

# Stofnesten en beeldschermen. De stand van zaken in de historische informatiekunde

*Jeroen Touwen*

## Rekenschap van het verleden

Computer betekent letterlijk rekenaar of rekenapparaat.<sup>1</sup> Wat moet de geschiedschrijving, een vak dat van oudsher bij de Letterenfaculteit thuishoort, met een rekenapparaat? Is het niet zo dat bij de bestudering van het verleden, of het verhaal dat de historicus vertelt, meten en rekenen slechts een kleine plaats in nemen? Verschillende opvattingen op methodologisch vlak impliceren een grotere of kleinere rol voor het tellen en turven, ofwel het stellen van kwantitatieve vragen. Bij het beantwoorden van die vragen kan de computer van dienst zijn. Maar het rekenapparaat bewijst ook op ander terrein zijn nut. Geschiedschrijving bestaat uit de visie van de historicus op het verleden, ondersteund met de door haar of hem waardevol geachte selectie van feiten. Deze selectie van feitenmateriaal stoelt op de systematische verwerking van informatie. De moderne technologie blijkt uiterst waardevol te zijn bij het zoeken en hanteren van deze informatie.

Nu is anno 1993 vrijwel iedereen het erover eens dat door de moderne informatietechnologie veel tijdovende arbeid overbodig wordt. Alleen al bij het begin- en eindpunt van het historisch ambacht heeft de computer zijn nut bewezen: het zoeken in een geautomatiseerde bibliotheekcatalogus of het aanvragen met behulp van de computer in het Algemeen Rijksarchief enerzijds, het schrijven met behulp van een tekstverwerker anderzijds. Maar op welke gebieden kunnen wij de computer nog meer inzetten?

Deze bijdrage geeft een overzicht van de verschillende toepassingsmogelijkheden van de computer in de geschiedenis. We onderscheiden daarbij vier hoofdonderdelen:<sup>2</sup>

- heuristische hulpmiddelen,
- archiefontsluiting met behulp van de computer,
- het analyseren van bronmateriaal (dataverwerking) en
- de uitwisseling van informatie.<sup>3</sup>

Nadat deze vier hoofdonderdelen aan bod zijn gekomen, wordt kort ingegaan op de historiografie. Welke literatuur is er verschenen op het gebied van de historische informatiekunde? In welke tijdschriften kan men te rade gaan voor specifieke vragen op dit gebied? Tenslotte komt de vraag aan de orde wat er in de nabije toekomst verwacht kan worden van deze hulpwetenschap.

## Toepassingsmogelijkheden van nieuwe informatietechnieken

De genoemde vier verschillende toepassingsgebieden van de computer in de geschiedschrijving bestrijken het hele historische ambacht. We volgen de historicus als het ware op zijn of haar tocht op zoek naar literatuur, op zoek

naar bronnen in het archief, bij het verwerken van het bronmateriaal tijdens de fase van analyse en tenslotte bij het rapporteren over bevindingen en conclusies. Maar laten wij allereerst stilstaan bij de vraag wat historische informatiekunde nu precies is.

Historische informatiekunde kan worden gedefinieerd als de historische specialisatie die zich bezighoudt met het toepassen van geautomatiseerde systemen (computers) voor de opslag, bewerking en bevraging van historische gegevens. Het gaat hier dus niet om informatica, in welk vak het inzicht in het informatie-systeem centraal staat, maar om de *toepassing* van de technieken in het onderzoeksgebied.<sup>4</sup> Men kan daarbij denken aan een veelheid van toepassingen, van elektronische kaartenbakken tot statistiek en tekstverwerking. Sinds de jaren zestig en zeventig, toen grote, ontoegankelijke mainframe systemen werden gebruikt voor gecompliceerde statistische berekeningen (voornamelijk van sociaal-economische historici), is er een aanzienlijke uitbreiding opgetreden. Nieuwe mogelijkheden om bestanden te sorteren of te doorzoeken, gebruiksvriendelijke database programma's en zelfs tekstanalyse-technieken die ontwikkeld zijn door linguïsten en literatuurwetenschappers worden tegenwoordig in historisch onderzoek gebruikt.<sup>5</sup>

Boonstra onderscheidt drie fasen in de historische informatiekunde.<sup>6</sup> Aanvankelijk ging de aandacht vooral uit naar de *software*. Historici beschreven hoe bepaalde programmatuur in hun onderzoek kon worden aangevend. Al snel verschoof de aandacht naar de *bronnen*: hoe kon een historisch informatiesysteem worden ontwikkeld dat rekening hield met inconsistenties en eigenaardigheden van een bron? Het probleem was hier dikwijls, dat als een grote hoeveelheid materiaal in de computer was ingevoerd, de bevraging daarvan pas mogelijk was als bijvoorbeeld de standaardisatie van namen, uiteenlopende notaties van de tijd, of soms zelfs verschillende tijdrekeningen uit de weg waren geruimd. Het gebruikte computerprogramma moet met deze zaken rekening houden. In de laatste fase gaat men nog een stap verder met deze problematiek en richt de aandacht zich op de *informatieproblemen*. In deze fase worden de specifieke informatie-aspecten van invoer, opslag en vergelijking bestudeerd en benaderd met behulp van speciale algoritmen. Pas recentelijk is deze fase ingetreden.

Thaller, die een speciaal database programma heeft ontwikkeld voor historici, heeft in een aantal artikelen gepleit voor het opzetten van een 'theorie van historische informatiekunde'. Deze theorie zou een sterk bron-georiënteerde onderzoeksmethodologie moeten vormen.<sup>7</sup> Anderen stellen echter dat een eenzijdige nadruk op bron-problemen de aandacht afleidt van de onderzoeksvraag, en zal leiden tot het bouwen van fraaie historische informatiesystemen zonder dat de praktische bruikbaarheid daarvan gewaarborgd is.<sup>8</sup> Ook bij sommige fraaie demonstratieversies van (kunst)historische database-applicaties dringt deze kwestie zich op.<sup>9</sup>

Een gevolg van wat men zou kunnen noemen de toenemende *informatisering* van de historische informatiekunde, is dat er twee opvattingen zijn ontstaan over het vak: in de ene visie richt de historische informatiekunde zich vooral op de informatica, om methoden en algoritmen te ontwikkelen die

gebruikt kunnen worden bij historisch onderzoek. Dit is een sterk specialistisch gebied: het gaat bijvoorbeeld om modellen en schema's die recente informaticatechnieken vertegenwoordigen. In de andere opvatting is de historische informatiekunde een breed vakgebied, waarbij het gaat om alle mogelijke bruikbare toepassingen van de informatica voor het historisch onderzoek, met speciale aandacht voor het gebruik van historische bronnen. In dit artikel wordt uitgegaan van deze brede omschrijving van de historische informatiekunde.

### **Heuristische hulpmiddelen**

Wij keren terug naar de historicus en zijn onderzoek. De eerste fase betreft uiteraard het literatuurzoeken. De afgelopen jaren zijn de aard en de omvang van het heuristische apparaat ingrijpend veranderd. In Nederland zijn alle universitaire bibliotheken inmiddels geautomatiseerd. Dit maakt het mogelijk om in een bibliotheekcatalogus zeer snel een boek te vinden en dat meteen aan te vragen. Men kan zoeken op titel of auteur, maar men kan ook, na een korte verkenning van de mogelijkheden van het menusysteem, via combinaties van zoekopdrachten onderzoeksmateriaal vinden dat anders veel moeilijker vindbaar was. Een recente ontwikkeling is het koppelen van de verschillende bibliotheekbestanden, zodat ook elders aanwezige titels kunnen worden gelocaliseerd.<sup>10</sup>

Ook thematische bibliografieën, die vroeger in uitdijende reeksen en met steeds vernieuwde cumulatieve indexen op de plank kwamen te staan, hebben geprofiteerd van de komst van de digitale technologie. Recenter dan de automatische catalogus is de komst van de CD-ROM: Dikwijls verschijnen nu thematische bibliografieën op CD-ROM. Een bekende CD-ROM collectie op historisch gebied is *Historical Abstracts*. Deze databank bevat meer dan een half miljoen publikaties (boeken en artikelen). Hoewel als kritiek wel wordt aangevoerd dat deze databank (a) een oververtegenwoordiging van Engelstalige literatuur bevat en (b) de oudere publikaties nooit zal aanbieden en dus altijd onvolledig zal zijn, biedt een dergelijk systeem een schat aan verwijzingen.<sup>11</sup> Zo kan ook een korte zoektocht in *Current Contents*, een bibliografische databank die elk kwartaal verschijnt, voor de historicus nuttige literatuur opleveren. Dikwijls worden dergelijke diensten tevens via centrale *on-line* databanken aangeboden (waar een bibliotheek of gebruiker via het computernetwerk of per telefoonlijn contact mee kan maken) en soms ook nog in de oude vertrouwde, papieren vorm. Zo is er de Biografie van Moderne Europese Geschiedenis (BIMEG, onlangs omgedoopt tot HinT: History in Titels), welke zowel op papier als per computer raadpleegbaar is. Een ander voorbeeld is de elektronische bibliografie met publikaties over de historische informatiekunde: deze is verkrijgbaar op diskette (of op te halen via het netwerk) en bevat een grondige, recent bijgewerkte bibliografie van dit vakgebied.<sup>12</sup>

Hoewel het bedrukte papier waarschijnlijk altijd belangrijk zal blijven, wordt door de nieuwe mogelijkheden van informatieontsluiting de oorspronke-

lijke functie van de bibliotheek bedreigd. Daarom zijn veel universiteitsbibliotheken bezig zich vergaand te oriënteren in de informatiekunde. Een sprekend voorbeeld hiervan is het feit de in het maandelijks (roze) bulletin van het Leidse rekencentrum de UB een meerbladig (blauw) katern vult.<sup>13</sup>

### Archiefontsluiting met behulp van de computer

Ook in de archieven treft men alweer een aantal jaren de informatietechnologie aan. Tussen de dozen met stoffige folio's knippen de beeldschermen. Het aanvragen van archiefstukken met de computer werd reeds genoemd. Wanneer het gaat om het raadpleegbaar maken van archiefmateriaal, waarbij machineleesbare bestanden worden aangelegd, kunnen we tegelijkertijd twee ontwikkelingen signaleren: vanuit de aanbodzijde (de archiverende instelling) en vanuit de vraagzijde (de historicus). Ten eerste zijn archieven en documentatie-instellingen in toenemende mate bezig om datacollecties via elektronische weg raadpleegbaar te maken voor de gebruiker. Een tot de verbeelding sprekend voorbeeld is het *Archivo General des Indias* in Sevilla.<sup>14</sup> Dit archief, dat in 1785 werd opgericht door de Spaanse koning Karel III, zal in zijn geheel worden gescand, zodat de documenten op beeldschermen geraadpleegd kunnen worden. Reeds 9 miljoen van de ca. 90 miljoen documenten zijn inmiddels gedigitaliseerd. Een soortgelijk project zal de komende jaren worden uitgevoerd in de Vaticaanse Bibliotheek, waarbij een groot deel van de collectie, bestaande uit boeken, manuscripten, prenten en miniaturen zal worden gedigitaliseerd.<sup>15</sup> Ook Nederlandse archieven bestuderen mogelijkheden om via scanning of CD-ROM's datacollecties te ontsluiten. Zo is de fotocollectie van het Internationaal Instituut voor Sociale Geschiedenis in Amsterdam geheel geautomatiseerd. Analoog aan het bibliotheekstelsel kunnen de circa 100.000 foto's uit de collectie worden opgezocht en op het scherm worden getoond.

De tweede ontwikkeling, het aanleggen van datacollecties door de historicus, is reeds langere tijd gaande. Het gaat hier om datasets die worden aangelegd ten behoeve van het eigen onderzoek. Wanneer het onderzoek is afgerond en de resultaten zijn gepubliceerd, blijft een database over die onder bepaalde voorwaarden voor ander onderzoek beschikbaar kan worden gesteld. In feite vormt het materiaal dan een bronnenuitgave. Wanneer een bron of een gedeelte van een bron correct, systematisch en gestructureerd in de computer is ingevoerd, kunnen in principe vele andere onderzoekers gebruik maken van de machineleesbare versie van de bron. Voordelen van het secundair gebruik van databestanden, hetgeen in de Sociale Wetenschappen al decennia voorkomt, zijn kostenbesparing, controleerbaarheid van onderzoek, beschikbaarheid van het materiaal voor onderwijsdoeleinden en het toepassen van nieuwe methodes op het materiaal. Machineleesbaar historisch materiaal verliest zijn waarde niet. De datacollectie kan als een publikatie worden gezien waarbij het volledige intellectuele krediet toekomt aan de oorspronkelijke onderzoeker.<sup>16</sup>

Een belangrijk onderscheid dat hier moet worden gemaakt is dat tussen het invoeren van totale teksten en het invoeren van geselecteerde gegevens (bijvoorbeeld op politiek-, economisch- of sociaal-historisch gebied) in een database. In het eerste geval wordt een *tekst-corpus* opgebouwd, dat een gehele oorspronkelijke bron bevat, dikwijls nog aangevuld met allerlei markeringen die analyse mogelijk maken. Dergelijke projecten, zoals Amerikaanse initiatieven om materiaal met betrekking tot de Amerikaanse en Franse Revoluties in te voeren, worden dikwijls uitgevoerd in samenwerking met onderzoekers uit de linguïstiek.<sup>17</sup> De tweede vorm komt vaker voor en betreft meestal het invoeren van een selectie van brongegevens aan de hand van een specifieke onderzoeksvraagstelling. Bij het selectief invoeren van gegevens ontstaat onder autoriteit van de onderzoeker een zogenaamde *meta-source*, waarin al dan niet gecodeerd of gestandaardiseerd een selectie van brongegevens bijeengebracht zijn.<sup>18</sup>

Een voorbeeld van een databank met een nog specifiekere thematische afbakening is de *Medieval and Early Modern Data Bank* (MEMDB). Deze databank biedt een zorgvuldig geselecteerde verzameling Middeleeuwse prijzen en wisselkoersgegevens en is recentelijk uitgebracht op CD-ROM. De brongegevens, oorspronkelijk afkomstig uit een hele reeks Europese archieven en ter beschikking gesteld door verschillende onderzoekers, zijn zorgvuldig geannoteerd. Op deze wijze ontstaat een specialistisch informatiesysteem, een *tool* waarmee een mediaevist in korte tijd waarden of prijzen uit zijn eigen brongegevens kan vergelijken met of converteren naar andere valuta in pre-industrieel Europa.<sup>19</sup>

De historicus kan in zijn zoektocht naar materiaal behalve het archief en de bibliotheek tegenwoordig ook terecht in het data-archief. Data-archieven of 'data libraries' vervullen een belangrijke rol bij het beschikbaar stellen van databases en text-corpora voor *secondary analysis*. Hier worden in bewaring gegeven datasets gedocumenteerd en gecatalogiseerd, en wordt gezorgd voor verantwoorde opslag. Ook bieden data-archieven assistentie bij de conversie van de datasets naar andere opslagformaten. In toenemende mate spelen zij een bemiddelende rol als 'information broker' binnen de wereld van informatiezoekers en informatie-aanbieders. In Nederland bestaat sinds 1989 het Nederlands Historisch Data Archief, dat gevestigd is bij de Vakgroep Geschiedenis van de Rijksuniversiteit Leiden.<sup>20</sup>

Dat in Nederland het gebruik van de databases door historici sterk is gestegen, blijkt uit de verschillende inventarisaties die zijn gehouden. Zo leverde in 1987 een landelijke enquête die historici om bestandsbeschrijvingen vroeg, 81 responsies op. In 1989 waren dat er al 131.<sup>21</sup> Het NHDA werkt momenteel aan een uitgebreide catalogus van Nederlandse databestanden. Ook goed bruikbaar is de door Schürer en Anderson samengestelde catalogus van Engelse historische databestanden. Deze gids bevat beschrijvingen van 376 Britse datasets, maar bevat ook beknopte overzichten van circa 350 databestanden in Oostenrijk, Denemarken, Duitsland, Noorwegen, Nederland, Hongarije, Zweden en de Verenigde Staten.<sup>22</sup>

Zowel archiefinstellingen die hun collecties willen digitaliseren, als individuele onderzoekers die hun bronnen met de computer willen bewerken, krijgen te maken met de enorme tijd die het kost om bronnen in te typen. Historici die met een draagbare computer in het archief werken, zullen dit bevestigen. In de techniek zoekt men al een jaar of vijftien naar bruikbare alternatieven, waarbij de laatste jaren vooral optische scanning voortgang boekt. Documenten worden met behulp van een soort fotokopieerapparaat gescand, waarna het ontstane plaatje in het geheugen van de computer wordt 'gelezen' met behulp van een speciaal programma. Men noemt deze programmatuur Optical Character Recognition (OCR)-software. Met steun van NWO is bij het Nederlands Historisch Data Archief een speciaal Scan-OCR-laboratorium ingericht, waar diensten worden aangeboden om documenten te digitaliseren. Ook opslag- en ontsluitingsproblemen krijgen hier de aandacht.<sup>23</sup>

De landelijke informatie-infrastructuur, bestaande uit computer-netwerken en centrale computerdiensten, maakt momenteel een snelle groei door. Zo heeft de Leidse Universiteitsbibliotheek eind vorig jaar een speciale *documentserver* in gebruik genomen. Dit is een speciale computer voor de opslag van elektronische documenten, die via het netwerk beschikbaar zijn voor de universitaire gebruiker. Om landelijke uitwisseling mogelijk te maken stelt SURFnet, het landelijke netwerk voor communicatie- en informatiediensten in onderzoek en hoger onderwijs, regels op. Het document-ontsluitingsproject van de Leidse Universiteitsbibliotheek heet OASIS. De OASIS-server bevat onder andere een aantal bestanden van het Nederlands Historisch Data Archief.

Ook elders in Nederland worden dergelijke projecten ontwikkeld.<sup>24</sup> Grote dataprojecten die genoemd kunnen worden zijn de Historisch-Ecologische Databank (HED), de Historische Steekproef Nederlandse Bevolking (HSN), de Databank Noord-Brabant (DNB) en een kartografische databank van de wijzigingen in de gemeentegrenzen in de afgelopen eeuwen (NL-KAART). Deze projecten voeren overleg om in de toekomst te komen tot een 'Nederlandse Geschiedenismachine' (of Dutch History Machine). Een dergelijke aggregatie maakt de toegankelijkheid van de dataverzamelingen voor de historicus natuurlijk groter. Informatie over bovengenoemde en buitenlandse databaseprojecten kan verkregen worden bij het Nederlands Historisch Data Archief.

Wat betreft archiefontsluiting kan tenslotte opgemerkt worden dat er zich in de nabije toekomst ook in de 'gewone' archieven een verandering zal voltrekken. Grote hoeveelheden materiaal zullen namelijk bij het Algemeen Rijksarchief worden aangeleverd in elektronische vorm, omdat in alle overheidsinstellingen inmiddels met tekstverwerkers, lokale netwerken en *electronic mail*-systemen wordt gewerkt.<sup>25</sup> De opslagcapaciteit van hard-disks, computertapes en CD-ROMs is enorm veel groter dan de traditionele kast met dozen, waardoor grote hoeveelheden materiaal kunnen worden opgeslagen. Een archief zal daardoor zeer grote collecties archiveren, waarvan het ontsluiten via een index of inventaris onbegonnen werk is. De machineleesbare vorm leent zich echter voor een andere aanpak: de teksten zijn

namelijk elektronisch doorzoekbaar. Met een elektronisch zoekprogramma kunnen in één ochtend in plaats van vijf dozen archiefmateriaal het equivalent van vierduizend dozen worden doorzocht op sleutelbegrippen, jaartallen, thema's of eigennamen. Nadeel hierbij is natuurlijk wel dat de vangst heel groot is en er sprake is van veel ruis. Ook zullen back-up tapes van netwerken veel dubbele versies van tekstbestanden bevatten, hoewel deze bij het archiveren grotendeels geautomatiseerd kunnen worden verwijderd. De nationale archieven van Zweden en Canada hebben reeds over de mogelijkheden en problemen van dergelijke nieuwe archiveringsmethoden gerapporteerd. In Canada zullen historici binnenkort alle documenten afkomstig van de onderhandelingen over het recent gesloten NAFTA-verdrag elektronisch kunnen doorzoeken.<sup>26</sup>

Dergelijke ontwikkelingen zetten de huidige discussie over archiefverniëting in een nieuw licht. In het recente discussiedossier over archiefverniëting in de *Bijdragen en Mededelingen betreffende de Geschiedenis der Nederlanden* besteedt gemeentearchivaris W.Chr. Pieterse aandacht aan de MLG's (machine-leesbare gegevensbestanden) vanuit de optiek van selectie en vernietiging van archieven. Zij stelt: 'De problematiek die bij de selectie en vernietiging van deze archieven een rol speelt, is vele malen groter dan waar we nu met onze papieren archieven mee hebben te kampen. Misschien is het beter al onze krachten aan die problematiek te wijden (...).'<sup>27</sup> De toekomst zal moeten uitwijzen of deze problemen op zodanige wijze het hoofd kan worden geboden dat de historicus er voordeel van heeft.

### **Het analyseren van bronmateriaal (dataverwerking)**

Wanneer historisch bronmateriaal eenmaal in de computer is opgeslagen, doen zich een groot aantal mogelijkheden voor verwerking en analyse voor. Het gaat hier niet alleen om het berekenen van statistische maten en tabellen, maar ook om andere bewerkingen die met de hand arbeidsintensief en tijdrovend zijn: sorteren, koppelen, zoeken, en classificeren.

Het zijn deze bewerkingen die dikwijls centraal staan in de aandacht van de computergebruikende historicus. In abstractere zin gaat het hier in feite om *methoden*, die door de computer worden aangereikt, terwijl het bij de overige hoofdonderdelen over het algemeen gaat om *hulpmiddelen*.

Als voorbeeld kijken wij naar een project waarbij historische Leidse criminele vonnissen worden ingevoerd in de computer.<sup>28</sup> Reeds geruime tijd wordt in het Leidse Gemeentearchief gewerkt aan de invoer van vonnisboeken uit de periode 1811-1838. In een eerder stadium werden de vonnissen uit de periode 1600-1811 gecodeerd ingevoerd. In de fase van dataverwerking zouden de volgende stappen kunnen voorkomen:

(a) het oudere, numerieke bestand en het numerieke gedeelte van het nieuwe bestand kunnen worden gekoppeld. Wellicht moeten beschrijvingen van straffen, bepalingen van tijdsduur en namen worden gestandaardiseerd. De

betekeningen van de gecodeerde variabelen zouden aan het bestand kunnen worden toegevoegd. Extra velden kunnen worden aangelegd om aggregatieniveaus of categorieën aan te geven.

(b) Nu kunnen over een lange periode optellingen worden verricht, frequenties van straffen worden berekend, aantallen veroordelingen bekeken, etc.

(c) Voor de analyse van soorten straffen kan de database worden gesorteerd per categorie, en binnen categorie naar zwaarte van straftoekenning of naar tijd. Samenvattende tabellen en grafieken waarin opvallende ontwikkelingen naar voren komen kunnen worden uitgedraaid.

(d) De ingevoerde teksten van de vonnissen uit de periode 1811-1838 kunnen worden geanalyseerd met behulp van woordconcordanties en andere tekstanalyse-methoden.

Deze stappen geven in een notedop enkele mogelijkheden weer van een dergelijk computerbestand. De resultaten zijn behalve van de gestelde vragen uiteraard ook afhankelijk van het soort bronmateriaal. Het ligt daarom voor de hand om ons af te vragen welke soorten bronnen binnen de geschiedwetenschap nu het meest dikwijls met de computer worden verwerkt.

Aanvankelijk waren het voornamelijk sociaal-economisch historici (waaronder de cliometristen), die met hulp van computerprogrammeurs gecompliceerde statistische bewerkingen lieten uitvoeren. In de jaren zeventig kwamen pakketten als SPSS (statistiek) en OCP (concordanties) op de markt, waarmee men zelf, zij het met omslachtige hulpmiddelen als ponskaarten, berekeningen kon doen. Tot 1987 had het meeste historisch onderzoek waarbij de computer werd gebruikt betrekking op de negentiende en twintigste eeuw, dit waren dus hoofdzakelijk sociaal-economische onderzoeksprojecten. Sinds die tijd is er een bredere verdeling ontstaan van de toepassing van de computer binnen de geschiedwetenschap. Er kwam meer nadruk op documenttoepassingen, waarbij het ging om de opbouw, ordening en raadpleging van gegevens. Problemen rondom bestandsstructuren en 'brongetrouwe' data-invoer krijgen hierbij relatief meer aandacht, naast de traditionele aandacht voor analyse van de data. Men gaat vaker tekst invoeren, meestal in combinatie met het kwantitatieve materiaal. Het grootste gedeelte van de onderzoeken richt zich op de nieuwe en nieuwste tijd, maar inmiddels hebben de Middeleeuwen en de vroeg-moderne periode (1500-1800) een aandeel van ongeveer twintig procent. Hoewel door de aard van de brongegevens sociaal-economische aspecten zich over het algemeen beter lenen voor analyse met behulp van de computer, valt er een toename waar te nemen van tekst-analyse en automatische record-linkage (het koppelen van informatie uit verschillende bronnen).<sup>29</sup>

In het algemeen zou men kunnen stellen, dat veel historici tegenwoordig van mening zijn dat de kwalitatieve, literaire geschiedbeoefening en de kwantitatieve, cliometrische geschiedbeoefening elkaar niet uitsluiten maar dienen aan te vullen. Dit heeft tot gevolg dat het gebruiken van ondersteunend materiaal uit een database of het uitvoeren van een speciale analyse met behulp van de computer bij allerlei onderzoek wordt toegepast. Een recent voorbeeld van een methodologische 'eye-opener' is het proefschrift van J. Brefeld.<sup>30</sup> Zij

analyseerde met behulp van de computer een aantal reisverslagen van Middeleeuwse pelgrims naar Jeruzalem en constateerde dat voor de pelgrims een soort 'Guide Michelin' bestond, die een route langs de heilige plaatsen beschreef en waaruit dikwijls informatie werd overgeschreven. Door middel van factor-analyse kon zij bepalen welk reisverslag de oorspronkelijke reisgids het beste benaderde. Dit is slechts één voorbeeld van een creatieve toepassing van de computer op een traditioneel historisch, niet-kwantitatief onderwerp. Te verwachten is dat bij vele historische onderzoeksthema's de komende decennia aanvullende analyses mogelijk worden met nieuwe methoden met behulp van de computer. Het 'traditionele onderzoek' zal hierdoor worden verrijkt.

### **Informatie uitwisseling**

Een toepassing die de laatste tijd sterk in de belangstelling staat is 'de elektronische snelweg': de mogelijkheid om de over de hele wereld uitgestrekte computernetwerken te gebruiken voor het verzenden van berichten. Uiteraard hebben alle wetenschappelijke disciplines baat bij *electronic mail*, die het gemakkelijker maakt om samen artikelen te schrijven, congressen te organiseren, elkaar op de hoogte te houden van nieuwe ontdekkingen, etc. Voor de historicus heeft het netwerk als extra voordeel dat zeldzame, moeilijk toegankelijke archieven wanneer zij eenmaal gedigitaliseerd zijn, op afstand geraadpleegd kunnen worden. Zo is er in 1993 een digitale tentoonstelling geweest van Vaticaanse documenten in de Library of Congress te Washington. Iemand die deze documenten wilde bestuderen, kon ze overal ter wereld 'ophalen' en op het eigen beeldscherm raadplegen. Ook de reeds genoemde datacollectie 'OASIS' is over Internet te raadplegen. De catalogus van historische bestanden van het Nederlands Historisch Data Archief is op deze wijze in principe vanuit de gehele wereld toegankelijk.

Een ander fenomeen dat zijn bestaan aan het netwerk te danken heeft is de *electronic discussion list*. Men kan zich via de computer opgeven voor een discussielijst omtrent een bepaald thema en ontvangt dan dagelijks de elektronische bijdragen van de andere deelnemers. Alleen dient de lijst door een strenge editor te worden geredigeerd, omdat men anders dagelijks grote hoeveelheden irrelevante informatie binnenkrijgt. Er is reeds een groot aanbod van discussielijsten over historische onderwerpen.<sup>31</sup>

De ontwikkelingen in de netwerkwereld gaan erg snel. Doorzichtige menu-systemen, die elke wildvreemde informatieplek op het netwerk op een herkenbare standaardmanier ontsluiten, volgen elkaar op. Zo zijn er 'Archie', 'Gopher' en 'World Wide Web'. De presentatie van de informatie op het beeldscherm wordt steeds fraaier en het zoeken en vinden van specifieke informatie wordt op deze wijze steeds goedkoper en sneller. Voor de historicus betekent dit, dat wanneer archief en documentatie-instellingen op grote schaal ontsloten worden, er een schat aan informatie beschikbaar komt die nu nog op kostbare reisafstand verwijderd ligt.<sup>32</sup>

## Historiografie

De afgelopen jaren is een toenemende hoeveelheid artikelen en boeken verschenen over het thema 'History and Computing'. We kunnen ons afvragen hoe deze literatuur het beste kan worden gezien: als een specifiek methodologisch sub-vakgebied (van en voor een geïsoleerde groep bekeerlingen binnen de historische wetenschap) of als een handreiking van nieuwe, efficiënte technieken aan de geïnteresseerde historicus?<sup>33</sup>

Uit het bovenstaande is inmiddels gebleken dat de computer enerzijds een hulpmiddel biedt bij datgene wat door historici altijd al werd gedaan (bronnen zoeken, literatuur zoeken, gegevens analyseren, een studie schrijven), maar anderzijds ook een methodologische vernieuwing biedt (nieuwe zoek-, sorteer- en analysemogelijkheden, tijdwinst bij bepaalde aspecten van het onderzoek, etc.). Bepaalde onderdelen van het historisch onderzoek die vroeger veel tijd kostten, kunnen nu sneller en vaker worden uitgevoerd, hetgeen andere accenten in het onderzoek met zich mee kan brengen of een geheel andere aanpak. Rapportage van diegenen, die met succes de computer bij hun onderzoek hebben gebruikt, kan anderen op ideeën brengen en tot voorbeeld dienen bij het ondernemen van nieuwe projecten.

Het gevaar is echter aanwezig dat de verslaglegging blijft steken op het individuele en thematische vlak. Na een beschrijving van de bron en het gebruikte computerprogramma volgen een paar tabellen of grafieken waarmee wordt getoond hoe waardevol de aangelegde database is, en het artikel eindigt met de conclusie, dat dit nog slechts een verkennende analyse van de database was en dat een zee aan mogelijkheden voor de onderzoeker van deze data openligt. Dergelijke bijdragen worden wel omschreven als '*me and my database*'. De eigenlijke reden van de verslaglegging van het onderzoek is dat de computer werd gebruikt om de gegevens in op te slaan - niet de inhoudelijke conclusies. Afgezien van deze thematische beperkingen hebben deze papers soms slechts journalistieke waarde: de informatica-toepassingen verouderen snel. Alleen de dikke historische studie op grond van het onderzoek (die hopelijk korte tijd later ook verschenen is) zal overleven. Maar ook als de database groot en veelzijdig van opzet is, zal zij voornamelijk interessant zijn voor onderzoekers die zich met hetzelfde onderzoeksthema bezighouden. Een dergelijk artikel kan, mits uitgebreid met een inhoudelijk-historiografische component, beter verschijnen in een historisch tijdschrift of een bundel over dat specifieke onderzoeksgebied.

Het nadeel van publikaties op het gebied van de historische informatiekunde is dus, dat een veelvoud van historische onderwerpen wordt gebundeld onder de gemeenschappelijke noemer dat de computer bij het onderzoek is gebruikt. Wanneer het gaat om verhelderende voorbeeldstudies, waarin nieuwe wegen worden geopend en methodologieën met behulp van de computer worden ontdekt waarvan iedere historicus kennis zou moeten nemen, dan is dit uiteraard nuttig.

## **Handboek**

Bij een nieuw vak hoort een nieuw handboek. Dat was de gedachte van de makers van het handboek 'Historische Informatiekunde', dat in 1990 verscheen.<sup>34</sup> In een twaalfstal hoofdstukken worden onderwerpen als historische database-systemen, analyse van teksten, verwerking van beeldmateriaal en tekstverwerken behandeld door verschillende auteurs. Als inleiding voldoet het boek goed en het strekt tot aanbeveling dat naast enkele aan de tand des tijds onderhevige hoofdstukken ('De anatomie van de computer', 'Tekstverwerken en het voorbereiden van publikaties') ook hoofdstukken zijn opgenomen waarin wordt gereflecteerd over de rol van de computer in historisch onderzoek. In hoofdstuk III, 'De computer in een historisch onderzoeksproject' probeert Boonstra antwoord te geven op de vraag waarom in sommige historische onderzoeksprojecten de computer wèl kan worden ingezet en in andere niet. Ter illustratie bespreekt hij toepassingen op het gebied van de historische demografie, materiële cultuur (boedelinventarissen), verhalende bronteksten (het onderzoek naar doodsbeleving door Breure) en tijdreeksen (roggeprijzen in Danzig 1562-1801). Eén van de conclusies luidt dat wanneer er systematiek aanwezig is in de gegevens, of wanneer de gegevens systematisch dienen te worden verwerkt, de computer een nuttige rol kan spelen bij de data verwerking of analyse.<sup>35</sup> In hoofdstuk V, 'Computer en geschiedenis', wordt de vraag gesteld: Is historisch onderzoek nog mogelijk zonder computer? Het moge inmiddels duidelijk zijn, dat de computer in vrijwel alle fasen van het historisch onderzoek bruikbaar is. De auteurs tonen aan dat in bijna alle takken van de geschiedenis de computer gebruikt wordt en dat het niet alleen meer gaat om kwantitatief of tekstanalytisch werk. Strijdvaardig roepen zij uit: 'De historicus die thans computerverwerking nog afwijst als overbodig sluit een breed gebied van onderzoeksmogelijkheden af en mag eigenlijk niet meer serieus worden genomen.'<sup>36</sup>

## **Tijdschriften**

Als belangrijkste vaktijdschriften voor de historische informatiekunde kunnen genoemd worden *History and Computing*, *Historical Social Research* en *Computers and the Humanities*.<sup>37</sup> *History and Computing* is het officiële orgaan van de Europese 'International Association for History and Computing' en verschijnt sinds 1989. Naast thematische artikelen die verslag doen van onderzoek waarbij de computer is toegepast, treft men boekbesprekingen, software besprekingen en aankondigingen van congressen. Een aantal artikelen levert algemene inzichten in het toepassen van de computer in historisch onderzoek. Sommige artikelen behandelen een methode of een bepaald softwarepakket en laten zien hoe deze gebruikt kan worden in historisch onderzoek. Andere gaan dieper in op de vernieuwende aspecten die de onderzoeker onder ogen zijn gekomen of richten zich volledig op de methodologische kaders van softwaretoepassing of software-ontwikkeling.

De Association for History and Computing heeft ook een Nederlandse pendant, de Vereniging voor Geschiedenis en Informatica. Deze geeft de *VGI Cahiers* uit, die gecentreerd zijn rond een thema, bijvoorbeeld onderwijs, of congresbijdragen bevatten.<sup>38</sup>

Binnen de geschiedbeoefening waren het de sociaal-economische historici, met name de historisch demografen, die van oudsher het meest van de computer gebruik maakten. Dezen zijn dan ook ruim vertegenwoordigd in *Historical Social Research/Historische Sozialforschung*, een tweetalig blad dat wordt uitgegeven door het Duitse QUANTUM. QUANTUM richt zich, zoals de naam al aangeeft, op de kwantitatieve geschiedbeoefening. Uitgangspunten zijn, dat men met behulp van computer-databases tot een betere onderbouwing van uitspraken betreffende de geschiedenis komt, verkeerde opvattingen kan corrigeren, massale bronnen kan ontsluiten en een nieuw instrument heeft om de kloof tussen theorie en empirie te overbruggen. Deze geschiedbeoefening is complementair van aard en beoogt niet om de traditionele geschiedbeoefening te vervangen.<sup>39</sup>

*Computers and the Humanities* is het oudst en bestrijkt alle humaniora aan beide zijden van de Atlantische Oceaan, hoewel in de praktijk het Noordamerikaanse continent wat sterker vertegenwoordigd is. De taal- en tekstanalyse nemen de belangrijke plaats in in de alfa-informatica. Het aantal bijdragen dat specifiek op historische toepassingen gericht is, is niet zo groot. Dit maakt het mogelijk om binnen het bestek van dit artikel een overzicht te geven van artikelen in *Computers and the Humanities* die voor de historicus interessant zijn.

Naast evaluaties van software-pakketten kunnen artikelen over taalanalyse met behulp van concordanties voor bepaalde historici bruikbaar zijn. Een voorbeeld is het artikel 'The language of Enlightened Politics: The *Société de 1789* in the French Revolution' van M. Olsen, waarin op systematische wijze de taal in het Franse politieke debat wordt geanalyseerd, zodat heersende opvattingen over macht etc. ten tijde van de Franse Revolutie aan het licht komen.<sup>40</sup> Verder verschenen na 1988 enkele artikelen over algemeen en specifiek historisch database beheer,<sup>41</sup> een bijdrage over kerkgeschiedenis en de computer<sup>42</sup> en een overzichtsartikel van H. Best over kwantitatieve methoden in de sociale geschiedenis. Tenslotte vindt men ook in *Computers and the Humanities* aandacht voor de gespletenheid van de alfa-informatica (*Humanities computing*): 'Is there any set of basic goals, methods and/or assumptions in our discipline which could be set forth clearly and proved either right or wrong in some sort of debate with the non-believers?'<sup>43</sup> Evenals de historische informatiekunde, is in ruimere zin de 'Letteren Informatiekunde' of 'Alfa-Informatica' ontstaan in een tijd dat al het computergebruik zich concentreerde rond het mainframe, waar het aantal programmeertalen beperkt en de gebruikersvriendelijkheid gering was. Na de opkomst van de PC is het vak 'snel en willekeurig alle kanten opgegroeid'.<sup>44</sup>

*Computers and the Humanities* toont ons twee zaken: ten eerste is de grote verscheidenheid van computertoepassingen, zowel bij geschiedenis als bij Letteren in het algemeen, een kenmerk dat de informatiekunde tot een zeer

pluriform geheel maakt. Deze heterogeniteit is de afgelopen 25 jaar sterk gegroeid. Ten tweede zien we dat binnen deze verscheidenheid bepaalde groepen sterk overheersen: zo zijn binnen Letteren de taalkundigen (linguïsten) de grootste computergebruikers.<sup>45</sup> We zagen reeds dat binnen de geschiedenis de sociaal-economische historici het meest dikwijls de computer bij hun onderzoek gebruiken.

### **Congresbundels**

In een aantal bundels zijn congresbijdragen samengebracht op het gebied van geschiedenis en informatica. Nuttige literatuur vormt de reeks 'History and Computing', die de congresbijdragen bundelt van de jaarlijkse congressen van de reeds genoemde International Association of History and Computing. Van de conferenties in Londen, Glasgow, Bordeaux en Montpellier zijn inmiddels de proceedings verschenen.<sup>46</sup> Voor de geïnteresseerde historicus geven deze delen een goede indruk van de veelzijdigheid aan onderzoeken en database-projecten die door historici worden geëntameerd. Ook kan hier een bundel genoemd worden van een eerder congres, dat onder de brede titel 'Use of the computer in the Humanities and Social Sciences' in Keulen werd gehouden. De reeds genoemde thematische versnippering treedt hier sterker op.<sup>47</sup>

Over computers en Letteren in bredere zin wordt nagedacht in de bundels van Miall, Lancashire, en Genet en Zampolli<sup>48</sup>. Zo stelt P. Denley dat 'some historians are getting wary of what the computer offers because they see computing techniques as emphasizing the repetitious at the expense of the unique.'<sup>49</sup> Genet benadrukt dat er zeker sprake kan zijn van een vak als 'computerkunde in historisch onderzoek', welk specialisme verder reikt dan slechts het behandelen van computertoepassingen bij één bepaald soort onderzoek. Ook benadrukt hij het nut van vakliteratuur op het gebied van de historische informatiekunde, ter aanvulling of vervanging van de handleidingen gebruikte computerprogramma's.<sup>50</sup>

Tenslotte kan hier genoemd worden de uitgebreide reeks *Halbgraue Reihe zur Historische Fachinformatik*, met serie A: *Historische Quellenkunde* en serie B: *Software Beschreibungen*. In deze reeks, die wordt uitgegeven door het Max Planck Institut für Geschichte in Göttingen, komt een veelheid van onderwerpen aan bod, van de analyse van beroepen of boedelinventarissen tot de stand van zaken van de historische informatiekunde in Oost-Europa.<sup>51</sup>

In het voorgaande is voorbijgegaan aan de grote stroom informatie die beschikbaar is van documentatie-instellingen, archieven, databanken en andere informatie-voorzieners. Sommige nieuwsbrieven worden inmiddels zelfs over het netwerk verspreid.<sup>52</sup> Ook de reeds genoemde *electronic discussion lists* of *listservers* zijn een manier om op de hoogte te blijven van recente ontwikkelingen.

Samenvattend kan gesteld worden dat de historiografie vakgenoten kan wijzen op nieuwe toepassingsmogelijkheden en via voorbeeldstudies op elk deelgebied onmisbare informatie levert voor degenen die met moeizaam te

ontcijferen handleidingen worstelen en proberen zich de wereld van knoppenbalk en pull-down-menu eigen te maken. Voor de nieuwe generatie historici, die opgroeit met het besef van de mogelijkheden van de computer, zijn het handboek *Historische Informatiekunde* en het tijdschrift *History and Computing* een basis van waaruit specifiekere expertise kan worden opgebouwd. Men kan verwachten dat de methodologie van computergebruik in toenemende mate integreert in het vakgebied, zodat uiteindelijk de historische informatiekunde zal opgaan in de historische methoden. De literatuur zal dan, behalve onder de noemer 'computer', ook te vinden zijn onder het betreffende historiografische thema in de 'gewone' historische tijdschriften.<sup>53</sup> Het wachten is dus op meer aandacht in deze tijdschriften, hetzij in voetnoten, hetzij in bijlagen, voor de nieuwe mogelijkheden die het gevolg zijn van het gebruik van de computer. De onlangs geuite klacht dat het vak geschiedenis en informatica teveel gericht zou zijn op de eigen navel, zal door deze ontwikkeling vanzelf worden weggenomen.<sup>54</sup> De algemene toepassingen zullen thematisch worden ingebed in het vakgebied, en er zal een kleine hulpwetenschap blijven bestaan, waarin de ontwikkeling van specialistische algoritmen en informaticatechnieken voor de doorgewinterde computeraars centraal staat.

## Conclusie

De historicus kijkt vanuit het heden naar het verleden. In deze tijd van snelle technologische ontwikkelingen zullen een aantal methodes en hulpmiddelen voor de studie van het verleden voortdurend vernieuwd worden. Wie zoekt nog een na 1962 verschenen boek in de Leidse boekjes van de Universiteitsbibliotheek, terwijl er een terminal staat die in seconden de gezochte titel toont?

Met name wanneer het gaat om bronnenuitgaven of bronnen-analyses, kan de computer de historicus van groot nut zijn. Dit wil niet zeggen dat de 'traditionele' aspecten van het onderzoek, zoals het stellen van intelligente vragen, het aanbrengen van verbanden, het trekken van conclusies en het schrijven van een heldere studie, minder aandacht hoeven te krijgen. Het is dan ook zeker niet zo dat voorstanders van het gebruik van de computer aan dergelijke onderzoeksfases minder waarde zouden hechten.

Behalve de titelbeschrijving of het inventarisnummer, zal ook het materiaal zelf steeds vaker op een beeldscherm worden getoond. De contemporain historicus die in de nabije toekomst de Golfoorlog of het Iran-Contra schandaal bestudeert zal grote hoeveelheden archiefmateriaal in elektronische vorm vinden. Voor deze archivalia dienen zich nieuwe zoekstrategieën aan. Maar ook de mediaevist die levensomstandigheden wil interpreteren aan de hand van reële prijzen, zal in een databank een volledig overzicht van prijs- en wisselkoersgegevens kunnen vinden, waaraan de gewenste (of althans de aanwezige) brongegevens gekoppeld zijn. Hier is de beschikbaarheid van het materiaal sterk verbeterd door de komst van de elektronische databank. Over het algemeen zal de moderne historicus de

computer erkennen als een waardevol en krachtig hulpmiddel, zij het dat de voordelen pas na geduchte inspanning en studie optimaal aangewend kunnen worden.

De informatie-infrastructuur in de wereld verandert en de historicus komt met die veranderingen in aanraking. Ligt de nadruk nu nog op vernieuwde ontsluitingsmogelijkheden van bronnen en literatuur, in de toekomst zullen ook de nieuwe analyse-methoden meer en meer gemeengoed worden in het historisch onderzoek. Iedereen die een PC gebruikt voor tekstverwerking, heeft een groot deel van die mogelijkheden reeds binnen handbereik.

De fase waarin de historische informatiekunde zich momenteel bevindt, is niet meer de emancipatie-strijd van vroege voorstanders, maar eerder een fase van consolidatie van methodologie. Voor het vakgebied is het van groot belang dat een generatie ontstaat, die de computer weet in te zetten op de gebieden waar dat voordelig is. Een toegankelijke historiografie van voorbeeldstudies op elk deelgebied is hiervoor van belang. De methodologische erkenning van computergebruik zal dan vanzelf volgen. Toch is het belangrijk dat de historici die zich met de historische informatiekunde bezighouden, de computeraars die vol vuur werken aan razendknappe oplossingen en prachtige toepassingen, zich niet in wederzijdse bewondering afzonderen in een eigen subdiscipline. Zoals het voornaamste overblijfsel van de Cliometrie wordt gevormd door de methoden en technieken die zijn opgenomen in de sociaal-economische geschiedenis, zo zal ook de toekomst van de historische informatiekunde bestaan uit het aanreiken en onderbouwen van technieken die iedere historicus van nut kunnen zijn. Het preken voor eigen parochie zal daarbij slechts een zijdelingse rol vervullen.

#### **Aanbevolen literatuur:**

Boonstra, O., L. Breure en P. Doorn, ed., *Historische informatiekunde. Inleiding tot het gebruik van de computer bij historische studies* (Hilversum 1990).

Denley, P., en D. Hopkin, ed., *History and computing* (Manchester 1987).

Denley, P., S. Fogelvik en C. Harvey, ed., *History and computing II* (Manchester 1989).

*History and Computing* 1 (1988) - 5 (1993)

Mawdsley, E., en T. Munck, *Computing for historians. An introductory guide* (Manchester en New York 1993)

Reiff, J.L., *Structuring the past. The use of computers in history* (American Historical Association 1991).

*VGI Cahiers* 5 (1992) -

Noten:

1. Met dank aan Ellen Leenarts voor het kritisch doorlezen van een eerdere versie van dit artikel.
2. Zoals op elk ander vakgebied kan ook bij het geschiedenisonderwijs de computer als hulpmiddel worden ingezet. Het aanschouwelijk maken van inzichten en informatie via een multi-media systeem, het aanleren van onderzoek met behulp van oorspronkelijk data-materiaal en het gebruik van modellen en simulaties zijn voorbeelden waarbij nuttig gebruik wordt gemaakt van de computer. Hoewel het geschiedenisonderwijs met behulp van de computer onderdeel uitmaakt van de historische informatiekunde, zullen wij hier in dit artikel niet nader op ingaan. Ter introductie kan gebruik gemaakt worden van D. Aberson e.a., ed., *VGI-Cahier 6* (Hilversum 1993) dat voor een groot deel in het teken staat van computers in het geschiedenisonderwijs. Ook hoofdstuk XI van O. Boonstra, L. Breure en P. Doorn, ed., *Historische informatiekunde. Inleiding tot het gebruik van de computer bij historische studies* (Hilversum 1990) (359-378) geeft een goede inleiding. Inspirerend is R. Trainor, 'Using computers in historical teaching and research: the British experience', *Tijdschrift voor Geschiedenis (TvG)* 103 (1990) 373-380.
3. Uiteraard komen ook bij de hoofdonderdelen archiefontsluiting en informatie-uitwisseling veelvuldig heuristische toepassingen voor. Zo is een groot voordeel van het Internet dat men met gemak literatuur kan zoeken in de catalogi van Amerikaanse bibliotheken. Met nadruk dient te worden vermeld dat andere indelingen dan de gebruikte zeer goed mogelijk zijn. De gebruikte indeling is puur praktisch van aard.
4. Boonstra e.a., *Historische informatiekunde*, 11.
5. E. Mawdsley e.a., ed., *Historians, computers and data. History and computing III: Applications in research and teaching* (Manchester 1990) ix.
6. O.W.A. Boonstra, 'Wat is historische informatiekunde. Een repliek' in: Aberson, ed., *VGI-Cahier 6*, 72-74. Zie ook: G. Welling en E. Kleijn, 'Wat is historische informatiekunde' in: *Ibidem*, 61-71.
7. M. Thaller, 'The need for a theory of historical computing' in: Denley, Fogelvik en Harvey, ed., *History and computing II* (Manchester 1989) 2-12. Het aan het Max Planck Institut für Geschichte te Göttingen ontwikkelde databasemanagement systeem heet 'Kleio'. Onlangs is een Engelse vertaling verschenen van dit programma (versie 5.1.1), dat oorspronkelijk met commando's in het Latijn werkte (!). Verg. M. Thaller, 'Methods and techniques of historical computations' in: P. Denley en D. Hopkin, ed., *History and computing* (Manchester 1987) 147-156 en M. Thaller, 'Databases and expert systems as complementary tools for historical research', *Tijdschrift voor Geschiedenis (TvG)* 103 (1990) 233-247.
8. Boonstra, 'Wat is historische informatiekunde', 73.
9. De Nederlandse CD-ROM 'Dutch Printers Devices' bijvoorbeeld bevat afbeeldingen en beschrijvingen van de collectie historische drukkersmerken van de Koninklijke Bibliotheek. Deze afbeeldingen zijn geclassificeerd volgens het Iconclass systeem, dat oorspronkelijk in Leiden door de kunsthistoricus prof.dr. van der Waal ontwikkeld is. De CD-ROM toont dat de computer het mogelijk maakt om op een elegante manier via dit systeem de gegevens te bevragen. Dit 'voorbeeld-nut' is groter dan de daadwerkelijke waarde van de data, omdat de verzameling drukkersmerken tamelijk klein is (900 drukkersmerken).
10. Veel oud bezit van universiteitsbibliotheken is alleen via de kaartcatalogus of de 'Leidse boekjes' toegankelijk. Bibliotheken zijn nu bezig in speciale projecten deze gegevens alsnog in te voeren, maar dit vergt veel tijd en moeite. Bij het standaardiseren en koppelen van de bestanden speelt PICA uiteraard een grote rol. Momenteel wordt ook gewerkt aan het koppelen van CD-ROM's die collecties tijdschriftartikelen bevatten aan het lokale tijdschriftenbestand van de bibliotheek.

11. Vergelijk O.W.A. Boonstra en H.J.C. Termeer, 'Gevraagd: een nieuwe algemene bibliografische databank voor historisch onderzoek' in: D. Abernson, L. Breure en G. Colletteur, ed., *VGI-Cahier 4* (Hilversum 1991) 7-16, aldaar 8.
12. Deze elektronische bibliografie werd gemaakt door de 'Association for History and Computing' (AHC) in samenwerking met het Queen Mary and Westfield College van de University of London en bevat 4000 verwijzingen. HCBIBL draait op elke computer. Als men een schijfje opstuurt ontvangt men deze bibliografie zelfs gratis.
13. Dit katern heet *Boek en byte* en behandelt nieuwe CD-ROMs, netwerktoegangen naar documentatiecentra, etc. Het rekencentrum heet overigens tegenwoordig instituut voor telecommunicatie en computerdienstverlening CRI.
14. P. Gonzales, 'Computerisation project for the "Archivo general de Indias"' in: P.K. Doorn, C. Kluts en E. Leenarts, ed., *Data, computers and the past. VGI Cahier 5* (Hilversum 1992) 52-67. Het project won de Europese 'Erasmus Prijs' in 1992.
15. Zie bijvoorbeeld de berichtgeving in *NRC Handelsblad* van 17 maart 1994 en *Computable* van 18 maart 1994. Andere voorbeelden van gedigitaliseerde archiefcollecties zijn Spaanse incunabelen op de CD-ROM 'Admyte' en de reeds genoemde Nederlandse CD-ROM 'Dutch Printers Devices'.
16. Zie bijvoorbeeld: L.J. Touwen, 'Experiences in data-archiving: developing the History of European Expansion Data Bank' in: P.K. Doorn, M.P.M. van Horik en L.J. Touwen, *Nederlands Historisch Data Archief I. Eindverslag van een pilot project* (Amsterdam 1990) 15-23; K.J. Kiecolt en L.E. Nathan, *Secondary analysis of survey data* (Beverly Hills 1985).
17. Bijvoorbeeld de Thesaurus Linguae Graecae (TLG), CETEDOC, the American and French Research on the Treasury of the French Language (ARTFL), de Dartmouth Dante Database, de Fontes Anglo-Saxonici, etc. Zie ook M. Neumann e.a., *The Georgetown University Catalogue of Projects in Electronic text* (Washington 1991). Tekst-archieven, die zich specialiseren in de ontsluiting van elektronische teksten zijn the Oxford Electronic Text Archive in Oxford en het Center for Electronic Texts in the Humanities (CETH) aan Rutgers University (New Jersey).
18. J.P. Genet en A. Zampolli, *Computers and the Humanities* (Dartmouth 1992) 32.
19. R.M. Bell en E. van Cauwenbergh, 'The Medieval and Early Modern Data Bank (MEMDB)', *TvG* 103 (1990) 260-278; L.J. Touwen, 'Electronic bookshelf or electronic library? An examination of the Medieval and Early Modern Data Bank', *Historical Social Research* 18,2 (1993) 213-229.
20. Informatie over het NHDA in: P.K. Doorn en N.S. van Hall, *Nederlands Historisch Data Archief. Eindverslag van een verkennend onderzoek* (Amsterdam 1989); Doorn e.a., *Nederlands Historisch Data Archief I*.  
Het NHDA publiceert (sinds 1994 samen met de Vereniging voor Geschiedenis en Informatica en de Historische Steekproef Nederlandse Bevolking) tweemaal per jaar een nieuwsbrief, waarop men zich kan abonneren.  
Voor inlichtingen: NHDA, Doelensteeg 16, 2311 VL Leiden, 071-272742 of 272719.
21. P.F.A. de Guchteneire en S.C.M. de Schaeppdrijver, *Historische databestanden Nederland 1987* (Amsterdam 1987); N.S. van Hall en W. van Tellingen, *Nederlands Historisch Data Archief III: Historische databestanden 1990* (Amsterdam 1990).
22. K. Schürer en S.J. Anderson, ed., *A guide to historical data files held in machine readable form* (Londen 1992).
23. Een grondige analyse van de toepassingsmogelijkheden voor historisch materiaal van een aantal OCR-software pakketten vindt men in R. van Horik, *Van Beeldpunt tot betekenis* (Amsterdam 1992). Dit boek bevat de verslaglegging van een onderzoek dat in 1991 bij het NHDA werd uitgevoerd. Samenvattende conclusies staan in R. van Horik, 'Optical Character Recognition and historical documents: some programs reviewed', *History and Computing* 4,3 (1992) 211-220. Het gaat hier om de OCR programmas Gigaread 1.52, Kurzweil K5200 6.01, LiOCR 2.0, Omnipage Professi-

- onal 2.0, ProLector 1.12, Recognita Plus 1.2, Topscan Professional 1.1 en TypeReader 1.0. Uiteraard zijn er inmiddels betere versies van deze pakketten en nieuwe programma's op de markt. Veel van de besproken problemen met het scannen van oude of beschadigde documenten blijven echter representatief. Wanneer het gaat om behalve het scannen ook het ordenen en ontsluiten van het ingelezen (heuristische) materiaal, geeft een helder overzicht van meerdere projecten: Doorn e.a., 'Digitaliseren en ontsluiten van historische bescheiden', *Nederlands Archievenblad* 3 (1993) 168 e.v..
24. Zo is er bijvoorbeeld in Groningen het 'Groningen Historical Electronic Text Archive' (GHETA), dat via het netwerk enkele bestanden beschikbaar stelt.
  25. *Documenten uit de tijd; behoud en beheer van digitale informatie*. Project MLG-Machine Leesbare Bestanden (Den Haag 1993).
  26. P. Marsden, 'The impact of electronic information technology on the philosophy and policies of the Government Archives Division' (Government Archives Division, National Archives of Canada) en M. Geber, 'Acquisition and use of electronic records in the National Archives of Sweden' (National Archives of Sweden). Congresbijdragen IASSIST/IFDO Conference 1993 (Edinburgh 1993).
  27. W.Chr. Pieterse, 'Praktijkervaringen van een gemeentearchivaris' in: 'Discussiedossier archiefvernieuwing,' *Bijdragen en Mededelingen betreffende de Geschiedenis der Nederlanden* 108 (1993) 722-785, aldaar 739.
  28. H.M. Van den Heuvel, *De criminele vonnisboeken van Leiden 1533-1811* (Leiden 1977-1978). Zie ook *NHDA-Nieuws* 4 (1992/1993) 5.
  29. Dit beeld is grotendeels gebaseerd op Boonstra e.a., *Historische informatiekunde* 380-384. Ook is gebruik gemaakt van Van Hall en Van Tellingen, *Nederlands Historisch Data Archief III*, 11. Van de 131 historische datasets die in dit overzicht zijn opgenomen, hadden 82 een sociaal-economische component. Schürer en Anderson wijzen op een grote spreiding binnen het gedigitaliseerde bronmateriaal, maar vinden een overgrote meerderheid aan databestanden met bevolkingstellingen en gemeenteregisters. De historicus onderscheidt zich van de sociale wetenschapper door zijn voorkeur voor bestanden met gemengde tekstuele en numerieke gegevens. 24,8 % van de data bestanden uit Schürer en Anderson bevat bronmateriaal uit de 17e of 18e eeuw, 55,8 % de 19e en slechts 3,8 % de 20e eeuw. Het hoge percentage voor de 19e eeuw wordt door de auteurs verklaard door het grote aantal ingevoerde volkstellingen. (Schürer en Anderson, ed., *A guide to historical data files*, 6-7.)
  30. J. Brefeld, *A guidebook for the Jerusalem Pilgrimage in the late Middle Ages. A case for computer-aided textual criticism* (Proefschrift Rijksuniversiteit Groningen 1994).
  31. Een uitgebreid overzicht van discussielijsten en LISTSERVERS (on-line databanken waar men zelf documenten kan opvragen) met historische onderwerpen wordt gegeven in: G.M. Welling, 'Historici en computernetwerken' in: Abersson e.a., *VGI-Cahier* 6, 103-126.
  32. E. Jannink, *De elektronische jungle van het Internet: Een onderzoek naar de mogelijkheden voor het Nederlands Historisch Data Archief*. *NHDA-Rapporten I* (Leiden 1993).
  33. Vergelijk M. Thaller, 'The need for a theory of historical computing' in: Denley e.a. *History and computing II* 2-12, aldaar 2.
  34. Boonstra e.a., *Historische informatiekunde*. Buitenlandse handboeken die genoemd kunnen worden zijn E. Mawdsley en T. Munck, *Computing for historians. An introductory guide* (Manchester en New York 1993); J.L. Reiff, *Structuring the past. The use of computers in history* (American Historical Association 1991).
  35. Boonstra e.a., *Historische informatiekunde*, 50-56.
  36. P. Doorn, O. Boonstra en L. Breure, 'Computer en geschiedenis' in: Boonstra e.a., *Historische informatiekunde*, 385.

37. Vergelijk L. Breure, 'Historische databasesystemen', *TvG* 103 (1990) 217-232, aldaar 230 en noot 25.
38. Doorn e.a., *Data, computers and the past. VGI Cahier 5.*
39. H. Best, 'Technology or methodology? Quantitative historical social methods in Germany', *Computers and the Humanities (CHum)* 25 (1991) 163-171, aldaar 165.
40. M. Olsen, 'The language of enlightened politics: The *Société de 1789* in the French Revolution', *CHum* 23 (1989) 357-364. Olsen werkte bij het ARTFL-project aan de Universiteit van Chicago, waar men de Franse literatuur en ander geschreven Frans materiaal in de computer invoert om de taal te kunnen analyseren. Het betreffende artikel baseert zich op de *Cahiers Général de Doléances* uit 1789.
41. M. Thaller, 'The Historical Workstation Project', *CHum* 25 (1991) 149-162; L.J. Touwen, 'The Medieval and Early Modern Data Bank', *CHum* 26 (1992) 237-247.
42. A. Melloni, 'Church History and the Computer', *CHum* 24 (1990) 393-395. Melloni beschrijft hoe de bekende pioniersarbeid van Roberto Busa, die het complete oeuvre van St. Thomas met behulp van de computer van een concordantie voorzag, zijn vervolg heeft gekregen in verscheidene andere Italiaanse projecten, zoals het computergestuurd indexereren van alle Italiaanse en Latijnse geschriften van Paus Johannes XXIII. (Verg. D.M. Burton, 'Automated concordances and word-indexes: the fifties,' *CHum* 15 (1981) 1-14.)
43. D. Brink, 'Input ... Output', *CHum* 24 (1990) 105-106.
44. J. Raben, bespreking van I. Lancashire en W. McCarthy, ed., *The Humanities Computing Yearbook*, *CHum* 24 (1990) 111-113. Vergelijk ook: J. Raben, 'Humanities Computing 25 years later', *CHum* 25 (1991) 341-350, dat een overzicht geeft van 25 jaar *Computers and the Humanities*.
45. G.P. Landow en P. Delany, ed. *The digital word; text-based computing in the humanities* (Washington 1993). Zie ook het tijdschrift *Literary & Linguistic Computing*.
46. De Association for History and Computing werd in 1986 opgericht en beoogt middels jaarlijkse congressen en een tijdschrift de inzichten op het gebied van de toepassing van de computer in de geschiedenis in Europa te verspreiden en bevorderen. In elk land zijn deelorganisaties gevestigd, zoals de Vereniging voor Geschiedenis en Informatica (VGI) in Nederland. Inmiddels zijn een aantal congresbundels verschenen, die een goede indruk geven van dit vakgebied: P. Denley en D. Hopkin, ed., *History and computing* (Manchester 1987); P. Denley, S. Fogelvik en C. Harvey, ed., *History and computing II* (Manchester 1989); E. Mawdsley e.a., ed., *Historians, computers and data. History and computing III: Applications in research and teaching* (Manchester 1990); B. Lavallé, M. Usselman en J. Vivés, ed., *L'ordinateur et le métier d'historien. Volume des actes du IVe congrès 'History and computing'*, Talence 1989 (Bordeaux 1990); J. Smets, ed., *Histoire et informatique. Actes du Ve congrès 'History and computing'*, Montpellier 1990 (Montpellier 1992). Van de congressen in Odense (1991), Bologna (1992), Graz (1993) en Nijmegen (1994) worden nog procecdings verwacht.
47. H. Best, E. Mochmann en M. Thaller, ed., *Computers in the Humanities and the Social Sciences. Proceedings of the Cologne computer conference. 1988: Use of the computer in the Humanities and Social Sciences* (München 1991).
48. D.S. Miall, ed. *Humanities and the computer. New directions* (Oxford 1990); I. Lancashire, ed., *The Humanities Computing Yearbook 1989-90* (Oxford 1991); J.P. Genet en A. Zampolli, *Computers and the Humanities* (Dartmouth 1992).
49. P. Denley, 'The computer revolution and "redefining the Humanities"' in: Miall, ed., *Humanities and the computer*, 13-25, aldaar 18.
50. Genet en Zampolli, *Computers and the Humanities*, 37.
51. In de serie *Halbgraue Reihe zur Historische Fachinformatik. Serie A: Historische Quellenkunde* verschenen onder meer de volgende delen:

- L.I. Borodkin en W. Levermann, ed., *History and computing in Eastern Europe* (Band 21) (St. Katharinen 1993);
- D. Greenstein, ed., *Modelling historical data: Towards a standard for encoding and exchanging machine-readable texts* (Band 11) (Göttingen 1991);
- Nederlands Historisch Data Archief/Nijmegen Institute for Cognition and Informatica, *Optical Character Recognition in the historical discipline, Proceedings of an International Workgroup* (Band 18) (St. Katharinen 1993);
- J. Nemitz, *Die historische Analyse städtischer Wohn- und Gewerbelagen: die Auswertung Socialtopographischer Quellen* (Band 3) (Göttingen 1989);
- K. Schürer en H. Diederiks, ed., *The use of occupations in historical analysis* (Band 19) (St. Katharinen 1993);
- D. Spaeth e.a., *Towards an international curriculum for history and computing* (Band 12) (St. Katharinen 1992);
- W. Werner en G. Grotum, *Sämtlich Hab und Gut ... Die Analyse von Besitzstandslisten* (Band 2) (Göttingen 1989).
52. Zoals de *Research-Newsletter* van de Humanities Computing Facility van de University of California at Santa Barbara, een gratis per post of netwerk verkrijgbare nieuwsbrief over computers en Letteren.
53. Een voorbeeld van een overzichtsartikel dat zich geheel richt op computertoepassingen in één thematisch gebied is: J.Th. Lindblad, 'Computer applications in Expansion History. A survey', *Itinerario* 12 (1988) afl. 2, supplement (Leiden 1988).
54. Vergelijk P.K. Doorn, 'AHC tussen Graz en Nijmegen. Een persoonlijke visie', *NHDA Nieuws* 6 (1993-1994) 4.

