaties;
- b) de catalografische regels;
- c) het systeem van classificatie en indexatie.

a) Het formaat

Voor dit punt is men tot nog toe het verst gevorderd.

Her MARC systeem (*Machine Readable Cataloging*) (13, 14, 15, 16, 17, 18 en 19), uitgewerkt door de Library of Congress (één der voornaamste Amerikaanse bibliotheken), is op weg, algemeen erkend te worden door het ISO (*International Standard Organisation*) als een internationale norm. In 1969 werd het trouwens reeds erkend door het USASI (*United States of America Standards Institute*) als een USA-norm (voor een formaat van uitwisseling van bibliografische informatie op magneetband) (20).

Het project is eigenlijk gegroeid uit de traditionele manuele methode voor de verspreiding van de catalogussteekkaarten van de Library of Congress (13). Het oorspronkelijk opzet van het MARC systeem beperkte zich tot onderzoek van de mogelijkheid, catalografisch materiaal, aangepast voor gebruik in plaatselijke bibliotheken onder machinaal-leesbare vorm te brengen. Het project heeft zich nu ontwikkeld tot een internationaal communicatiesysteem voor uitwisseling van bibliografische informatie.

Door de technische onmogelijkheid van directe verenigbaarheid van verschillende systemen, werd de norm zeer breed en zeer soepel opgevat. Enerzijds heeft men rekening gehouden met de diversiteit der machines en programmeertalen, anderzijds met de specifieke objectieven van bepaalde bibliotheken.

We geven hier een kort overzicht van het MARC system (16). De bibliografische elementen worden gekarakteriseerd door een zogenaamde "tag" (= kenteken). Aangezien dergelijke informatie zeer sterk in lengte varieert, is het niet verantwoord met vaste veldlengtes te werken. Op de magnetische band zijn voor elk document, de controle, de identificatie en de informatie volgens een vast schema gestructureerd (cf. fig. 3).

Hierna volgt een uittreksel van de lijst elementaire bibliografische informaties, voorzien voor de beschrijving van een monografie.

zone 1	zone 2	zone 3	zone 4
Leader (Leider,gids)	Record directory (leidraad)	Control fields (Control-velden)	Data fields (Velder der gegevens)

Fig. 3

- Zone 1 : beslaat een vast veld van 24 karakters, bevattende de lengte van het record de toestand van het bestand (verbeterd record ...) en nog enkele specifieke gegevens (tijdschrift, monografie ...)
- Zone 2 : is een veranderlijk veld waarin alle tags gegroepeerd zijn en waarin de lengte en het adres van elke elementaire bibliografische informatie aangegeven is.
- Zone 3 : een veranderlijk veld waarin o.a. het Library of Congress-identificatienummer opgenomen is, dat een testgetal bevat.
- Zone 4 : is een veranderlijk veld met de eigenlijke bibliografische informatie.

Tags Bibliografische informatie

Controlenummers

0 1 0 LC*-kaartnummer
0 1 1 Verwijzings LC*-kaartnummer
0 1 5 Nummer in de Nationale Bibliografie
0 1 6 Verwijzingsnummer in de Nationale

Bibliografie

0 2 0 SBN**
0 2 1 Verwijzings-SBN**
0 3 5 Lokaalnummer
0 3 6 Lokaal verwijzingsnummer
0 4 1 Talen

Identificatienummers

0 5 0 LC*-boeknummer
0 8 0 UDC-notatie
0 8 2 Dewey classificatienummer
0 9 0 Lokaal boeknummer

Hoofdwoord

1 0 0 Auteursnaam
1 1 0 Corporatieve auteur
1 1 1 Congressen e.d.
1 3 0 Eenvormig titelhoofdwoord

* LC = Library of Congress

** SBN = Standaard Boek Nummer

Tags Bibliografische informatie

Gefingeerde titels

2 4 0 Eenvormige titel
2 4 1 Getranslittereerde titel
2 4 2 Vertaalde titel

Collatie

3 0 0 Collatie
3 5 0 Prijs in oorspronkelijke munt
3 6 0 Omgerekende prijs

Aanvullende nota's voor reeksen

4 0 0 Auteursnaam als hoofdwoord
4 1 0 Corporatieve auteursnaam als hoofdwoord
4 1 1 Congrestitel als hoofdwoord
4 4 0 Titel

Bibliografische aanvull. nota's

5 0 2 Thesissen
5 2 0 Excerpten

Onderwerpswoorden

6 0 0 Auteursnaam
6 5 1 Plaatsnaam
6 9 0 Trefwoorden lokaal systeem

In het totaal zijn er een zeventigtal gegevens opgenomen.

Het grote belang van dit project blijkt onmiddellijk uit de interesse die andere vooraanstaande bibliotheken van bij de aanvang ervan betoonden. De Engelse Nationale Bibliotheek heeft in nauwe samenwerking met de Library of Congress het British National Bibliography (BNB) Marc II-systeem ontworpen, dat slechts in enkele kleine details van het oorspronkelijke Marc-systeem afwijkt (21).

Ook in Duitsland (22) en Frankrijk (23) heeft men dit project van nabij gevolgd en een eigen formaat ontwikkeld dat verenigbaar is met het Marc-systeem.

b) *De catalografische regels*

Voor het oogenblik is er nog geen vooruitzicht voor een overeenkomst op internationaal vlak betreffende de normalisatie der catalografische regels. Voor de Britse Bibliografie gebruikt men de Anglo-Amerikaanse regels (21), die een grote overeenkomst vertonen met de zogenaamde Principes van Parijs. Deze laatste werden in 1961 op een internationale conferentie vastgelegd (24). De Parijse Principes zijn zodanig opgevat dat een aanpassing aan locale catalografische werkmethodes geen moeilijkheden biedt. Deze zouden dus als basis kunnen dienen voor de uitwerking van een standaard.

c) *Het universeel systeem van classificatie en indexatie*

Ook hier is men tot het besluit gekomen dat de klassieke technieken niet willekeurig in een geautomatiseerd systeem mogen overgenomen worden. Er is nl. een evolutie waar te nemen waarbij men tracht over te schakelen van de strakke hiërarchische classificatie-systemen zoals bvb. de UDC (= *Universele Decimale Classificatie*) naar meer soepele indexatie-systemen. Aangezien echter de opbouw van een goed universeel indexatie-systeem een zeer moeilijke en zeer dure onderneming is, trachten de meeste instellingen zich voorlopig nog met de U.D.C. te behelpen (25, 10, 11). In Frankrijk heeft men plannen om op nationaal vlak een universele thesaurus bestaande uit een vierduizendtal termen op te bouwen. Ook in de Duitse Universiteitsbibliotheken (1) wordt dit probleem bestudeerd. Op internationaal vlak zijn er echter nog geen contacten gelegd.

2.2.2 *De verdeling der taken*

Welke zijn nu de vooruitzichten voor de toekomst ?

De volle waarde van het MARC-project zal pas tot uiting komen, op het ogenblik dat een internationaal netwerk gecreëerd wordt. Dit zou als volgt kunnen verwezenlijkt worden. Voor elk land of groep van landen wordt een bibliotheek aangeduid die de nationale publicaties centraliseert en tot taak heeft de bibliografische beschrijving daarvan onder machinaal-leesbare vorm te brengen, in een formaat dat verenigbaar is met een internationaal erkend MARC-formaat.

Tevens zou deze bibliotheek instaan voor een selectieve verspreiding van alle bibliografische informatie, gecreëerd binnen het netwerk onder de verschillende bibliotheken in eigen land (cf. fig. 4).

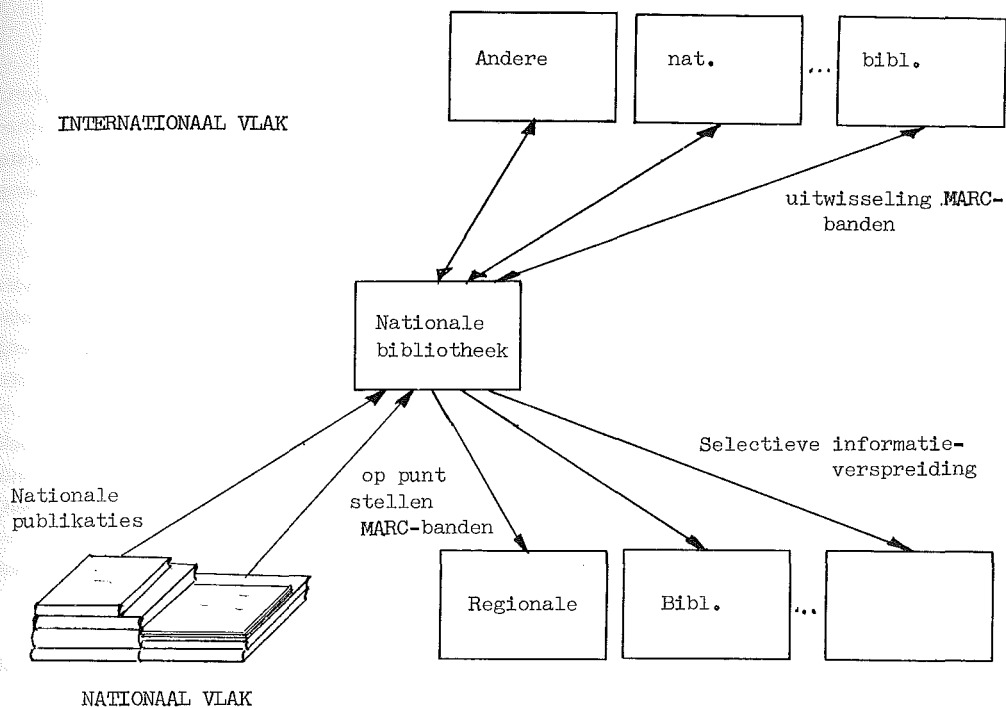


Fig. 4

2.3 Verdere vereenvoudiging op het niveau van de bibliotheek zelf

Deze nieuwe vooruitzichten zullen een hele ommekeer in de huidige organisatie van een bibliotheek teweeg brengen.

Van bij de keuze der aan te kopen werken kunnen de MARC-banden reeds gebruikt worden (11). De computer drukt keuzelijsten af van de werken die in het belangstellingsgebied van de betrokken bibliotheek vallen (volgens bepaalde criteria vb. UDC-notatie, wetensch. uitgever ...). Aan de hand van deze lijsten kan dan een verantwoorde keuze voor de aankoop der werken gebeuren. Per programma worden de nodige catalografische gegevens (deze verschillen nu nog sterk van bibliotheek tot bibliotheek) van de MARC-tape overgenomen en dienen bij de ontvangst van het boek enkel de eventueel ontbrekende gegevens nog aangevuld te worden. Dit betekent dus nog een grote vereenvoudiging van de werkwijze voorgesteld in paragraaf 1.

We leggen er nogmaals de nadruk op dat het beschreven netwerk slechts volkomen efficiënt kan werken, indien internationale akkoorden worden bereikt voor de catalografische regels en voor een universeel indexatie-systeem.

3. DE DOCUMENTATIE- EN INFORMATIEVERSTREKKING

Men verwacht dat de computer voor dit punt, de lezer zelf het meeste diensten zal kunnen bewijzen. Zoals blijkt uit veelvuldige rapporten, heeft onze maatschappij de laatste jaren met een enorme informatieëxplosie te kampen. De taak der documentatie- en informatiespecialisten wordt hierdoor ten zeerste bemoeilijkt. Men hoopt dat een volledige automatisering van het opslag- en terugzoekwerk een oplossing kan brengen.

In verschillende landen heeft men van overheidswege de dringende noodzaak van een grondige studie van deze problemen ingezien. Het rapport "Science, Government and Information" in 1963 in de U.S.A. door de werkgroep van de President's Science Advisory Committee onder leiding van professor A.M. Weinberg uitgebracht, betekende een grote stimulans. Ook in vooraanstaande private bedrijven zijn projecten voor research in die richting aan gang.

De grote moeilijkheden die zich bij de ontwikkeling van informatiesystemen voordoen, zijn in hoofdzaak te wijten aan het dualisme dat bestaat tussen de vraagformulering en de documentbeschrijving. Deze leemte vindt men ook in de klassieke systemen. De aanvrager en de indexeerder zijn twee verschillende personen, waardoor het praktisch onmogelijk is, tot een eenvormigheid in de terminologie te komen.

Aldus krijgen we onmiddellijk een idee van de complexiteit en de moeilijkheidsgraad der problemen die zich bij de automatisering van informatie-opslag en -terugzoeking zullen voordoen.

3.1 De Documentbeschrijving

De technieken der automatische documentanalyse zijn in twee groepen onder te brengen. Deze die uitsluitend gebruik maken van elementen die zich in het document bevinden, en deze die de inhoud interpreteren en transformeren (26).

Een gekend voorbeeld der eerste groep is de KWIC-index (= Keyword in context), op punt gesteld door Luhn (27). In een KWIC-index vindt men voor een bepaald document een ingang op elk betekenisvol woord in de titel. De titel zelf wordt meestal cyclisch gepermuterd zodat het betreffende woord vooraan komt te staan.

Vb. het werk :

Introduction to Numerical Methods and Fortran Programming

zal in een alfabetische KWIC-lijst de volgende ingangen krijgen :

Fortran Programming. ** Introduction to Numerical Methods and Programming.

** Introduction to Numerical Methods and Fortran Numerical Methods and Fortran

Programming. ** Introduction to waarbij de twee asterieksen het begin van de titel aangeven.

De keuze van de relevante woorden gebeurt per programma door uitsluiting van lidwoorden, voegwoorden, adjectieven, enz. die in een afzonderlijk bestand in de computer opgeslagen zijn.

Een dergelijke techniek betekent een grote hulp, doch kan niet als een wetenschappelijk gefundeerd indexatiesysteem aanvaard worden. Verschillende problemen worden totaal over het hoofd gezien waaronder :

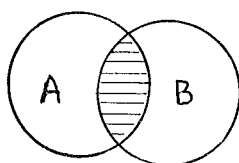
- de mate waarin de titel de inhoud dekt;
- de synoniemen;
- de homoniemen;
- de grammaticale vormen;
- de verschillende spellingswijzen, enz.

Voor de studie van deze laatste onderwerpen belanden we volop in het linguïstisch onderzoek en meteen in de tweede groep methodes. De tot nog toe ontwikkelde theorieën voor de automatische interpretatie van teksten, zijn zeer uiteenlopend van aard. Sommige zijn gebaseerd op een grammaticale analyse, andere op zuivere indexatie of op het opstellen van een korte inhoud (26, 28, 29, 30).

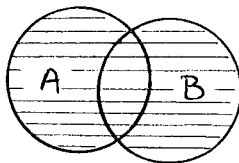
Zoals uit de literatuur blijkt (26, 30, 31), hebben de tot nog toe ontwikkelde technieken weinig bevredigende resultaten opgeleverd. Dit verklaart, dat in vooraanstaande informatiesystemen, waaronder MEDLARS (= *Medical Literature Analysis and Retrieval Systems*, ontwikkeld door de U.S. National Library of Medicine) (32) en het ENDS (= *Euratom Nuclear Documentation System*) (33), de inhoudsomschrijving nog door de vakspecialisten zelf gebeurt. Bovendien heeft men in het huidige stadium slechts efficiënte resultaten bekomen voor projecten die beperkt blijven tot gespecialiseerde verzamelingen (geneeskunde, nucleaire wetenschappen...) waarvoor goede thesauri bestaan en de taalkundige problemen van synonymie en homonymie minder frekwent voorkomen.

3.2 De zoekprogramma's

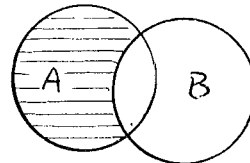
De opzoeking van de relevante documenten gebeurt hoofdzakelijk aan de hand van een aantal indextermen gecoördineerd door de drie logische functies : "en", "of" en "niet" uit de Booliaanse algebra (29, 34). De documenten die over het onderwerp A en B handelen worden verkregen als de doorsnede van de verzamelingen documenten, die de indextermen A en B bevatten (cf. fig. 5). Deze die of over A of over B handelen bekomt men als de unie van de twee verzamelingen (cf. fig. 6). Tenslotte de documenten die over A handelen doch niet over B zijn het resultaat van het verschil van A en B (cf. fig. 7).



A en B
Fig.5



A of B
Fig. 6



A Niet B
Fig. 7

In meer gesofistikeerde systemen zijn nog supplementaire logische bewerkingen voorzien. Zo bvb. kan aan elke indexterm een constant waardecijfer toegekend worden. De computer telt voor een document de waardecijfers van de relevante indextermen samen en neemt dit werk slechts in aanmerking indien een bepaalde waarde overschreden wordt. Deze werkwijze kan evenwel ook bij de beschrijving van de documenten toegepaste worden (28).

Een andere techniek bestaat erin, van bepaalde woorden slechts de "stam" in te voeren. De overige karakters worden als parameter aanzien. Vb. Industr.. kan bij vergelijking met de indextermen der documentbeschrijving in overeenstemming gebracht worden met : Industry, Industries, Industrial... Men spreekt hier van masker-methodes.

In information retrieval-systemen, gebaseerd op vrije tekst (free text) - onderzoek (waaraan geen gecontroleerde lijst van indexatietermen geassocieerd is) worden nog andere procedures aangewend. Bij Text-Pac (IBM Technical information retrieval center) wordt o.a. nog gebruik gemaakt van het zoeken naar groepen van aangrenzende woorden, termen die zich in eenzelfde zin bevinden (35), enz.

3.3 De vraagformulering

De gebruiker stelt de vraag in een natuurlijke taal. Een documentalist die goed op de hoogte is van het indexatiesysteem en de zoekprogramma's, transformeert deze onder een voor de machine geschikte vorm. Men spreekt van een "zoekprofiel" dat bestaat uit een aantal indextermen die volgens een bepaald patroon logisch gecombineerd zijn.

De rechtstreekse aanwending van een dergelijk profiel schenkt echter geen voldoening.

Het onderzoek naar de bruikbaarheid der ontvangen referenties leverde, zelfs voor systemen met manuele indexatie, weinig bevredigende resultaten op. Het ENDS (33) bvb. geeft een relevantieverhouding (= het percentage bruikbare afgeleverde documenten) van minder dan 50 %. Hieraan tracht men enigszins te verhelpen door toevoeging van een techniek van terugkoppeling (feedback).

Aan de documentatie-gebruiker wordt een gedetailleerde beoordeling van de ontvangen resultaten gevraagd. Men tracht ondermeer te weten te komen hoeveel referenties werkelijk pertinent zijn, hoeveel er het onderwerp niet helemaal dekken doch nuttig zijn, en welke referenties ontbreken die de gebruiker kent en die van fundamenteel belang zijn. Zo kan een onderzoek ingesteld worden naar de oorzaken van het slechte resultaat. Hiermee rekening houdend past men het gebruikersprofiel aan, en de procedure wordt herbegonnen. Dank zij deze terugkoppelingsmethode kon bij het ENDS reeds een gemiddelde relevante verhouding van 75 % bereikt worden.

In een verdere poging de informatiebenutting volledig te automatiseren zijn er projecten die de documentatiespecialist bij de vraagformulering uitschakelen. De aanvrager communiceert hier rechtstreeks met de computer. De vraag wordt in een natuurlijke taal gesteld. De machine transformeert de vraag in een aantal indexatietermen. De gebruiker duidt de passende termen aan in hun logisch verband. Een aantal documenten die beantwoorden aan dit zoekprofiel worden visueel op het scherm weergegeven. De lezer controleert deze naar hun relevantie (= systeem van instand feedback). Aan de hand van deze resultaten doet het systeem een nieuw voorstel voor het zoekprofiel, waarna heel het proces kan herbeginnen.

Deze projecten zijn zeer vooruitstrevend. Een experimenteel voorbeeld hiervan is het INTREX project (= *Information transfer experiment*) (36), ontwikkeld door het MIT (*Massachusetts Institute of Technology*). In dit systeem worden bovendien de mogelijkheden microfilm-computer gecombineerd. De informatie wordt onder sterk gecodeerde vorm in de computer opgeslagen en verwijst door middel van een plaatscode naar de relevante microfilmbeelden die via cassettes snel toegankelijk zijn.

3.4 Besluit

Schematisch kan het besproken systeem als volgt voorgesteld worden (cf. fig. 8).

Deze information retrieval-methode geeft interessante vooruitzichten. Zij is het onderzoekstadium echter nog niet ontgroeid.

Men hoopt, door een verder doorgedreven linguïstisch onderzoek, het rendement der bestaande semi-automatische systemen op te drijven en in een volgend stadium, een verdere automatisering door te voeren.

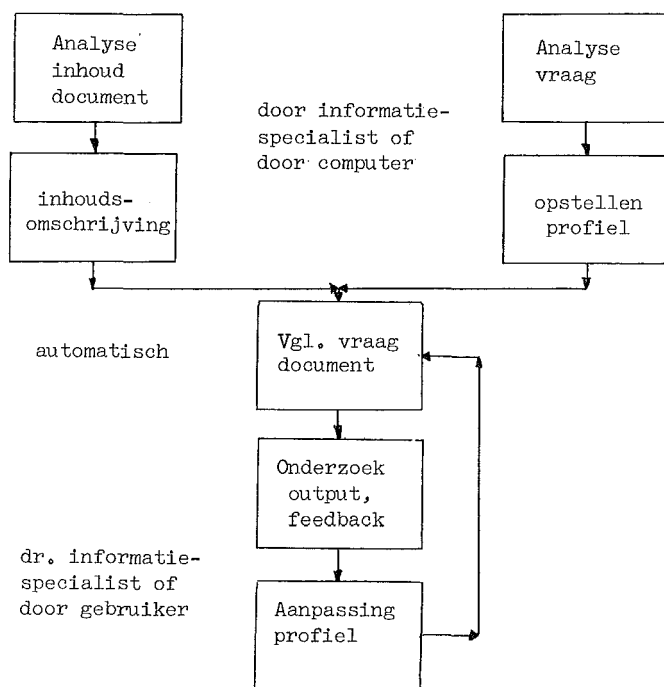


Fig. 8

4. REALISATIES EN PROJECTEN BINNEN DE KONINKLIJKE BIBLIOTHEEK

Daar de Koninklijke Bibliotheek er in de eerste plaats naar streeft de lezer zoveel mogelijk in zijn opzoeken te helpen, behandelen de tot nog toe verwezenlijkte projecten in hoofdzaak de catalografie en de documentatie (37).

4.1 De operationele systemen zijn :

4.1.1 De permanente inventaris van de Belgische wetenschappelijke productie (Key of Belgian Science)
In dit project wordt er naar gestreefd een zo volledig mogelijke inventaris der wetenschappelijke artikels van Belgische auteurs samen te stellen.

De bibliothecaris ziet een aantal belangrijke tijdschriften, jaarverslagen e.d. na.

Op gedrukte standaardformulieren worden de bibliografische gegevens ingevuld (cf. fig. 9). In de linker kolom van deze formulieren zijn de codes die de verschillende elementen identificeren vooraf aangebracht. Bij de aanduiding der trefwoorden maakt men gebruik van de Amerikaanse E.J.C.-thesaurus (= *Engineering Joint Council*).

De gegevens worden op ponskaarten geperforeerd en maandelijks door de computer verwerkt. Als output bekomt men cumulatieve lijsten (maximale periode : 1 jaar) waaronder : een basislijst die gerangschikt is volgens *E.J.C. nummer* en een reeks registers volgens *onderwerp, auteur, corporatieve auteur en tijdschrift* waarin gepubliceerd werd. In 1970 werden een 10.000-tal artikels ingevoerd.

We hopen dat dit belangrijk project zich tot een informatiebron van internationaal formaat zal ontwikkelen.

4.1.2 De Kalender der Nationale en Internationale Congressen (38)

Hier wordt op volkomen analoge wijze tewerk gegaan als in voorgaand project. Een aantal kalenders, tijdschriften en losse aankondigingen worden nagezien. Na selectie vult men de standaardformulieren voor de congressen in. Als resultaat levert de computer een reeks steekkaarten : het basisbestand volgens *E.J.C. nummer*, een bestand volgens *data, plaats* en een volgens *trefwoorden der E.J.C.-thesaurus*. De computer drukt ook brieven af waarin uitgebreide gegevens over het congres gevraagd worden aan de organiserende instelling. Jaarlijks worden een 4.000-tal congressen behandeld.

In een volgende stap zullen de opzoeken naar deze verschillende gegevens ook geautomatiseerd worden. Men denkt hierbij aan de ontwikkeling van SDI systeem (*Selective Dissemination of Information*). Aldus zou elke belangstellende vorser, aan de hand van een geschikt zoekprofiel, regelmatig op de hoogte gehouden worden van de komende congressen in zijn interessegebied.

7	13	17 18	29				
§ § E	197. A	50100L934704					
S/A	13 14 TYPE	15 16 SUPPORT	21 22 CONTENU	13 14 REL. PHYSIQUE			
	<input checked="" type="checkbox"/> P P Periodique <input type="checkbox"/> C E Conference <input type="checkbox"/> T H These <input type="checkbox"/> S R Rapport	<input checked="" type="checkbox"/> D D Imprime <input type="checkbox"/> M R Copie <input type="checkbox"/> I N Information PAYS 17 <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> 19	<input checked="" type="checkbox"/> W W Scientifique <input type="checkbox"/> E D Vulgaris. <input type="checkbox"/> B I Bibliogra. <input type="checkbox"/> W T Tech.-Commer	<input checked="" type="checkbox"/> C E Entier <input type="checkbox"/> I N Incomplet <input type="checkbox"/> E X Extrait <input type="checkbox"/> U S Avec supplém. <input type="checkbox"/> U C Dans Serie			
S A	AUTEURS (NOM, PRENOMS; QUALIFICATION.) (NOM. INST SUP, INST INF; LIEU.) ... F.A.B.R.Y., G. ... AC Acad. Lukenhous. St. Raphael. Wilgevoild.						
LANGUE	TITRE (MOTS NON SIGNIFICATIFS-KWIC-ROMAINS) Use of computers in medicine, with special emphasis on its use in the orthopedic clinic						
§ T <input checked="" type="checkbox"/>							
§ PP <input checked="" type="checkbox"/>	TITRE PERIODIQUE (MOTS NON SIGNIFICATIFS-ROMAINS) Acta orthopaedica belgica						
S/B	30 ANNEE	34 VOLUME	38 N° PER	42 NUMERO EDITEUR	51	52 PAGE	61
	1970	36	2			255-271	
§ RR	COTE RANGEMENT B 3858 / 12065						
§ YS	REF EJC. MOT VEDETTE (N°. MOT). (N°. MOT) 09.00 Computers 14.00 Automation 08.00 Data processing 09.00 Data storage 05.00 " " 05.00 " " " " " " 05.00 Scientific research 14.00 " " 06.05 Orthopedics						

Fig. 9

4.1.3 *De programma's voor de catalografische verwerking der aanwinsten van het Quetelet Fonds (Centrale Bibliotheek van het Ministerie van Economische Zaken) (10)*

In het vooruitzicht van een verdere automatisering binnen de Koninklijke Bibliotheek, was het interessant, eerst enige ervaring op te doen bij een project van kleinere omvang. Einde '68 werd gestart met de analyse van de bibliotheek van het Quetelet Fonds. Deze is niet zo uitgebreid als de Koninklijke Bibliotheek en blijft tevens grotendeels beperkt tot een bepaalde discipline. De programma's voor de verwerking der catalografische gegevens, ontwikkeld in samenwerking met IBM, zijn sedert april '70 operationeel. Maandelijks wordt een catalogus der aanwinsten samengesteld, waarin de bibliografische eenheden automatisch volgens de *U.D.C. notaties* worden gerangschikt. Verder drukt de computer internationale steekkaarten af, gerangschikt volgens *geografische U.D.C., auteur en titel* (voor de anoniemen).

4.2 *De lopende projecten*

4.2.1 *De Belgische en Luxemburgse Centrale Catalogus van Lopende Buitenlandse Tijdschriften*

Op dit ogenblik wordt de bestaande catalogus (een 40.000-tal titels; uitgegeven in 1965) onder machinaal-leesbare vorm gebracht. De aanvullingen van bezit, titels, titelwijzigingen, bibliotheken, enz. zullen per programma automatisch toegevoegd worden. Eens het systeem op punt staat, zal het mogelijk zijn, zonder veel administratief werk, regelmatig een cumulatieve catalogus uit te geven.

De andere projecten liggen volledig in de lijn van dit artikel :

4.2.2 Ontwerp van een MARC-formaat voor de Belgische publicaties

4.2.3 Studie voor een volledig geïntegreerd systeem binnen de Koninklijke Bibliotheek

4.2.4 Het op punt stellen van een S.D.I.-systeem voor bestaande informatiediensten op magneetband zoals Chemical Abstracts, MEDLARS, Engineering Index, enz.

4.2.5 Opzoekingswerk in verband met informatieopslag en -terugzoeking.

Alvorens dit artikel te besluiten, vermelden we nog het interessante werk : "Annual Review of Information Science and Technology" (39), waarin jaarlijks een uitgebreid overzicht wordt gegeven van de nieuwe ontwikkelingen op het gebied der documentaire informatieverwerking.

Het is een feit dat slechts met behulp van de computer een oplossing zal worden gevonden voor de ingewikkelde problemen waarmee de bibliotheken en informatiecentra nu reeds te kampen hebben. Wat de technische apparatuur betreft, beschikt men op dit ogenblik over ruime mogelijkheden. Voor de systemen zelf, blijft er echter nog veel interessant werk te verrichten.

- (1) WIEDER Joachim, Computer techniques in the libraries of the Federal Republic of Germany. Applications and trends. *Unesco Bull. Libr.*, vol. 23, n° 5, September-October 1969, p. 245-252.
- (2) PFLUG G. en ADAMS B., *Electronische Datenverarbeitung in der Universitätsbibliothek Bochum*. Bochum, Universitätsbibliothek, 1968, 147 p.
- (3) Die automatisierte Buchausleihe (mit Beiträgen von BOSSMEYER C., ADAMS B., HEIM H. und PFLUG G.), Bochum, 1967.
- (4) HELAL Ahmed H., Circulation, Recording and Control of Interlibrary Loans. **Referat für das Unesco-seminar : Electronische Datenverarbeitung in Bibliotheken, Regensburg, 13-18 april 1970**, 10 p.
- (5) NIEWALDA Paul en PREUSS Gertrud, Die Electronik im Dienste der Katalogisierung der Universitätsbibliothek Regensburg. Ein Erfahrungsbericht. *Zeitschrift für Bibliothekswesen und Bibliographie*, vol. 16, 1969, p. 86-118.
- (6) NIEWALDA Paul, Input and Processing in the University Library of Regensburg. Execution (Abridged version). *Unesco-seminar : Electronic data processing in Libraries, Regensburg, 13-19 april 1970*, 5 p.
- (7) PAUER M., Datacollecting and processing at the University Library of Regensburg. Planning. *Unesco-seminar : Electronic data processing in Libraries, Regensburg, 13-19 april 1970*, 12 p.
- (8) STOLTZENBURG Joachim en RABE Günther, Das Konstanzer Schema zur Erfassung bibliographischer Daten. *Zeitschrift für Bibliothekswesen und Bibliographie*, vol. 16, 1969, p. 119-153.
- (9) FELLMANN Ulrich, Computer-aided Circulation Control : On-line Systems. *Unesco-seminar : Electronic data processing in Libraries, Regensburg, 13-19 april 1970*, 9 p. + bijl.
- (10) BARTHOLOMEUS M., L'aspect informatique de la catalographie automatique. (L'automatisation des bibliothèques, 1ère partie). Centre National de Documentation Scientifique et Technique, Bibliothèque Royale Albert 1er, Bruxelles, 1970, 63 p. + bijl.
- (11) GIBUS, Paris, Impr. Berger-Levrault, 1970, 4°, 8 p. Groupe informatiste de bibliothèques universitaires et spécialisés.
- (12) Un système unifié d'information pour la Bibliothèque Nationale du Canada. Résumé du **Rapport** du Projet d'Analyse et d'Automatisation des services. Ottawa, Bibliothèque Nationale du Canada, 1970, 88 p.

- (13) AVRAM Henriette D., The MARC Pilot Project. Washington, Library of Congress, 1968, 183 p.
- (14) The MARC II Format. A communications Format for Bibliographic Data. Prepared by Henriette D. AVRAM, John F. KNAPP and Lucia J. RATHER. Washington, Library of Congress, 1968, 167 p. ; Supplement One, 1968, 7 p.
- (15) "Preliminary Guidelines for the Library of Congress, National Library of Medicine, and National Agricultural Library Implementation of the Proposed American Standard for a Format for Bibliographic Information Interchange on Magnetic Tapes as Applied to Records Representing Monographic Materials in Textual Printed Form (Books)". Journal of Library Automation, 2, p. 68-83.
- (16) MARC Manuals used by the Library of Congress. Prepared by the Information Systems Office, Library of Congress, Information science and automation division, American Library Association; Chicago, 1969, 4°, 219 p. + bijl.
- (17) A MARC Format. Specifications for Magnetic Tapes containing Monographic catalog Records in the MARC II - Format. Washington, Library of Congress.
- (18) Serials : A MARC Format. Preliminary edition. Information Systems Office. Washington, Library of Congress, 1970, 72 p.
- (19) Maps : A MARC Format. Specifications for Magnetic Tapes Containing Catalog Records for Maps. Information Systems Office. Washington, Library of Congress, 1970, 45 p.
- (20) United States of America Standards Institute. Committee on Library Work and Documentation. Sub-Committee on Machine Input Records. U.S.A. standard for a format for bibliographic information interchange on magnetic tape. USASI, New York, 1969. (USAS Z 39.2-1969).
- (21) COWARD Richard E., MARC International. Journal of Library Automation. Vol. 2/4, December 1969, p. 181-186.
- (22) Entwurf eines Kategorienschema zur Erfassung bibliographischer Daten. Informationen (3). Arbeitsstelle für Bibliothekstechnik bei der Staatsbibliothek Preussischer Kulturbesitz. 1970, p. 2-21.
- (23) CHAUVEINC Marc, Monocle. Projet de mise en ordinateur d'une notice catalographique de livre. Grenoble, Bibliothèque Universitaire, 1970, 156 p. + bijl.

- (24) International Federation of Library Association. International Conference on Cataloguing Principles; Paris, 9th-18th October. Report Ed. by A.H. CHAPLIN.
- (25) Proceedings of First Seminar on U.D.C. in a Mechanized Retrieval System. Conducted by R.R. FREEMAN and Pauline ATHERTON, Copenhagen, 2th-6th September, 1968. Copenhagen, Danisch Centre for Documentation, 1969.
- (26) COYAUD M., SIOT-DECAUVILLE, L'analyse automatique des documents (Informatique I). Mouton - Paris, 1967, 141 p.
- (27) LUHN H.P., KWIC-Key-word-in-context Index for Technical literature. 1959, IBM techn. Publ.
- (28) SALTON G., Automatic information organization and retrieval. Computer Science Series, New York, McGraw Hill, 514 p.
- (29) CROS R.C., GARDIN J.C., LEVY F., L'automatisation des recherches documentaires. (Documentation et Information). Un modèle général : le SYNTOL, Paris, Gauthier-Villars, 1968, 260 p.
- (30) LEVERY P., Documentation automatique. Traité pratique d'informatique. (Techniques de l'ingénieur) (p H8660), 1970.
- (31) DE PAUW G., Problematiek van de geautomatiseerde informatie benutting : een overzicht. Intermediair 5, 20 november 1970, p. 13.
- (32) DE GRAAFF Edo, MEDLARS en Index Medicus. Open 1 (1969) 7/8 (juli/aug.) p. 432-435.
- (33) VERNIMB Carlo, Een met computer uitgerüst informatiesysteem en zijn clientèle. Euro Spectra, 9, p. 11-19.
- (34) LE DU Jacques, Les travaux de la section d'automatique documentaire du Centre National de la Recherche Scientifique. Documentaliste. Numéro Spécial, premier trimestre 1966, p. 15-33.
- (35) KAUFMAN Samuel, Text-Pac Design Considerations. ASIS. Proceedings of the American Society for Information Science, Vol. 6, 1969, p. 77-81.
- (36) Massachusetts Institute of Technology. Project INTREX. Semiannual activity report. 15 sept. 1969. Cambridge, Massachusetts, 1969, 39 p.
- (37) COCKX A., Bibliotheek, Documentatie en Computer. Handelingen van het XXVIIe Vlaams Filologencongres, p. 257.

- (38) LEBRUN L., COCKX A., Expérience belge d'information rétrospective, sélective et future des congrès nationaux et internationaux. Vienna, International Atomic Energy Agency, 1970, 6 p. (Overdruk uit Handling of Nuclear information, p. 521-526).
- (39) Annual Review of Information Science and Technology. Editor : CUADRA C.A. Assistant editor : LUKE A.W. Encyclopaedia Britannica, Inc. Chicago.