

Artikel 6

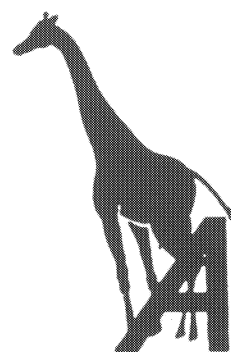
GIRAF

Een information retrieval systeem
voor universiteiten, hogescholen en
andere pluriforme organisaties

1984

*Het herkennen van de cultuur en het weten
hoe daarmee om te gaan zijn belangrijk om
te voorkomen dat de cultuur zich als een
virus gaat gedragen in de ontwikkelde
informatiesystemen.*

Mentink



Informatie, mei 1984, jaargang 26 nr. 6 pag. 445 t/m 455

Dit artikel is eerder verschenen in:

GIRAF

EEN INFORMATION RETRIEVAL SYSTEEM VOOR UNIVERSITEITEN, HOGESCHOLEN EN ANDERE PLURIFORME ORGANISATIES

door ir. A. J. van Dijk

In 1970 is bij de Technische Hogeschool Delft het (batch-)systeem 'BIBLIOSYSTEM' ontwikkeld. Over dit systeem is gepubliceerd in het blad 'Informatie' van oktober 1971. Hoewel het systeem bedoeld was om het samenstellen van bibliografieën met behulp van de computer sneller en prettiger te laten verlopen, is het in de periode 1970-1984 zowel binnen als buiten de TH-Delft ook gebruikt voor het bewerken van allerlei bestanden zoals adresbestanden, ledenbestanden, apparaatbestanden, diabeestanden, boekbestanden, bestanden met gegevens van congresgangers, etc. In 1976 ontwikkelde het Rekencentrum van de TH-Delft het on-line retrieval systeem 'BIBINFO'. Dit systeem sloot aan op 'BIBLIOSYSTEM'. De voortdurende vraag binnen de TH-Delft naar dergelijke faciliteiten heeft er toe geleid dat het, naar de normen van 1984, gebruiksvriendelijke information retrieval systeem 'GIRAF' is ontwikkeld. Bij de realisatie van 'GIRAF' heeft het database-management- en data-communicatie-systeem IDMS DB/DC als draagsysteem gefungeerd. Bovendien werd gebruik gemaakt van het ontwikkeltool 'Application Development System (ADS/ONLINE)' dat, evenals IDMS DB/DC, door de firma Cullinet Software geleverd wordt. Dit artikel behandelt de functionele aspecten van 'GIRAF'.

1 INLEIDING

In 1970 is bij de Technische Hogeschool Delft (kortweg TH-Delft) het systeem 'BIBLIOSYSTEM' [3] ontwikkeld. Aanleiding voor de bouw van dit systeem was de grote hoeveelheid tijd, die onderzoekers moesten besteden aan het verzamelen, registreren en ordenen van literatuur. Het bestuderen van literatuur kreeg nl. een steeds belangrijkere plaats in het onderzoek en het samenstellen van bibliografieën en literatuurlijsten werd onderdeel van vrijwel ieder researchproject van enige betekenis. Een grote flexibiliteit in het bijwerken en classificeren bleek daarbij van grote waarde.

De klassieke wijze van rapporteren van een literatuurstudie was zeer tijdrovend o.a. vanwege de omvang van het typewerk. Dit typewerk vormde een ware bottleneck. Bibliografieën waren soms op het moment dat ze gereed waren voor verspreiding alweer verouderd. Een en ander leidde tot het idee om het samenstellen van bibliografieën te automatiseren. 'BIBLIOSYSTEM' werd begin 1971 door de TH-Delft in gebruik genomen. Binnen de TH-Delft bleek het in een grote behoefte te voorzien. Maar niet alleen binnen de TH-Delft was men geïnteresseerd; na een publikatie in het blad 'Informatie' (oktober 1971) [3] kreeg de TH-Delft meer dan honderd verzoeken om de systeembeschrijving en programma's beschikbaar te stellen. In de loop der jaren bleek bovendien, dat door de 'abstracte' opzet het systeem bruikbaar was voor allerlei toepassingen. Naast bestanden met literatuurreferenties (zie figuur 2) werden met behulp van 'BIBLIOSYSTEM' adresbestanden, bestanden voor ledenadministraties, apparaatbestanden, diabeestanden, boekbestanden, etc. vervaardigd, bijgewerkt en afgedrukt.

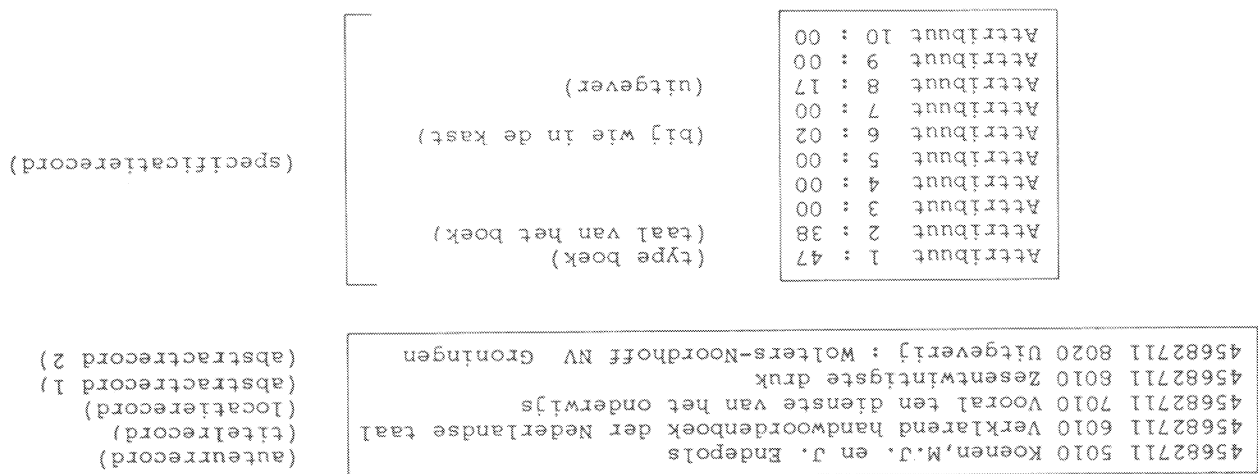
In 1975 werd het Rekencentrum geconfronteerd met het

verzoek van een congresorganisatiecomité om tijdens een te organiseren congres in de zomer van 1976 een real-time congresinformatiesysteem beschikbaar te stellen. Dit leidde tot de ontwikkeling van 'BIBINFO' [4]. Dit systeem werd als het ware bovenop 'BIBLIOSYSTEM' gebouwd. Het opbouwen van de bestanden gebeurde met 'BIBLIOSYSTEM'. Het raadplegen en het on-line selecteren gebeurde met 'BIBINFO'. In 1976 was het systeem voor de congresorganisatie beschikbaar. In de periode 1976-1983 is 'BIBINFO' regelmatig gebruikt door de 'BIBLIOSYSTEM'-gebruikers en door congresorganisaties.

'BIBLIOSYSTEM' (1970) en 'BIBINFO' (1976) zijn ontwikkeld voor een beperkte doelgroep. De systemen zijn ontwikkeld met methoden en hulpmiddelen, die in die jaren modern waren. Het grote succes van deze systemen binnen de TH-Delft en de voortdurende vraag naar dergelijke faciliteiten heeft de directie van het Rekencentrum in 1982 doen besluiten een nieuw information retrieval systeem te laten ontwikkelen, dat op een ruime doelgroep moest worden gericht. Het moest de systemen 'BIBLIOSYSTEM' en 'BIBINFO' vervangen en zeer gebruiksvriendelijk zijn. Dit nieuwe systeem is het *programmatuursysteem* 'GIRAF' (General Information Retrieval Facilities) geworden.

In dit artikel zullen de *functionele* aspecten van 'GIRAF' worden behandeld. De technische aspecten komen slechts zeer zijdelings ter sprake. In een volgend artikel zal worden ingegaan op enkele technische aspecten van 'GIRAF'. Bovendien zal in dat artikel worden ingegaan op de ervaringen opgedaan met conversationale programmering ('BIBINFO' is volledig conversationeel geprogrammeerd), pseudo-conversationele programmering [5] en programmering met behulp van een 'ontwikkeltool' ('GIRAF' is ontwikkeld met behulp van het ont-

Figuur 2: Voorbeeld van een 'BIBLIOSYSTEM'-referentie



attribuut een waarde uit de betreffende waardeverzamel-
 ming worden toegekend. Op deze manier kon een ge-
 bruiker aan een referentie kenmerken en kenmerkwaar-
 den toekennen. Aan een boekreferentie kon bijvoor-
 beeld in attribuutnummer 2 de waarde 38 en in attribuut-
 nummer 8 de waarde 17 worden toegekend (zie figuur 2).
 Een en ander betekende dat het boek in de Engelse taal
 was geschreven en door een bepaalde uitgever was uitge-
 geven. Het selecteren van een deelverzameling was te
 realiseren door de attributen, waarin men voor de betref-
 fende selectie was geïnteresseerd, op te geven. Per attri-
 buut werd de waardeverzameling, die voor deze selec-
 tie geldig was, opgegeven. Met behulp van EN- en OF-
 relaties op attribuutniveau was het mogelijk de gewenste
 deelverzameling te construeren.

Aangezien 'GIRAF' voor een ruimere doelgroep moest
 worden gemaakt, moest het meer algemeen bruikbaar
 zijn dan de oude systemen (eis 2). Daaruit werd, mede
 naar aanleiding van verzoeken van gebruikers, de con-
 clusie getrokken, dat het codesysteem ruimer zou moe-
 ten zijn (eis 3).

De TH-Delft is een pluriforme organisatie, die veel zelf-
 standig opererende groepingen herbergt. Sommige
 van die groepingen willen (delen van) hun informatie-
 voorziening op een geïsoleerde manier realiseren (bijv.
 promovendi) (eis 4). Andere groepingen willen de TH-
 Delft. Voor die groepingen is het noodzakelijk om
 vanuit verschillende locaties, met behulp van real-time-
 verwerking, bestanden te kunnen opbouwen, raadple-
 gen en/of muteren (eis 5).

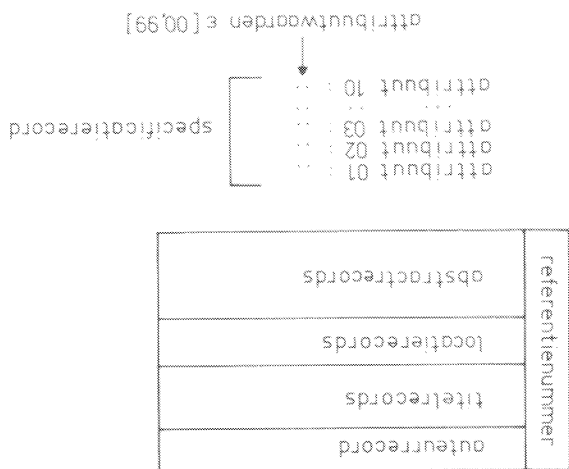
'GIRAF' moet kunnen worden toegepast ten behoeve
 van veel informatiesystemen binnen de TH-Delft. Daar-
 om moet 'GIRAF' worden aangeboden als *confectiesys-
 tem*. De opzet moest echter dusdanig zijn, dat het met
 weinig inspanning kan worden aangepast. Op die manier
 kan 'GIRAF' voor sommige gebruikers als *maatconfec-
 tiesysteem* fungeren (eis 6).

Bij de TH-Delft werken en studeren ook buitenlanders.
 Het is daarom wenselijk dat 'GIRAF' kan werken met
 andere talen. Het systeem moest zodanig worden opge-
 zet, dat in een later stadium *andere taalversies* snel kun-
 nen worden geïmplementeerd (eis 7).

Het Rekencentrum van de TH-Delft is een dienstverle-
 nende instantie. Zoals het een goede dienstverlenende

referentie geschiedde via het referentienummer of de
 auteur. Toegang tot een *verzameling* referenties was al-
 leen mogelijk via een codesysteem. Dit codesysteem be-
 stond uit een tiental attributen met voor ieder attribuut
 een waardeverzameling. Aan iedere referentie kon per

Figuur 1: Gegevensmodel van een referentie binnen 'BIBLIOSYSTEM' (1970)



BIBLIOSYSTEM (1970)

Uit het succes van de systemen 'BIBLIOSYSTEM' en
 'BIBINFO' is de conclusie getrokken, dat de filosofie,
 waarop deze systemen waren gebaseerd, voor de TH-
 Delft een goede filosofie is. De eerste voorwaarde, die
 aan 'GIRAF' is gesteld, is het handhaven van deze filoso-
 fie. Dit betekende o.a., dat 'GIRAF' een eenvoudig en
 doeltreffend informatie retrieval systeem moest wor-
 den (eis 1).

De gegevensseenheid in de oude systemen was de refe-
 rentie (zie figuren 1 en 2). Toegang tot een *individuele*

2 UITGANGSPUNTEN 'GIRAF'

wikkelrool 'Application Development System (ADS/
 ONLINE)' van de firma Cullinet Software [1, 2]).

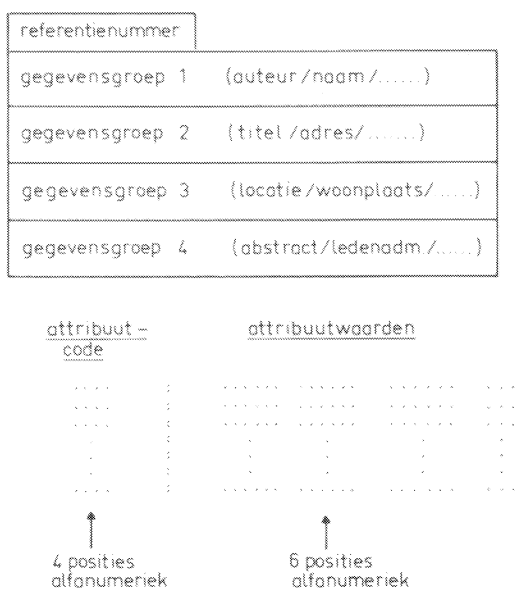
instantie betaamt, geeft zij een hoge prioriteit aan de belangen van haar huidige gebruikers. 'GIRAF' diende daarom voorzieningen te hebben om de gebruikers van 'BIBLIOSYSTEM' en 'BIBINFO' op een zeer soepele wijze te kunnen laten overstappen naar het nieuwe systeem (eis 8).

Door het Rekencentrum van de TH-Delft is als draagsysteem ten behoeve van database-management- en datacommunicatie-activiteiten gekozen voor IDMS DB/DC, dat wordt geleverd door de firma Cullinet Software. 'GIRAF' diende met behulp van IDMS DB/DC te worden ontwikkeld (eis 9). 'GIRAF' kan daardoor in principe functioneren op elke computer waarop een recente versie van IDMS DB/DC is geïnstalleerd. Vaak zal dat een computer uit de serie IBM 4300 zijn, maar ook andere computers kunnen werken met IDMS DB/DC. Het Rekencentrum van de TH-Delft heeft 'GIRAF' operationeel op een Amdahl 470V/7B computer.

3 INDELING VAN DE REFERENTIES

De eenheid in de systemen 'BIBLIOSYSTEM' en 'BIBINFO' was de referentie. Hoewel oorspronkelijk bedoeld voor het behandelen van literatuurreferenties (zie figuur 2) werden de systemen voor allerhande toepassingen gebruikt. De naamgeving was echter gebaseerd op literatuurreferenties. Zo was er sprake van een auteur, een titel, een locatie en een abstract. Daarnaast werd aan een referentie voor een aantal (maximaal 10) attributen een attribuutwaarde toegekend. Aangezien de term 'referentie' bij veel gebruikers van de oude systemen is ingeburgerd, is besloten deze term in 'GIRAF' over te nemen. 'GIRAF' moest een algemeen karakter hebben, daarom kon niet worden volstaan met de bestaande naamgeving. Binnen 'GIRAF' bestaat een referentie uit vier gegevensgroepen (zie figuur 3). Iedere gebruiker die verantwoordelijk is voor een gegevensverzameling (in

GIRAF (1984)



Figuur 3: Gegevensmodel van een referentie binnen 'GIRAF' (1984)

'GIRAF'-termen: iedere ToepassingsBestandsBeheerder (TBB)) kan zelf de namen kiezen, die hij aan deze gegevensgroepen wenst te geven. Overal waar de gegevensgroepen worden gebruikt (op beeldschermen of papier) worden de door de TBB opgegeven namen gebruikt. 'GIRAF' kent zelf aan iedere nieuwe referentie een referentienummer toe. Het codesysteem, waarmee attribuutcodes en attribuutwaarden aan een referentie kunnen worden toegekend, is uitgebreid (zie figuren 3 en 4). Het aantal attribuutcodes is nu onbeperkt. Bovendien kunnen per referentie aan een attribuutcode verschillende alfanumerieke attribuutwaarden worden toegekend.

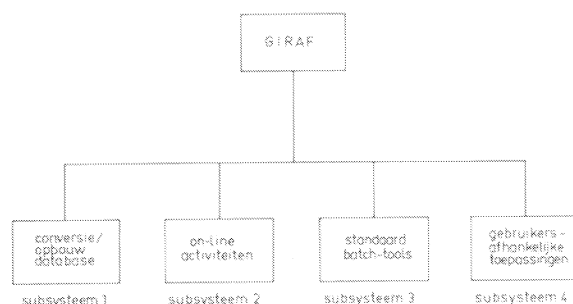
| | | |
|--------------------------------|---|--|
| 00000427 | | |
| Roos, R. de | | |
| Speelgrageingel 392 Hofje 3 | | |
| 2689 KL Delft bij Een Haag | | |
| Betaling | : | behoeft niet meer dan fl. 110,- te betalen |
| Telefonisch bereikbaar | : | 079-896157 |
| Organisatie | : | is in de toekomst beschikbaar voor een bestuursfunctie |
| Bijzonderheden | : | spreekt chinees |

| Attribuutcode | Attribuutwaarde | |
|---------------|-----------------|----------------------------|
| BETA | B4-2 | (betaling) |
| INST | PIANO | (bespeelt instrument) |
| INST | GITAAR | |
| INST | KLARNET | |
| CURS | DRUMS | (volgt cursus) |
| CURS | JAZZ | |
| TAAL | REC | (spreekt taal) |
| TAAL | ENG | |
| TAAL | CHIN | |
| STAT | TH-MED | (status) |
| JAAR | 1980 | (jaar van de inschrijving) |

Figuur 4: Voorbeeld van een 'GIRAF'-referentie

4 SUBSYSTEMEN

Het programmatuursysteem 'GIRAF' is opgedeeld in 4 subsystemen (zie figuur 5), te weten:
 subsysteem 1: conversie/opbouw database
 subsysteem 2: on-line activiteiten
 subsysteem 3: standaard batch-tools
 subsysteem 4: gebruikersafhankelijke toepassingen
 Allereerst worden nu de subsystemen 1, 3 en 4 kort toegelicht. Daarna komt subsysteem 2 aan de orde.



Figuur 5: Subsystemen 'GIRAF'

4.1 Conversie/opbouw database (substelsysteem 1)

Met behulp van dit subsysteem kunnen bestanden in 'BIBLIOSYSTEM'-formaat worden geconverteerd naar een 'GIRAF-Toepassings-Database' (GTD). Met behulp van stuurinformatie kan tot op zekere hoogte rekening worden gehouden met individuele conversiewensen. Gebruikers die in het bezit zijn van een gegevensverzameling, die niet met 'BIBLIOSYSTEM' is opge-

