

FUNCTIONELE ASPECTEN VAN DE BESTANDSORGANISATIE TEN BEHOEVE VAN AUTOMATISCHE INFORMATIEVERWERKING

door Dr. G. Bresser

De voortdurende ontwikkelingen op het gebied der automatische informatieverwerking confronteren ons veelvuldig met ogenschijnlijk nieuwe begrippen. Het *bestand* is één van de vele. In de grote Van Dale komt van dit begrip echter geen verklaring voor die de lezer direkt aansluiting geeft op de informatieverwerking. De betekenis van het bestandsbegrip in de bosbouw - verzamelnaam voor al wat op een bepaalde bosgrond groeit - biedt nog het meest aanvaardbare aanknopingspunt. Wij omschrijven het bestand in het kader van ons onderwerp als *een verzameling informatie*.

Aannemende dat het verzamelen om der wille van het verzamelen in een organisatie niet aan de orde is, zal de verzameling steeds met een bepaald oogmerk worden opgebouwd. Dit doel is mede bepalend voor de gegevens die in het bestand worden opgenomen en de wijze waarop dat gebeurt. Men streeft naar een systematische verzameling. Alleen door een bewuste ordening kan het bestand een onderdeel van het „geheugen” van een organisatie zijn. Hiermee is niets nieuws gezegd. In alle organisaties komen bestanden voor al hebben ze vaak niet die naam gekregen. Het inkoopboek, het magazijnvoorraadboek en het debiteurenboek zijn duidelijke voorbeelden. Ook in de vertrouwde boekhouding vormen bestanden een essentieel onderdeel van het gehele informatie registratie-, verwerkings- en verstrekkingssysteem. We zullen echter moeten proberen het geheel vanuit een andere gezichtshoek te benaderen om daarmee de mogelijkheden van het bestand in relatie tot automatische informatieverwerking beter te kunnen bepalen.

Het gaat in dit artikel niet in de eerste plaats om de betekenis van de gegevens die worden verzameld, maar om een beoordeling van de ordening der gegevens op grond van de betekenis die het bestand heeft voor het functioneren van de organisatie. Wij zullen proberen de criteria die bij de ordening van de informatie van belang kunnen zijn in hun onderlinge verband te presenteren. Bij dat alles zal steeds de functie die het bestand voor de organisatie als complex systeem vervult maatgevend zijn.

In het voorgaande kwam de noodzakelijke relatie tussen bestand en administratie reeds ter sprake. Een administratie zonder bestanden - van welke omvang dan ook - is niet goed denkbaar. Ter beoordeling van de functies van het bestand is het dan ook nuttig nog eens terug te gaan naar de functies van de administratie die de meeste lezers van dit tijdschrift wel eens in één of andere vorm zullen hebben bestudeerd. Een dergelijk overzicht geeft terzelfder tijd een indicatie van de betekenis van het bestand voor een organisatie.

In aansluiting op de talrijke varianten die door andere auteurs in dit verband worden onderscheiden noemen we de volgende elementen:

- 1 De administratie maakt het mogelijk op grond van de vergelijking van plan en realiteit een *stuurimpuls* te geven. (Terugkoppeling.)

- 2 De administratie als bron van informatie ten behoeve van de *beslissingsvoorbereiding*, op elk niveau.
- 3 De administratie als bron van informatie waarin binnen de organisatie *opdracht* en *taakstelling* zijn gegeven.
- 4 De administratie als bron van informatie ten behoeve van de *communicatie bij de uitvoering*.
- 5 De administratie als bron van informatie *ter controle* van beleid, uitvoering, beheer en bewaring.
- 6 De administratie als een *historisch verslag*; een middel tot bewijs.

Een dergelijke brede schaal van functies is ontstaan in een langdurige evolutie. Het valt gemakkelijk in te zien dat bestanden voor deze functies van groot belang zijn. De stuurimpuls die in de eerste functie als het resultaat van een terugkoppeling werd gepresenteerd heeft slechts betekenis wanneer ten behoeve van beslissingsvoorbereiding, opdracht, taakstelling en communicatie bij de uitvoering over adequate informatie kan worden beschikt. Deze informatie wordt verzameld, bewerkt en in bestanden vastgelegd om, onder andere, het functioneren van de organisatie als geheel te vergemakkelijken.

Reeds voordat computers op ruime schaal in de administratie werden toegepast groeide geleidelijk een inzicht in de samenhang van de informatie die in verschillende bedrijfsonderdelen ontstaat of wordt gebruikt. Daarnaast sluit ook de ontwikkeling van het normatieve element in de administratie, dat zijn oorsprong vindt in de wens met behulp van algemeen aanvaarde regels (normen) tot een doelmatige besturing te komen, goed aan op de technische mogelijkheden van de huidige generatie computers. Het wordt met behulp van de moderne technieken van bestandsorganisatie - technieken die overigens in dit artikel buiten beschouwing blijven - mogelijk een veelheid van informatiebehoefte uit enkele soorten van basisgegevens te bevredigen. De functies van de administratie worden op deze manier koppelingselementen die een geleidelijke opbouw van een geïntegreerd systeem van informatieverwerking mogelijk maken. Het begrip administratie wordt in dat verband steeds minder vaak gebruikt; liever spreekt men van bestuurlijke informatieverwerking. Het lijkt ons op grond van het voorgaande geoorloofd een tweetal hoofdfuncties van de informatieverwerking te onderscheiden: allereerst een *stuurfunctie*, daarnaast een *verantwoordingsfunctie*. Het is niet noodzakelijk deze beide functies uitsluitend op de administratie als geheel van toepassing te achten. In het voorgaande is ook de betekenis van het bestand - als onderdeel van de administratie - voor het functioneren van de organisatie aan de orde gesteld.

Wij willen er nog eens uitdrukkelijk op wijzen dat deze beide hoofdfuncties van het bestand weliswaar als gevolg van de ontwikkeling van administratief-technische hulpmiddelen duidelijker kunnen worden waargenomen, ze zijn echter op zichzelf niet een gevolg van de toepassing van automatische bestuurlijke informatieverwerking. In een organisatie zijn vele vormen van informatieverwerking denkbaar, waarbij het bestand dezelfde functie vervult, zonder dat van een computer gebruik wordt gemaakt.

Motieven tot inschakeling van elektronische rekenapparatuur liggen veelal op een geheel ander vlak. Vaak wordt bijvoorbeeld de personeelschaarste

genoemd. Het is mogelijk dat door veranderingen in het onderwijssysteem een structurele schaarste aan bepaalde categorieën werknemers ontstaat die alleen kan worden bestreden door machinale activiteiten de menselijke te laten vervangen. Nauw hiermee verbonden is het motief waarin een mogelijke kostenverlaging centraal wordt gesteld. Bij een toenemende complexiteit van organisaties zijn vanuit bestuurlijk zowel als verantwoordings oogpunt vooral die motieven van belang waarin de nadruk wordt gelegd op de grote betrouwbaarheid, accuratesse en snelheid in de informatieverwerking waarmee computers zich van mensen onderscheiden. In dergelijke situaties komen mogelijkheden aan de orde die voordien vrijwel uitgesloten waren. De grenzen der toepassingen kunnen door de ontwikkeling van de apparatuur steeds verder worden verlegd. De inventiviteit van de systeemontwerper zal zich dientengevolge steeds meer moeten richten op de organisatie van bestanden om de stuur- en verantwoordingsfunctie ten volle tot hun recht te laten komen. In dit artikel beperken wij ons verder tot bestanden die met behulp van een computer worden behandeld.

De beide functies die in het voorgaande werden besproken maken nog eens duidelijk dat door middel van een systematische opbouw van het bestand de *toegankelijkheid* van de informatie te allen tijde verzekerd moet zijn. Deze ordening is in aansluiting daarop enerzijds gericht op de manier waarop het bestand wordt *bijgewerkt* terwijl anderzijds de wijze van *raadplegen* centraal staat.

De informatie met betrekking tot één object - een klant, een artikel, een personeelslid -, kan vanuit verschillende gezichtspunten (aspecten) worden benaderd. De keuze van de aspecten aan de hand waarvan het bestand toegankelijk moet worden gemaakt is afhankelijk van de functie die het betrokken bestand in de organisatie vervult. Een klantenbestand kan bijvoorbeeld worden opgebouwd aan de hand van klantnummers; het is echter evenzeer denkbaar de naam als primaire ingang te gebruiken. Zodra een bestand verschillende nevenfuncties, dan wel van de hoofdfunctie afgeleide subfuncties vervult bestaat de mogelijkheid dat vanuit elk der functies een ander aspect wordt aangewezen van waaruit toegankelijkheid gewenst wordt. Bij het ontwerpen van systemen voor automatische informatieverwerking wordt in dergelijke gevallen veelal gebruik gemaakt van tabellen. Is in ons voorbeeld het bestand op klantnummers georganiseerd dan kan een tabel worden opgebouwd met behulp waarvan de alfabetische ingang wordt gerealiseerd. In die tabel staat achter elke naam - soms een afgekorte of zelfs een gedeeltelijk gecodeerde naam - het klantnummer. Via het klantnummer kan daarna alle informatie over het object in het hoofdbestand worden teruggevonden. Met deze wijze van organiseren wordt voorkomen dat alle informatie over de objecten in een voortdurend wisselende volgorde in verscheidene bestanden moet worden gedupliceerd. Er bestaat slechts één hoofdbestand met een naar de behoefte variërend aantal registers (indexen).

Wij zijn met dit voorbeeld al midden in het vraagstuk van de bestandsorganisatie terecht gekomen. Omdat wij in dit artikel de technische varianten buiten beschouwing zullen laten ten gunste van een korte analyse van de functionele aspecten van de bestandsorganisatie, beperken wij ons tot de

behandeling van enkele factoren die in hun onderlinge verband voor de organisatie van bestanden van belang zijn.

In ons betoog hebben wij tot dusverre de toegankelijkheid als uitgangspunt gekozen. Daarnaast werd het aspect genoemd als sleutel waarmee de toegang kan worden gerealiseerd. De toegang tot het bestand bleek nodig om het, in relatie tot de functie, te kunnen bijwerken en/of raadplegen. Bijwerken van bestanden heeft betrekking op (1) het muteren van de reeds in de bestanden aanwezige informatie; (2) het inbrengen van informatie over nieuwe objecten; (3) het verwijderen van informatie over objecten die in het kader van de organisatie niet langer van belang zijn. Het zal duidelijk zijn dat het bijwerken van een bestand nooit op zichzelf staat. Alleen een bestand dat voortdurend aan de aktuele situatie wordt aangepast heeft als bron van informatie waarde. De wisselwerking tussen bijwerken en raadplegen zal bij het organiseren van bestanden dan ook voortdurend in de beschouwing worden betrokken. Naarmate er meer aspecten in het geding zijn, die mogelijkerwijze voor bijwerken en raadplegen verschillen, gaat de betekenis van de genoemde relatie zwaarder wegen. Alleen door de elementen die hierna aan de orde zullen komen in beide gevallen te operationaliseren kan een aanvaardbare keuze ten aanzien van de bestandsorganisatie worden gemaakt.

Een eerste belangrijk gegeven voor de systeemontwerper is het aantal mutaties in het bestand per tijdséénheid. In de tweede plaats zal hij willen weten op welke aspecten die mutaties meestal betrekking hebben, anders gezegd in welke gegevens regelmatig wijzigingen komen. Een derde gegeven dat naar onze mening zeer belangrijk is betreft de geografische spreiding van de mutaties. We zullen er gemakshalve van uit gaan dat het bestand als één geheel op één plaats in de organisatie wordt gehanteerd. Geografische spreiding van de mutaties betekent in dat geval dat de mutaties op één of andere manier naar het bestand moeten worden gebracht. Naast de ruimtelijke factor speelt daarin opnieuw de tijd een belangrijke rol. Elk transport vraagt immers tijd. Afhankelijk van het gewicht dat in de organisatie aan de tijd wordt toegekend zal een antwoord moeten worden gegeven op de vraag of het bestand seriegewijze dan wel postgewijze moet worden bijgewerkt. Het seriegewijze bijwerken is gebaseerd op het verzamelen van de mutaties. Deze worden vervolgens gecodeerd om daarna achtereenvolgens in het bestand te worden verwerkt. Postgewijze verwerking vindt plaats wanneer de mutaties bij het ontstaan ook direkt individueel aan de computer ter verwerking worden aangeboden. Een toenemende spreiding van de mutaties gekoppeld aan de wens het bestand snel bij te werken vraagt steeds duidelijker om de realisatie van on-line verbindingen. Van een on-line verbinding is sprake wanneer tussen een computer - en daarmee een bestand - en een aantal eindstations, die geografisch willekeurig kunnen worden gespreid, direkte overdracht van gegevens mogelijk is via datatransmissielijnen. Het bijwerken van bestanden betekent in dat geval dat gegevens via het eindstation naar de computer worden gezonden om in aansluiting daarop direkt te worden verwerkt. Met hetzelfde station kan het bestand in principe ook worden geraadpleegd. De overdracht van gegevens vindt dan in omgekeerde richting plaats. Seriegewijze en postgewijze benaderingen bestaan in tal van varianten. In

een met de hand gevoerde administratie zullen de rekeningen bijvoorbeeld vaak direkt na het ontstaan van de mutaties worden opgezocht en bijgewerkt. Een zelfde gedragslijn kan worden gevolgd nadat voor het bijwerken van de rekeningen boekhoudmachines zijn ingevoerd. Het wordt in dat geval bij een toenemende omvang van de administratie echter aantrekkelijk de mutaties van frequent voorkomende soorten te verzamelen, op het gewenste aspect te sorteren en pas daarna te verwerken. Op die manier kan worden voorkomen dat de programmaliniaal veelvuldig moet worden verwisseld, terwijl met het bijwerken van de rekening terzelfdertijd een totaaloverzicht per mutatiesoort kan worden samengesteld. Over het algemeen komen bij gebruik van boekhoudmachines zowel postgewijze als seriegewijze bewerkingen voor. Met de invoering van ponskaarten apparatuur behoort de postgewijze benadering als systeem tijdelijk tot het verleden. De mutaties worden na codering naar soort en aspect gesorteerd en vervolgens seriegewijze verwerkt. Een dergelijke gedragslijn wordt eveneens gevolgd bij computersystemen waarin de bestanden op magneetband zijn vastgelegd. Een aangewezen rekening kan pas worden gemuteerd nadat alle voorliggende rekeningen die op dezelfde band werden vastgelegd zijn doorlopen. De opbouw van het bestand op een magneetband wijst in de richting van een seriegewijze benadering van dat bestand. Pas met de toepassing van direkt toegankelijke geheugens, waarin iedere rekening geheel onafhankelijk van de anderen kan worden benaderd, behoort de postgewijze bewerking weer tot de mogelijkheden. Wij zullen na deze korte illustratie de reeks van toepassingen tussen de uitersten van een seriegewijze en een postgewijze benadering in het kader van dit artikel verder buiten beschouwing laten.

Ons betoog voert tot de slotsom dat in het vraagstuk van de bestandsorganisatie, beoordeeld vanuit de noodzaak de inhoud van een bestand aktueel te houden, het *aantal* mutaties, de *aard* van de mutaties, de *tijdsfactor* en de *ruimtelijke spreiding* van de mutaties in hun onderlinge verband moeten worden bestudeerd. De technische realisatie kan daarbij variëren tussen de uitersten van serie- en postgewijze informatieverwerking.

In onze visie op het bijwerken van bestanden was voortdurend ook het raadplegen van die bestanden aan de orde. Wij wezen reeds op de sterke binding tussen beide activiteiten. Bij de beoordeling van een bestandsorganisatie vanuit de noodzaak om dat bestand te raadplegen spelen dan ook dezelfde factoren opnieuw hun rol. De systeemontwerper zal het aantal raadplegingen willen opsporen en tevens nagaan van welke *aard* en *omvang* - gemeten in bijvoorbeeld cijfers of letters - de informatie is die gevraagd wordt. Hij zal deze gegevens willen afwegen tegen de *tijd* waarop de informatie bekend moet zijn, rekening houdend met de *geografische plaats*. Technisch ligt de keuze opnieuw tussen de uitersten van seriegewijze en postgewijze raadplegen. Zowel ten aanzien van het bijwerken als het raadplegen van bestanden zullen „last but not least” de kosten moeten worden geanalyseerd. Dat daarbij vooral de tijdsfactor en de ruimtelijke spreiding een rol spelen ligt voor de hand. Een postgewijze verwerking met behulp van een computersysteem is in vele gevallen kostbaar, maar mogelijkwijze vanuit de functie die het bestand in het geheel van de organisatie vervult volledig

verantwoord. Veel hangt af van de beleidsnormen die door de leiding van een organisatie als uitgangspunt zijn gegeven. Wij zijn dan ook genoodzaakt ons in dit artikel te beperken tot het aangeven van de factoren die naar onze mening vanuit een functioneel gezichtspunt voor het organiseren van bestanden ten behoeve van automatische informatieverwerking van belang zijn.

Wij hebben ons in het voorgaande beperkt tot de ogenschijnlijk zelfstandige bestanden. Bezinning op de betekenis van de stuurfunctie maakt echter duidelijk dat voor het besturen vaak de relatie tussen de bestanden van evengroot belang is. In de Amerikaanse literatuur spreekt men van een common data base of een corporate data base (gegevensbank). In die gevallen worden verwijzingen tussen bestanden, zowel via objecten als aspecten, gerealiseerd zodat bij navraag ook de relaties tussen data kunnen worden gepresenteerd. Voor het hanteren van operationele research, simulatie en andere moderne bestuurstechnieken is die ontwikkeling van groot belang.

In een volgend artikel in deze reeks zal aan de technische aspecten van de bestandsorganisatie aandacht worden geschonken.