Danskernes Historie Online
Danske Slægtsforskeres Bibliotek

Dette værk er downloadet fra Danskernes Historie Online

Danskernes Historie Online er Danmarks største digitaliseringsprojekt af litteratur inden for emner som personalhistorie, lokalhistorie og slægtsforskning. Biblioteket hører under den almennyttige forening Danske Slægtsforskere. Vi bevarer vores fælles kulturarv, digitaliserer den og stiller den til rådighed for alle interesserede.

Støt vores arbejde – Bliv sponsor

Som sponsor i biblioteket opnår du en række fordele. Læs mere om fordele og sponsorat her: https://slaegtsbibliotek.dk/sponsorat

Ophavsret

Biblioteket indeholder værker både med og uden ophavsret. For værker, som er omfattet af ophavsret, må PDF-filen kun benyttes til personligt brug.

Links

Slægtsforskernes Bibliotek: https://slaegtsbibliotek.dk
Danske Slægtsforskere: https://slaegt.dk

EDB I DE HISTORISKE VIDENSKABER

RAPPORT FRA EN KONFERENCE 16. - 17. JANUAR 1973

KØBENHAVN
FÆLLESUDVALGET FOR HISTORISK FORSKNING
1975

EDB I DE HISTORISKE VIDENSKABER

RAPPORT FRA EN KONFERENCE 16. - 17. JANUAR 1973

KØBENHAVN
FÆLLESUDVALGET FOR HISTORISK FORSKNING
1975

COPYRIGHT 1975

by Fællesudvalget for historisk Forskning

Redaktion:
Erik Stig Jørgensen

ISBN 87 7423 043 3

FORORD

Fællesudvalget udsender hermed rapport fra konferencen vedr. anvendelsen af EDB i de historiske videnskaber, 16. - 17. januar 1973. Vi takker IBM's forskningsfond for støtte til konferencens afholdelse, og Ministeriet for kulturelle anliggender for tilladelse til at benytte midler af en tidligere givet bevilling til fremstilling og udsendelse af rapporten.

August 1975

Niels Knud Andersen

Fællesudvalgets formand

INDHOLD

Folketingsbibliotekar Kristian Hvidt:
EDB og historie. Nogle synspunkter 7

Forskningsassistent Ole S.D.Hansen:
Databanker for samfundsvidenskaberne i USA 11

Arkivar Grethe Ilsøe:
Er EDB-registrering af tingbøger et alternativ til udgivelse? 26

Universitetsadjunkt Jørgen Elklit:
Det politisk-økologiske EDB-arkiv ved Århus Universitet 33

Forskningsassistent Svend Bang:
Problemer omkring tilgængeligheden af data i nogle større offentlige EDB-registre 43

DELTAGERLISTE 46

Folketingsbibliotekar Kristian Hvidt:
EDB og historie. Nogle synspunkter.

Indtil for få år siden var begrebet databehandling omgivet af en air af magi og mystik, det var noget, teknikere på avanceret niveau talte om, men som historikere ikke drømte om. Men alle de fremtidsvisioner om dataalderen, der i begyndelsen af tresserne mest havde karakter af science fiction, er nu ved at blive til realitet, og visionerne synes ikke at have været overdrevet. EDB alderen er kommet nu og selv historikere har nu mulighed for at tage databehandlingen med i betragtning.

Den tekniske udvikling har stillet historieforskningen i lidt af et dilemma, i en slags spændingsforhold mellem på den ene side påvirkning fra nye metoder i de omgivende eksakte fag og på den anden side den rent fortællende, "episke" forpligtelse, der ligger i historiens humanistiske baggrund. Den stærke tiltrækning, som f.eks. statitik og sociologi udøver på historien truer til en vis grad med at gøre dette fag til "noget andet", historieforskningen er ved at blive et eksakt fag, der udtrykker sig i matematisk-statistiske termer, ulæseligt for det publikum, der ellers læser, hvad historikere skriver. EDB kan bidrage til at få denne udvikling i hi-

storieforskningen til at accellerere og man kan spørge sig, om det er en rigtig vej og i hvert fald advare mod, at historien går hen og taber sin identitet.

Det er umuligt at overskue eller opregne de nye muligheder, som EDB giver historieforskningen. Det er en helt ny verden, der åbner sig, hvor hidtil upåagtede kildegrupper kan give nye indblik i socialhistorien, de kollektive fænomener i fortidens samfund. Grænserne for mulighederne ligger ikke i de tekniske begrænsninger, men i arkivernes beholdninger og tidligere tiders kassationspolitik. Men relativt store arkivmasser er da bevaret i Danmark, man har været meget tilbageholdende med kassationer, bl.a. af hensyn til den personalhistoriske forskning.

Statistisk behandling i historieforskning er jo langt fra noget nyt, det nye ligger i muligheden af at behandle de meget store datamængder under et. Men tendensen til at kvantificere i historisk fremstilling, komme bort fra de mere vage vurderinger omgivet af ord som "formentlig", "måske" osv., er gammel. Man kan ikke se bort fra, at der i denne tilbøjelighed til kvantificering ligger lidt af et status-symbol for historieforskning. Man nærmer derved faget til de eksakte fag. Tabeller synes at have samme funktion i moderne historiske fremstillinger, som latinske citater havde det i 1600- og 1700-tallets historiebøger.

I løbet af dette århundrede har man fra den "klassiske" historie udskilt den økonomiske historie efter det kriterium, at den sidstnævnte anvendte kvantitative metoder i modsætning til den førnævnte. Denne deling må vel nu betragtes i nogen grad som fiktiv, idet kvantitative metoder nu anvendes overalt, bl.a. med statistisk materiale fra talrige ikke-økonomiske grupper: demografi, valgstatistik o.l..

En bedre undervisning i statistik for historikere vil blive en absolut nødvendighed i de kommende år. Men spørgsmålet er, i hvor høj grad historikere selv skal gå ind i programmeringsvidenskaben og selv fremstille de programmer, der skal anvendes i deres forskning. Jeg mener, at det er bedst at samarbejde med en EDB-specialist. Mange historikere kan sikkert lære EDB-teknikken selv og efterhånden på egen hånd gennemføre et projekt gennem alle faser. Men det vil let føre til, at historikerens opmærksomhed vil vende sig unødvendigt meget mod de tekniske vanskeligheder og i nogen grad miste overblikket over det væsentlige: den historiske sammenhæng, ideerne, hypotesen.

Datamaskiner har som bekendt en eventyrlig kapacitet til at absorbere og "fordøje" store datamængder. Derimod - hvad der måske er mindre erkendt - er historikernes "fordøjelsesevne" af output stærkt begrænset. Når først materialet er på båndene, er muligheden for tabelfremstilling så stor, at man meget let drukner i tabeller. Det er et stort og møjsommeligt arbejde at bearbejde det udkomne materiale. Kunsten ved dataprojekter er ikke så meget at konstatere, hvad tabellerne udsiger, men at fortolke dem, finde ud af hvorfor og kombinere de almindelige historiske kendsgerninger med tabellernes udsagn. Den tunge ende af dataprojekter er overgangen fra behandlingen af tabellerne til den historiske fremstilling. Mange projekter med EDB-behandling er gået i stykker, fordi ambitionsniveauet fra begyndelsen er sat for højt, og materialet let har været for stort til, at det har været muligt at udmønte det i fremstilling.

Det er nærliggende at spørge om EDB-behandling med dens kapacitet til at beregne præcise kvantitative udtryk for funktioner i fortiden, har bragt historieforskningen op på et højere objektivitetsniveau. I første række må svaret være be-

kræftende. De tabeller, man kan få ud på basis af store statistiske materialer, er i sig selv - forudsat at de er fejlfri - udtryk for en objektiv sandhed. Men når det kommer til historikernes fortolkningsforsøg af tabellerne, er man faktisk på det samme niveau som altid. Alle der har skullet fortolke tabeller vil have været ude for de rent semantiske vanskeligheder at finde de rette udtryk, der hverken overdriver eller underdriver den tendens, der ligger i en række tal. Hverken korrelationskoeficienter eller significansprøver kan befri historikerne fra forpligtelsen til rent ordmæssigt at veje de talmæssige resultaters betydning og sætte dem ind i en større sammenhæng.

Forskningsassistent Ole S. D. Hansen:
Databanker for samfundsvidenskaberne i USA.

Når jeg har valgt at indlede om samfundsvidenskabelige databanker i USA, skyldes det dels, at USA er det land, som er længst fremme på dette område, dels at Det samfundsvidenskabelige Forskningsråd gennem etableringen af et forsøgsprojekt (Dansk Data Arkiv = DDA), har gjort problemerne aktuelle omkring samfundsvidenskabelige databanker. Jeg finder, at en gennemgang af de samfundsvidenskabelige databanker giver en god anledning til at diskutere problemer og muligheder omkring DDA.

I bilag 1 er de 24 databanker opført, som i 1967 var tilsluttet Council of Science Data Archives. Jeg har forsøgt at få nyere oplysninger om disse databanker, men har erfaret, at rådet er hensygnende efter at dets dynamiske formand, Ralph L. Bisco er død.

Når man ser oversigten igennem, er det muligt at opdele databankerne i flere grupper. (Der henvises til numrene i bilag 1).

1. Databanker indeholdende data fra marketingsbureauer:
   Roper Public Opinion Research Center (1)
   Louis Harris (8)
2. Databanker indeholdende data fra den offentlige statistik:
   Bureau of Labor Statistics (20)

3. Databanker indeholdende biografisk materiale om politiske ledere:
   Archive on Political Elites in Eastern Europe (19)
   Archive on Comparative Political Elites (13)
   Latin American Data Bank (24)
   ICPR (2)
4. Databanker indeholdende materiale om politisk sociologi (Political Science):
   ICPR (2)
   Charleton University (11)
5. Databanker indeholdende materiale fra ét stort sammenhængende projekt:
   Project Talent Data Bank (14)
6. Databanker indeholdende materiale fra kultursociologi m.v.:
   Human Relations Area Files (21)

Imidlertid kan databankerne også opdeles i banker, der modtager data fra større, faste leverandører:

Roper Public Opinion Research Center (1)
ICPR (2)
National Opinion Research Center (5)
Louis Harris (8)

og databanker, der indsamler materiale ved henvendelse til enkelte forskere:

MIT Social Science Data Bank (4)

eller blot korrespondere med forskere og fører et katalog over materialer (som forbliver hos de enkelte forskere) som

Social Science Data and Program Library Service (SSDPLS) (16)

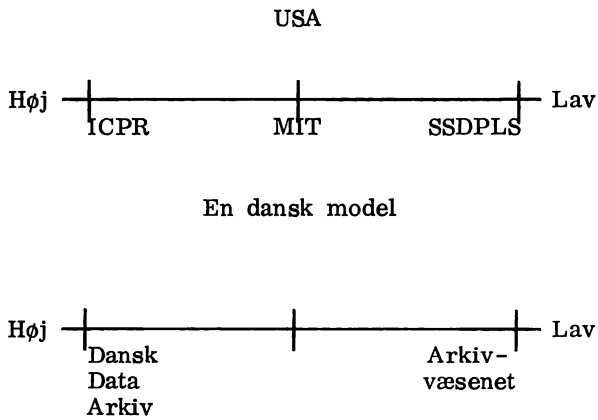
Databankerne kan også vurderes ud fra deres dataindsamlingspolitik, hvor vi i den ene ende af spektret finder data-

banker, der driver indsamlingsprojekter, som fx ICPR der oparbejder al USA's valgstatistik til maskinlæsbar form, og i den anden ende SSDPLS, som end ikke har datamaterialerne i banken.

Endelig kan man opdele databankerne efter de i banken værende materialers bearbejdningsgrad.

Det er yderst relevant at diskutere hvilken bearbejdningsgrad man skal tilstræbe fx. i DDA. I figur 1 har jeg illustreret en hypotetiks situation for, hvordan DDA og det øvrige arkivvæsen vil forholde sig med hensyn til bearbejdningsgrad.

Fig. 1. Bearbejdningsgrad.



Ud fra mange begrundelser ønsker jeg, at DDA får en så høj bearbejdningsgrad som muligt inden for givne rammer. Det betyder, at DDA efter min mening bør prioritere således, at man hellere må indskrænke antallet af datamaterialer og til gengæld sørge for, at disse bliver dels veldokumenterede og dels let til-

gængelige. Jeg tror, at veldokumenterede materialer vil medføre en høj udnyttelsesfrekvens blandt brugerne, både de videnskabelige brugere og de studerende fra samfundsvidenskaberne.

Der foreligger ikke egentlige statistiske oplysninger om USA-databankernes distribution, men af litteratur og årsrapporter fremgår det klart, at ICPR, som vel nok har den mest grundige bearbejdning af datamaterialer samt dokumentation, har den største efterspørgsel . ISPR's datasamlinger er langt fra de største, men der er bred enighed om, at ICPR's succes er baseret på den grundige dokumentation.

Det bør også nævnes, at der omkring de amerikanske samfundsvidenskabelige databanker er opstået livskraftige edb-miljøer, som har ført til udviklingen af program-pakker specielt for de samfundsvidenskabelige forskere. ICPR og ISR har således udviklet flg. systemer: OSIRIS/40, OSIRIS I, II, og III. NORC har udviklet SPSS, ligesom SSDB ved MIT har udviklet ADMINS. SSDPLS er ved at udvikle systemet SIMS.

Jeg vil tillade mig allerede nu at drage visse konklusioner på grundlag af erfaringerne fra USA, erfaringer som får virkning for DDA-projektet.

Emnekredsen for DDA vil ikke blive så specialiseret som i USA. DDA vil naturligt blive stillet overfor kravet om flere fagretningers data. Den politiske sociologi (Political Science) vil blive en betydelig gruppe, men ikke så dominerende som i USA, og endelig tyder oplægget for DDA på, at man her vil oprette en databank med høj bearbejdningsgrad, bl.a. fordi man er blevet pålagt at afprøve uddannelsesmæssige aspekter ved en samfundsvidenskabelig databank. På den anden side er der en lang række uafklarede problemer, særlig med hensyn til forholdet omkring omkostninger ved en bearbejdelse og ud-

byttet af højt bearbejdede data fremfor lavt bearbejdede. Det er væsentligt at indhøste disse erfaringer, idet meget store bearbejdelsesomkostninger måske drastisk vil indskrænke muligheden for indsamling af datamaterialer. Det er dog overvejende sandsynligt at der indenfor rimelige grænser vil være gode muligheder for at opretholde en vis standard uden at reducere antallet af materialer alt for drastisk.

Imidlertid kender vi ikke økonomien i et dataarkiv, og det er muligt, at forholdet er det, at en lille ændring i bearbejdelsesniveau betyder en udvidelse, henholdsvis indskrænkning i antallet af materialer, der kan indgå.

Selvom økonomien bag bearbejdningsgraden var kendt, kunne dette forhold ikke alene være afgørende for, hvilket niveau af bearbejdningsgrad man burde beslutte sig for. Sammenhængen mellem bearbejdningsgrad og anveldelse måtte også tages i betragtning.

For at tydeliggøre, hvad jeg forstår ved bearbejdning, henvises til bilag 2, som er hentet fra ICPR. Bilag 3 er en oversigt over materialer i Socialforskningsinstituttet, der er "klar" til at indgå i en samfundsvidenskabelig databank.

Bilag 1. Samfundsvidenskabelige databanker i USA (Status 1967)

Nr.	Navn	Adresse	Antal materialer	Antal kortbilleder i 1000	Bemærkninger
1	Roper Public Opinion Research Center	Williams College, Williamstown, Massachusetts	7000	12000	Samlingen er oprindelig opstået som arkiv for "Marketing-bureauet" Elmo Roper
2	Inter-University Consortium for Political Research	Institute for Social Research, P.O. Box 1248, Ann Arbor, Michigan 48106	50	3500	ICPR & ISR har produceret programpakkerne OSIRIS/40, OSIRIS I, II, III. ICPR er en "datter-organisation" af ISR
3	International Data Library and Referende Service	Survey Research Center, University of California, 2220 Piedmont Avenue, Berkeley, California 95720	500	2000	

Nr.	Navn	Adresse	Antal materialer	Antal kortbillider i 1000	Bemærkninger
4	M.I.T. Social Science Data Bank	Center for International Studies, E53-365, Hermann Building, Massachusetts Institute of Technology, Cambridge, Mass.02139	1100	1250	M.I.T.-SSDB har produceret programmellet ADMINS (under MAC) til brug for arkiv-analyse
5	National Opinion Research Center	University of Chicago, 6030 South Ellis Avenue, Chicago, Illinois 60637	95	1000	NORC har produceret programpakken SPSS
6	Political Science Research Library	Yale University, 89 Trumbull Street, New Haven, Connecticut 06520	-	-	Redistribuerer materialer fra andre databanker (væsentligst fra ICPR)
7	UCLA Political Behavior Archives	Dept. of Political Science Stastitical Laboratory, University of Calif., Los Angeles, California 90024	100	500	

Nr.	Navn	Adresse	Antal materialer	Antal kortbilleder i 1000	Bemærkninger
8	Louis Harris	Dept. of Political Science, Caldwell Hall, University of N. Carolina, Chapel Hill, North Carolina 27514	150	40	Data er væsentligst hulkort i multi-punch-form. Samlingen stammer oprindelig fra "Marketing-bureauet" Louis Harris
9	Columbia University School of Public Health and Administrative Medicine, Research Archives.	630 West 168th Street, New York N.Y. 10032	20	850	
10	Data Repositary Section Survey Research Laboratory	437 David Kinley Hall, University of Illinois, Urbana, Illinois 61801	14	-	
11	Carleton University, Social Science Data Archives	Dept. of Political Science, Carleton University, Colonel By Drive, Ottawa 1 Canada	200	300	

Nr.	Navn	Adresse	Antal materialer	Antal kortbilleder i 1000	Bemærkninger
12	International Development Data Bank	Michigan State University, 426 S.Kedzie Hall East Lancing Michigan 48823	50	-	
13	Archive on Comparative Political Elites	Dept.of Political Science, University of Oregon, Eugene Oregon 97403	-	-	Personorienteret arkiv for politikere i USA og en række europæiske stater.
14	Project Talent Data Bank	P.O.Box 1113 Palo Alto, California 94302	1	300 magnetbånd	PTDB er et forløbsprojekt for ca. 400.000 unge under uddannelse, startende i 1965, planlagt afsluttet i 1980
15	Laboratory for Political Research	Dept. of Political Science, University of Iowa, Iowa City, Iowa 52240	16	750	
16	Social Science Data and Program Library Service	Social Systems Research Institute, Social Science Building Univ. of Wisconsin, Madison, Wisconsin 53703	12	-	Overvejende meget omfangsrige tidsrækker. Udvikler programmellet SIMS. Distribuerer information om materiale og programmel

Nr.	Navn	Adresse	Antal materialer	Antal kortbilleder i 1000	Bemærkninger
17	Public Opinion Survey Unit	Research Center, School of Bussiness and Public Administration University of Missouri, Columbia, Missouri 65201	22	-	
18	Yale Economic Growth Center	Yale University 52 Hillhouse Avenue New Haven Connecticut	35	-	
19	Archive of Political Elites in Eastern Europe	Dept. of Political Science, 2312 Cathedral og Learning University of Pittsburgh, Pennsylvania	-	75	Biografiske oplysninger om ca. 1200 oersoner
20	Bureau of Labor Statistics	United States Dept. of Labor, Washington D.C. 20212	6 emner	22600	

Nr.	Navn	Adresse	Antal materialer	Antal kortbilleder i 1000	Bemærkninger
21	Human Relations Area Files	Yale University P.O.Box 2054 Yale Station, New Haven, Connecticut	-	40	Beskrivelse af 250 samfund. Over 450.000 tekstsider
22	Data Archive, Graduate School of Industrial Administration	Carnegie Institute of Technology, Pittsburgh, Pennsylvania 15213	120	300	
23	Council for Inter-Societal Studies	Northwestern University, 1818 Sheridan Road, Evanston Illinois 60201	50	100	
24	Latin American Data Bank	Dept. of Political Science, 107 Peabody Hall University of Florida, Gainsville Florida 32601	6	2600	

Kilde: Social Science Data Archives in The United States, 1967. Council of Social Sience Data Archives. 605, West 115th. Street, New York, N.Y. 10025 (September 1967)

Bilag 2.

Kort beskrivelse af ICPR's arkiveringsprocedure for survey data.

Kilde: ICPR: Annual Report 1970-71

ICPR har opstillet en beskrivelse for arkiveringsproceduren for et "standard-materiale" på 300 variable og 1.000 respondenter. Der gøres også rede for, hvilke kontroller, der ikke udføres.

1. (3-6 manduger)

Kontakt med den ansvarlige forsker med det formål at sikre sig en gennemgang af dokumentationen og at fornøden dokumentation og fornødne data er til stede, afmærket og i orden. Herunder en foreløbig kontrol af om tingene "plus", "minus", "blank" er anvendt som valide koder, og om der er anvendt alfabetiske koder.

Hvis der er en kompliceret datastruktur (fx forløbs-, panelundersøgelser) skal sammenhængen mellem de enkelte delmaterialer afklares. Det kodede materiale skal sammenholdes med spørgeskemaer, med det formål, at konstatere om sekvensen af de stillede spørgsmål er som de kodede variable. Er dette ikke tilfældet gøres der notat herom.

Rådata (hulkort) læses op på magnetbånd.

2. (under 1 manduge)

Data sorteres, flettes. Manglende hulkort erstattes, dubletter bringes i orden. Totalt antal afstemmes.

3. (under 1 manduge)

Materialet listes.

4. (2 manduger)

Marginaltabeller for samtlige variable optælles. Det kontrolleres, om de felter, der er angivet som ikke indeholdende data, er blanke.

5. (Estimat umuligt).

Kontrol for invalide koder. Hvis marginalfordelingen viser, at ulovlige værdier forekommer, undersøges det, om kodenøglen er komplet. Lovlige frekvenser på nul undersøges ligeledes. Multiple huller "spredes ud".

(Fejl i kodenøgle rettes - fejl i data føres på liste, rettes senere).

6. (3 manduger).

Dictionary kodes til OSIRIS og materiale indlæses.

7. (2 manduger).

Kodenøglen kodes ind som "codebook"-poster, der flettes ind i dictionary.

8. (5 manduger).

Dictionary (med codebook) rettes efter korrektur (dog kun for grove fejl).

9. (6 manduger).

Endelig korrektur og rettelse af dictionary og codebook.

10. (Estimat umuligt).

Vægte og problemer omkring forløbs- og panelundersøgelser bringes i orden.

11. (2 manduger).

Marginaltabeller køres igen for alle variable (indsættes samtidig i dictionary - codebook).

12. (2 manduger).

Indholdsfortegnelse og introduktion til materialet skrives og indlæses.

13. (under 1 manduge).

Det endelige materiale listes.

14. (2 manduger).

Codebook trykkes og gennemlæses endnu en gang.

Følgende udføres ikke:

Tekst for logisk sammenhæng på individniveau. Materialerne omkodes ikke til standardværdier som fx:

0 = logisk udelukket, spørgsmål ikke stillet

9 = uoplyst (nægtet).

Bilag 3.

Groft skøn over materialer for afsluttede undersøgelser i Socialforskningsinstituttet.

År	Navn	antal personer ca.	antal kortbilleder ca.
1961	Enlige ældre (Vesterbro)	1200	2400
1962	Unge mænds alkoholvaner	450	900
1964	Fysisk Handicappede I	10000	10000
1964	" " II	1100	5000
1965	Ældres levevilkår	2800	30000
1965	Privat alderdomssikring	3100	6200
1966	Fritidsvaner i Danmark	4000	40000
1968	14-20åriges uddannelse	3150	50000
1968	Børns opvækstvilkår	1300	5000
1969	30-45årige mænds alkohol-vaner	450	2000
1969	Gifte kvinder i erhverv	2600	15000
1970	Treskiftearbejde	600	6000
1970	Socialreformundersøgelsen	2300	11000
1970	Sociale konsekvenser af landbrugsvandringer	1450	7500
1971	Kommunale samarbejdsudvalg	1500	7000
1971	Saneringskvarterer	2500	10000
1972	Plejebehov hos ældre	4200	25000
17 materialer		42700	233000

Arkivar Grethe Ilsøe:

Er EDB-registrering af tingbøger et alternativ til udgivelse?

Baggrunden for dette indlæg er en EDB-behandling af 2.100 tingbogsindførsler på basis af Register til Malt herreds tingbøger 1657-1719 (ved Valdemar Andersen, udg. af Historisk Samfund for Ribe amt, 1969), som forsøgsvis er foretaget i Landsarkivet for Sjælland i samarbejde med Ole S.D. Hansen, Jens Højrup og Andreas Jørgensen. Forsøget er detaljeret beskrevet og kommenteret i Fortid og Nutid xxiv, s. 675-99, 1971. Der skal i almindelighed henvises til denne trykte redegørelse og her kun med henblik på den alternative problemstilling ridses visse hovedlinjer op, dels vedr. tingbogsmaterialet, dels vedr. forsøget.

Ved siden af kirkebøgerne (og rækkende længere tilbage i tid end disse) er tingbøgerne den eneste i princippet landsomfattende kildegruppe i tiden indtil omkring år 1700 til belysning af samfundet som det fungerede på menigmands plan og af det anonyme menneskes vilkår og reaktioner. Tingbøgerne blev i deres mest omfattende form påbudt i 1551. De skulle føres ved de sideordnede retter, herreds-, birke-, og byting samt ved landstingene og indeholder referater af det, der passerede på tinget, domme tingsvidner, skødebreve, skiftebreve

og alle andre "mærkelige" breve, der blev tinglæst. I slutningen af det 17. årh. udskiltes efterhånden protokoller med særligt indhold (skifteprotokoller, skøde- og panteprotokoller) af den egentlige tingbog. Tingbøgerne er særdeles uoverskuelige, de er ført kronologisk, tingdag for tingdag, de mangler registre og de er naturligvis skrevet med gotisk skrift, mere eller mindre vanskelig at læse.

Der kan ikke herske tvivl om det for den historiske forskning, forstået i bred forstand, værdifulde i at skabe en indgang til tingbogsmaterialet i 16. og 17. årh. og muligvis endnu højere op i tiden, idet en åbning for dette materiale vil være ensbetydende med at åbne for muligheden af en omfattende erkendelse af samfundsmæssige forhold og sammenhænge. Dette er naturligvis også baggrunden for at to af vore fornemme og traditionsrige udgiverselskaber, Kildeskriftselskabet og Landbohistorisk Selskab, er gået i gang med udgivelse af tingbøgerne. Udgiverarbejdet er som regel overdraget en enkelt eller to forskere, der sædvanligvis udfører det uden for vedkommendes egentlige faste ansættelse i en videnskabelig institution, altså som bierhverv, og med en eller flere tilsynsførende, udpeget af udgiverselskabet. Midlerne til udgivelsen søges i hvert enkelt tilfælde og kommer hovedsagelig fra offentlige fonds. Siden 1952, da udgivelsesarbejdet påbegyndtes og til dato er der på denne måde udgivet i alt ca. 3.500 tryksider, omfattende tingbøger fra en begrænset periode (fra nogle får år og op til en snes år) fra 3 herreder, 1 birk og 2 købstæder. Til sammenligning kan det oplyses, at antallet af bevarede tingbogsbind indtil år 1700 fra det sjællandske landsarkivs område (d.v.s. Sjælland, Møn, Lolland, Falster og Bornholm) når op på 251 herreds- og birketing-

bøger (repræsenterende 51 jurisdiktioner) og 288 bytingbøger (repræsenterende 18 jurisdiktioner), altså i alt 499 bind. Det totale antal bevarede tingbøger for hele landet indtil år 1700 balancerer skønsvis omkring 1.000 bind, snarere lidt flere.

Naturligvis har udgiverselskaberne aldrig haft til hensigt at udgive samtlige tingbøger, men kun ønsket at udgive nogle enkelte udvalgte og repræsentative med henblik på at demonstrere tingbogsmaterialets værdi for den historiske forskning i almindelighed og give mulighed for nærstudier i det konkrete, d.v.s. det udgivne materiale. Imidlertid må man notere sig, at udnyttelsen af de udgivne tingbøger er forbløffende beskeden. Kun sprog- og navneforskerne har benyttet materialet efter fortjeneste i forskellige systematiske undersøgelser. Herudover er der ikke fremkommet historiske arbejder, hverken i form af monografiske eller sammenlignende undersøgelser, hvori tingbogsudgaverne er udnyttet som centralt kildestof. Dette forhold må give anledning til at rejse spørgsmålet, om der evt. er andre muligheder for at åbne tingbøgerne for forskningen end gennem den fuldstændige og bogstavrette tekstudgivelse.

Når sprogforskerne har kunnet udnytte tingbogsudgaverne, skyldes det utvivlsomt de til hver udgave omhyggeligt udarbejdede, omfattende og konsekvente sted- og personregistre, altså navneregistre, som man umiddelbart kan arbejde ud fra. Derimod mangler de sagregistre, der, som flere allerede har påpeget, er en nødvendighed for at forskningsgrene som f.eks. social-, landbrugs-, rets- og kulturhistorien kan udnytte tingbøgerne. Fra tanken om at etablere saglige registre til de trykte tingbøger er springet ikke stort til tanken om at etablere saglige indgange til de utrykte (eller til dele heraf). I sin yderste konsekvens har en sådan ide i sig indbygget et alterna-

tiv til udgivelse. Og med EDB-teknikken har vi fået et hjælpemiddel i hænde, som skulle muliggøre opbygningen af et sagligt register og en varieret udnyttelse heraf.

Det nævnte registreringsforsøg kom i stand gennem et heldigt sammentræf af tilfældige omstændigheder. Og skønt det i enhver henseende må kaldes primitivt, er det dog af en så principiel karakter og dets resultater så overbevisende inden for deres begrænsning, at det må være rimeligt at gå videre med ideen om et EDB-registreringsprojekt af tingbøgerne. Maltregistret er opbygget i et kodesystem, hvorefter lokaliteter og emner angives ved hver sin talrække. Koden rummer betegnelser for 67 lokaliteter og for 139 emner. Hver tingbogsindførsel registreres i kronologisk opstilling under indførselsdato ved at kombinere betegnelse for hhv. emne og lokalitet, f.eks.:

datering	emne	lokalitet
3.5. 1662	86	019

Det principielle i denne registreringsmåde opfylder umiddelbart de krav, der stilles til et datamateriales komponenter. Derfor besluttede man at overføre registret til hulkort og fik det behandlet først efter et program udført i FORTRAN, derefter suppleret med et program i COBOL og med direkte læsbart output. Foruden den meget store fordel for forskeren, at udskrifterne forekommer i egentlig tekstform og ikke i symboler, indebærer sidstnævnte teknik, at de forskellige alfabetiske sorteringer af materialet direkte muliggøres.

Følgende variationer fremkom ved registerbehandlingen: a) sagtypefordelingen over hele perioden, b) sagtypernes fordeling på lokaliteter, c) den kronologiske fordeling af de enkelte sagtyper på år og måned, d) informationsudskrifter vedr. hver

enkelt sagtype, e) informationsudskrifter vedr. hver enkelt lokalitet, f) alfabetisk sortering efter sagtype med indførslerne vedr. den enkelte sagtype i kronologisk orden, g) alfabetisk sortering efter sagtype med udskrift af lokaliteterne i alfabetisk orden og indførslerne opført i kronologisk rækkefølge inden for den enkelte lokalitet, h) alfabetisk sortering efter lokalitet med indførslerne vedr. den enkelte lokalitet i kronologisk rækkefølge, i) alfabetisk sortering efter lokalitet med udskrift af sagtyperne i alfabetisk orden og med indførslerne opført i kronologisk rækkefølge inden for den enkelte sagtype.

Erfaringerne fra Malt-forsøget er i korthed disse: 1) en EDB-registrering af tingbogsmaterialet er realisabel og 2) det centrale problem er etableringen af de enkeltbegreber (=sagtyper), som sagregistreringen skal omfatte.

Der er fra datalogisk side peget på muligheden af at lade computeren selv foretage registreringen. Det kan ske ved at man indlæser hele tingbogsteksten, lader computeren alfabetisere alle forekommende ord og derefter skønsomt sletter de almene (og, men, hvad, der, som etc.). En anden mulighed er at foretage en registrering af overordnede emnebegreber. Herved vil registreringen fremstå som en bearbejdelse, en systematisering, der vil være meget værdifuld for et forskningsmæssigt overblik. Vel taber man i første omgang den enkelte slagbænk eller høstrive af syne, i stedet får man overordnede betegnelser som "inventar", evt. "bohave" og "landbrugsredskaber" foræret og muligheden (og sikkerheden) for gennem systematiske opslag at etablere kontakten til samtlige slagbænke og høstriver. Den utilfredshed, der evt. vil være med en overordnet emneregistrering, kan man imødekomme ved at programmere de etablerede overordnede emnebetegnelser så smidigt, at de

til enhver tid, hvis det bliver arbejdsmæssigt ønskeligt og muligt, kan opløses i underordnede begreber. Således vil man kunne forestille sig, at kompetente repræsentanter for de enkelte historiske forskningsgrene med tiden påtog sig denne opløsning eller specificering af de overordnede begreber, der specielt falder inden for deres fagområde. Sidstnævnte er en mulighed, ingen nødvendighed. En overordnet emneregistrering af et større antal tingbøger vil under alle omstændigheder være en storstilet forskningshjælp.

Primært er tingbogen en retshistorisk kilde, det vil derfor være både nærliggende og værdifuldt sideløbende med den skitserede emneregistrering at etablere et register, hvori hver enkelt sag klassificeres juridisk (f.eks. arvesag, gældssag, ran, tyveri, æressag etc.). Ved en kombination af emneregister og juridisk register opnår man dels at enkeltoplysningerne i førstnævnte får en extra dimension og nuanceres til lettelse ved den direkte indgang til materialet, - dels at de juridiske begreber karakteriseres gennem et konkret indhold og således vil kunne analyseres og udviklingen beskrives.

Det vil være rimeligt, om de udgiverselskaber, der hidtil har arbejdet med tingbøgerne, tager initiativet op og evt. gennem experimenter på baggrund af de udgivne tingbøger tilrettelægger et registreringsprojekt. At forsøge sig frem på de foreliggende udgaver vil reducere selve læsningsproblemet under experimentet og man vil i det mindste opnå at etablere en saglig indgang til de 3.500 tryksider. Hvis de nævnte udgiverselskaber ikke giver sig i kast med opgaven, kan den med lige så stor rimelighed tilfalde arkiverne, i hvis magasiner tingbøgerne jo befinder sig. I så fald vil den kunne anbefales til at indgå som et af Dansk Data Arkivs projekter.

Ved en afvejning af forholdet mellem den traditionelle udgivelse og en sagregistrering for EDB må forskellen i omfanget af det materiale, der lader sig tilgængeliggøre ved de to publiceringsformer, stærkt understreges. Med samme arbejdsindsats, målt i tid, som udgivelsen af et enkelt tingbogsblad kræver, vil man, når retningslinjerne for emneregistreringen er fastlagt, skønsmæssigt kunne skabe indgang til et tidobbelt bindantal. Endvidere må fremføres, at skønt man ved den traditionelle udgave opnår direkte tilgang til kildematerialet, herunder alle formuleringer, alle nuancer og alle detaljer, har den vist sig vanskelig at udnytte uden et sagregister. Sagregistret bliver således det springende punkt i afvejningen. Et traditionelt sagregister vil oftest kun unuanceret henvise til det blotte sagforhold uden angivelse af dets forekomst i forskellige kombinationer. Dette ses f.eks. tydeligt i traditionen for den praktisk og økonomisk betingede adskillelse mellem sagregistre og hhv. person- og stedregistre, evt. kronologiske eller juridiske registre - et forhold der i høj grad vanskeliggør muligheden for at få forskningsmæssige givende og relevante kombinationer. En sagregistrering for EDB vil derimod principielt kunne indebære en systematisk registrering af enhver brugbar information med henblik på udnyttelse af denne i principielt enhver kombination, at begrænsningerne for kombinationsmulighederne kun ligger i den begrænsning, der knytter sig til valget af maximum for informationernes antal.

Universitetsadjunkt Jørgen Elklit:
Det politisk-økologiske EDB-arkiv ved Århus Universitet.

Jeg skal i denne indledning søge at belyse:

1. arkivets indhold
2. arkivets struktur
3. arten af data-bærere
4. det institutionelle tilhørsforhold.

Hermed være dog ikke sagt, at jeg tror at kunne etablere sikker og præcis viden. Jeg har selv kun haft ansvaret for oprettelsen af én af de 7 dele, som vort dataarkiv for nærværende består af, og selv om jeg i beskedent omfang har medvirket ved opbygningen af et par af de andre, er det derfor tænkeligt, at jeg ikke er i stand til at fremlægge alt entydigt eller besvare alle spørgsmål. Men formålet er vel også snarere at lægge op til diskussion om nogle af de involverede problemer, end det er at manuducere tilhørerne i, hvad vi har, og hvad vi ikke har, i vore forskellige arkiver.

Et andet forbehold, jeg straks vil tage, er mht. min egen datalogiske kunnen og viden. Den er beskeden. Jeg er, som ca. halvdelen af de tilstedeværende, historiker af uddannelse og er nu ansat ved et institut for politisk videnskab.

Det er min faste overbevisning, som jeg bl.a. bygger på mine egne erfaringer, at man inden for et fag som historie eller statskundskab kan klare sig ganske langt - og skal kunne klare sig ganske langt - med hensyn til udnyttelse af de fleste edb-mæssige faciliteter med blot et beskedent kendskab til databehandlingens basale principper. Arbejdsdelingen som princip er også her af afgørende betydning for en fortsat udvikling. Hvis man samarbejder med en konsulent på et for den pågældende opgave passende datalogisk stade, kan man få løst de fleste problemer - og især da hvis han inddrages i en tidlig fase af projektet. I nogle tilfælde må man betale sig fra det, men det vil som regel alligevel være det mest rentable. Forskningsrådene har hidtil gennemgående set med milde øjne på ansøgninger om støtte til sådanne formål i de situationer, hvor der ikke gennem andre projekter eller lign. på den pågældendes institution allerede var skabt kontakt med passende datamatisk know-how. Personlige kontakter har også for nogle historikere løst problemer af denne art uden udgift for de pågældende.

Hertil kommer, at der nu forskellige steder er implementeret såkaldte "program-pakker", dvs. færdige systemer af programmer, der umiddelbart kan udnyttes til blandt andet de fleste af de gængse former for kvantitative analyser, dersom blot forskerens data er indlæst på den for det pågældende programsystem foreskrevne måde, fra hulkort eller magnetbånd.

Vi har således ved mit institut adgang til at bruge det såkaldte SPSS-system (1), som er installeret ved Århus Universitets regnecenter, indkøbt for den beskedne sum af 400$ af et af vore forskningsprojekter, hvor deltagerne fandt, at det var en god ting at investere i. Instituttet har siden overtaget det og psykologer, økonomer o.a. er derefter også begyndt at bruge det. Det er me-

get nemt at bruge, kan lave hvad man normalt har brug for ved databehandling af et samfundsvidenskabeligt - herunder historisk - datamateriale, giver velredigerede og nemt læselige udskrifter. Den store hage ved SPSS er, at det ikke er ret fleksibelt i sine indre dele, og at det derfor kræver forholdsvis megen plads i computeren, når det kører. Derfor er vi vist ikke særligt populære på RECAU, men det er et andet problem. Desuden er det næsten for fintfølende over for ubetydelige småfejl, som det reagerer kraftigt på. Man abonnerer på systemet og får derfor oplysning om fundne fejl og rettelseskort hertil og en gang imellem en ny og forbedret udgave, som kan noget mere. En anden af disse færdigsyede program-pakker er det såkaldte OSIRIS, som f.eks. Socialforskningsinstituttet bruger.

Mit principielle synspunkt er altså, at de forskellige, store og små, problemer af teknisk art, som anvendelsen af edb inden for de historiske og andre samfundsvidenskaber rejser, normalt må søges løst - ja, ofte kun kan løses - i et snævert samarbejde med folk, der har det som deres fag. Herved bliver det muligt for os andre - på et væsentligt forbedret og væsentligt mere overskueligt data-grundlag - at beskæftige os med det, som vi i princippet har forstand på.

Dataarkivet ved Institut for Statskundskab består faktisk af to helt adskilte arkiver. Ældst af de to arkiver er:

I. Det politisk-økologiske kommune-arkiv. (2)

Dette består igen af tre arkiv-dele:

1. Et arkiv med oplysninger hentet fra den offentlige (og for det meste også offentliggjorte) statistik vedr. folketingsvalg, folketællinger o.a. for perioden 1920-45.

2. Arkivets oprindeligste del, nemlig de samme typer af oplys-

ninger fra tiden fra og med valget i 1947.

3. Kommunevalgsdata 1909-66.

Det andet arkiv ved instituttet - eller den anden gruppe af arkiver, om man vil - er de såkaldte:

II. Politiker-arkiver (3).

Her er der 4 underafdelinger:

1. Et arkiv, der omfatter alle folketingsmedlemmer i perioden 1849-1968, i alt godt 1700 .
2. Tilsvarende for medlemmerne af landstinget frem til 1953.
3. Et minister-arkiv for samme periode, som MF-arkivet, dvs. 1849-1968, og endelig det seneste og lidt afvigende skud på denne stamme:
4. Folketingskandidat-arkivet for perioden 1960-71.

Jeg ser ingen grund til i detailler at gennemgå arkivernes indhold, men jeg finder anledning til at understrege noget, som jeg finder meget væsentligt. Det er nødvendigheden af en omhyggelig dokumentation. Det gælder ved alle former for databehandling, at man selv og andre skal kunne rekonstruere det gjorte arbejde, dersom det er nødvendigt. Det gælder om muligt i endnu mere absolut form ved data-arkiver, uanset deres indhold, hvor andre er direkte afhængige af det indsamlings- og etableringsarbejde, der er blevet gjort, og hvor brugerne derfor har et helt indiskutabelt krav på at få at vide, hvilket materiale, der ligger, hvorledes det er indsamlet, kodet, renset - etc., og hvorledes arkivet er etableret set fra den edb-mæssige side - altsammen for at evt. udefra kommende brugere kan afgøre, om de kan have glæde af at trække på det indsamlede materiale, og om det er muligt for dem.

Det fører mig over i en anden problemstilling, og det er problemerne omkring koordinationen med andres kendte eller ukendte behov i forbindelse med en registrering af et data-sæt. Der er selvfølgelig kun en vis valgfrihed med hensyn til, hvilke oplysninger der kan medtages om de enheder, man vil opbygge et data-arkiv omkring (eller et register til andet formål, der siden kan bruges af andre forskere) - man er afhængig af sit kildemateriale, arbejdsmæssige begrænsninger etc. Der er også ofte visse naturlige grænser for antallet af måder, det kan kodes på. Ønsket om at kunne komparere med udenlandske anstrengelser på det samme område eller med tidligere forskning inden for feltet begrænser også ofte handlefriheden, selv om det selvfølgelig er bånd, man selv lægger på sig. Hvis der overhovedet bliver nogen handlefrihed tilbage, rejser der sig et ikke uvæsentligt spørgsmål: hvor meget skal man tage hensyn til evt. kommende, andre brugere? Intet, fordi man selv har haft alt besværet og udmærket kan indrette tingene efter sit eget hoved, på grundlag af de hypoteser, man selv arbejder med, og som man selvfølgelig finder er nogle af de vigtigste og mest interessante, der nogen sinde er blevet formuleret inden for området? Eller alt, fordi man ønsker sine resultater prøvet og anvendt af andre og derfor vil lette dem i deres arbejde, og fordi man mener, at man ikke ene mand har lov til at sætte sig for tungt på de knappe forskningsmidler, der er her til lands - materialet skal være fuldt anvendeligt for andre. Gennemførelsen af det sidste, filantropiske synspunkt støder ofte på den hindring, at man ikke ved, hvilke behov fremtidig forskning vil have på det pågældende område. Som regel træffes vel de relevante afgørelser ud fra en sammenblanding af de forskellige motiver, men jeg tror dog, at hensynet til egne forskningsproblemer rent fak-

tisk ofte er dominerende - så må andre "take it or leave it".

Der ligger her et betydeligt koordinations- og kommunikationsproblem - hvorledes skal man på et område, der hidtil i alt væsentligt har været båret af forskere, der mere eller mindre har arbejdet alene, sikre, at anvendeligheden af disse materialer maksimeres? Og vel at mærke uden at genere de pågældende væsentligt! Og princippet om forskningens frihed, hvorledes skal det gå med det, hvis man etablerer et koordinationsorgan, der vel skulle vejlede etc., men som hurtigt kunne komme til at stå som dirigerende? Arbejdet med etablering af et samfundsvidenskabeligt data-arkiv er som bekendt nu kommet ind i mere faste rammer, blandt andet fordi der efterhånden forskellige steder følte et betydeligt behov herfor; jeg tror dog at Fællesudvalget for historisk Forskning alligevel bør overveje at oprette et mere permanent organ, der på en eller anden måde kunne påtage sig visse koordinerende, rådgivende og - måske - retningsgivende funktioner. Måske bør de historikere, der er i Dansk Dataarkivs referencegruppe, have en baggrundsgruppe - det er jo så moderne. Men det kan også være nyttigt - også for de pågældende.

Jeg er efterhånden kommet lidt langt væk fra det, der egentlig var mit udgangspunkt - data-arkiverne ved Institut for Statskundskab. Det skyldes, at jeg i en mere generel form har villet pege på nogle af de problemer, som vi har haft at arbejde med, og som også vil dukke op andre steder, hvor et tilsvarende arbejde tages op.

Det ældste af vore arkiver, det politiks-økologiske kommunearkiv, blev i sin opbygning fra 1962 ledet af afdøde afdelingsleder Jan Stehouwer, senere af ham og nuværende lektor Ole Borre i snævert samarbejde. Ved det i 1959 oprettede In-

stitut for Statskundskab ville man indsamle materiale om dansk vælgeradfærd og bl.a. relatere den til den sociale struktur - af hensyn til et sådant materiales anvendelse i forskning og i undervisning. Interviewundersøgelser var dengang som nu dyre, hvorimod en reproduktion af den offentlige statistik for enheder af en passende størrelse - kommunerne - var overkommeligt i penge og i arbejdskraft. Man startede med valget i 1960 og folketællingen samme år, og udvidede siden arkivet frem og tilbage i tid, så det til sidst omfattede hele perioden 1947-68. Man bestemte sig på et eller andet tidspunkt til at komme ud over problemet med de mange kommuneændringer ved at regne alle oplysninger tilbage fra 1947 om, så de svarede til den kommunale inddeling, der var gældende ved folketingsvalget i 1966. Derved fik Borre og Stehouwer løst ét af deres problemer, idet de nu havde godt 1150 enheder, der var veldefinerede og til at bruge som udgangspunkt for deres analyser, der især koncentrerede sig om den politiske udvikling i 60'erne ud fra en økologisk synsvinkel. Problemet var bare, at en del af de enheder, der så arbejdedes med, var kunstprodukter, kun med eksistens i dataarkivet. Da jeg siden overtog arbejdet med at udvide arkivet til også at omfatte mellemkrigstiden og besættelsen, kunne jeg ikke se nogen fornuft i at lade 1966 være mit kronologiske fikspunkt, og jeg udviklede derfor et andet system, der tager hensyn til ændringer i den kommunale struktur. Det er nøjere beskrevet i vores arkivfører. Hermed var et direkte samarbejde mellem de to arkivdele udelukket, i hvert fald hvis der ikke udvikledes store og vanskelige konverteringsnøgler. Jeg nævner dette som et eksempel på de vanskeligheder, der kan være ved et arbejde af denne art, selv når det udføres på én og samme institution af mennesker, der personligt, ideolo-

gisk og på anden måde står hinanden nær. Problemer med afgrænsninger og ændringer deri har været centrale alle de steder, her og i udlandet, hvor man på den ene eller den anden måde har opbygget arkiver med oplysninger over tid om geografisk-administrative enheder, og forskellige løsningsforslag har været bragt i anvendelse - ikke uden konsekvenser for det internationale samarbejde på feltet, som alligevel er ganske betydeligt.

Men det centrale er altså, at vi i det politisk-økologiske data-arkivs tre dele har samlet - på kommune-niveau - en meget betydelig del af den offentlige statistik siden tiden omkring 1. verdenskrig og frem til kommunalreformen af 1970. Det drejer sig især om oplysninger vedr. politiske fænomener og om den sociale sammensætning af de samme enheder. For en fuldstændigheds skyld bør jeg vel nævne, at vi i vore aktuelle forskningsprogrammer - muliggjort blandt andet af mere rigelige bevillinger fra Statens samfundsvidenskabelige Forskningsråd og andre organer - også arbejder med interview-undersøgelser af enkelt-individer.

Politiker-arkiverne er, på samme måde som kommune-arkivet oprindelig var det, i deres udspring knyttet til en enkelt mand, der på instituttets vegne har modtaget en bevilling til at få dem etableret. Det drejer sig om Mogens N. Pedersen, der i øvrigt nu har forladt instituttet. Hans interesser, teoretiske baggrund etc. har i et vist omfang været afgørende for, hvad der er kommet med i arkivet fra 1849 og frem - selv om det selvfølgelig er sket i diskussion med blandt andet historikere. Kildegrundlaget har sat andre grænser, idet det har været et ønske, at de medtagne variabler helst skulle kunne belyses for hele perioden.

Endelig vil jeg trække et andet væsentligt problem frem, nemlig spørgsmålet om adgangen til sådanne arkiver - skal der være fri adgang, eller skal nogen - f.eks. den der har stået for oprettelsen af det - have ret til at udelukke nogen fra at bruge det? Det sidste lyder selvfølgelig ikke så pænt - specielt ikke når arkiver af denne type normalt er oprettet for offentlige midler og evt. i konkurrence med andre projekter.

På den anden side er det vel også rimeligt at sikre den stakkel, der har brugt år af sit liv på at oprette et sådant arkiv en vis mulighed for også at få tid til at arbejde med det inden andre kommer og - fulde af gode ideer, ikke nedslidte af indsamlingsarbeidet -bruger det bedste derfra i deres egen forskning. Normalt har man derfor sat en eller anden - ret vilkårlig - tidsgrænse på et par år, hvorefter der i princippet var fri adgang. Nogle steder - blandt andet i Norge - han man dog indrettet sig sådan, at nye, udefra kommende brugere, skal give noget til gengæld - f.eks. i det aktuelle tilfælde den fil, hvor den nye bruger har sammenstillet ensartede oplysninger om danske og norske parlamentarikere.

Til allersidst vil jeg gøre opmærksom på et rent teknisk problem. Det er de mange forskellige måder, vore arkiv-dele ligger på. Noget ligger på hulkort, noget på bånd, noget på disc packs. Et par af arkivdelene findes samtidig og i næsten identisk form på flere af disse medier. Specifikationerne er også givet ret forskelligt i detaljeringsgrad. Forklaringen på dette - som jeg kun ser som et mindre, teknisk problem - er dels savnet af et overordnet styrende organ med ansvar for hele systemet af arkiver, hos os eller andet steds, dels at vi i disse oversigter har villet lade de forskellige medarbejdere fremlægge deres rent faktisk forskellige ting selv.

Ved disse tekniske forskelligheder understreges kun det, jeg har været inde på flere gange, og som Dansk Data-arkiv måske kan hjælpe os til at få gjort noget ved: data-arkiverne ved Institut for Statskundskab er ikke en enhed - det er mindst 5 disparate arkiver, som blot er etableret i én og samme institutions regi.

Noter:

1. Norman Nie, Dale H. Bent, and Hadlai Hull, SPSS, Statistical Package for the Social Sciences, New York: McGraw-Hill, 1970.

2. Karl-Henrik Bentzon, Ole Borre og Jørgen Elklit, Kommunedata-arkivet, 1909-68. En foreløbig oversigt over det politisk-økologiske data-arkiv ved Institut for Statskundskab, Århus, Institut for Statskundskab, 1972.

3. Mogens N. Pedersen, Danske Politiker-Arkiver - en oversigt over data-arkiver indeholdende oplysninger om danske politikere 1848-1971. Århus: Institut for Statskundskab, 1972.

Forskningsassistent Svend Bang:
Problemer omkring tilgængeligheden af data i nogle større offentlige EDB-registre.

Tanken med dette indlæg er at opridse nogle af de problemer m.h.t. bevaring og tilgang til de store offentlige administrative registre (CPR-registret, kildeskatteregistret osv.), som rejser sig derved, at disse registre ikke foreligger på traditionelle "papirmedier", men udelukkende på maskinlæsbare medier.

Traditionelt har forholdet været det, at alt administrativt materiale bevaredes med mindre det offentlige arkivvæsen gav tilladelse til kassation. Gennem sin forvaltning af kassationstilladelser har arkivvæsenet kunnet sikre og varetage forskningens tilgang til relevant materiale.

Denne situation er med opbygningen af de administrative EDB-registre radikalt forandret, idet der i selve disse registres systemopbygning indgår løbende kassation i form af sletning af ukurante eller ikke længere administrativt relevante data. Der foregår således en faktisk kassation af data, som arkivvæsenet ikke har haft mulighed for at tage stilling til.

Hvis forskningens interesser skal varetages, vil det være nødvendigt, at man explicit udtrykker, hvilke data og hvilke

relationer mellem disse, man ønsker udtrukket af registrene til bevaring. I et vist omfang vil det sikkert være nødvendigt, at der gribes ind allerede i den fase, hvor de datamatiske systemer, som ligger til grund for registre, udformes.

Denne stillingtagen til bevaring af data og relationerne mellem dem kan ikke ske alene på grundlag af en traditionel faglig arkivmæssig kompetence, men forudsætter en syntese af datamatisk og arkivmæssig kompetence.

En anden kreds af problemer, som er opstået ved anvendelsen af EDB i de offentlige administrative registre, knytter sig til den fysiske holdbarhed af de maskinlæsbare medier og til de maskinmæssige muligheder for genanvendelse på kortere og længere sigt. Vi må i dag på den ene side konstatere, at vi ved meget lidt om maskinlæsbare mediers fysiske holdbarhed, og at der på grund af EDB-teknikkens hurtige udvikling knytter sig en betydelig usikkerhed til mulighederne for genanvendelse af medier, der blot er få år gamle. På den anden side er det evident, at den praktiske mulighed for forskningen til at skaffe sig tilgang til EDB-registrenes data er snævert knyttet til, at der kan udvikles metoder og procedurer, som sikrer stadig overvågning af medierne og i nødvendigt omfang konvertering, enten til nye maskinlæsbare medier eller f. eks. til film.

Det kan som konklusion fastslås, at opbygningen af de offentlige administrative EDB-registre stiller krav til en kombination af datamatisk og arkivmæssig kompetence, som endnu ikke er kendt. Det bør dog samtidig understreges, at denne syntese af viden og kunnen er den samme som udgør forudsætningen for, at man ved hjælp af EDB-teknikken kan gøre de store masser af traditionelle "papirarkivalier" tilgængelige for

forskningen gennem massebehandling af råtekster, bearbejdning af manuelle registre, informationssøgning osv.

DELTAGERLISTE

Bang, Svend, forskningsassistent
Bjerg, Hans Chr., arkivar
Bunzel, Henning, cand. oecon.
Boje, Per, forskningsstipendiat
Dalgård, Sune, overarkivar
Due-Nielsen, C., universitetsadjunkt
Elklit, Jørgen, universitetsadjunkt
Gissel, Svend, dr. phil.
Gunn, J. H., direktør
Hansen, C. Rise, arkivar
Hansen, H. Bernt, universitetsadjunkt
Hansen, Ole S. D., forskningsassistent
Hartyani, M. de, forskningsbibliotekar
Hvidt, Kristian, folketingsbibliotekar
Ilsøe, Grethe, arkivar
Jørgensen, Frank, arkivar
Jørgensen, Erik Stig, arkivar
Kristensen, Karl Evald, forskningsbibliotekar
Leunbach, G. H., afdelingsleder
Liversage, David, museumsinspektør
Pedersen, Finn Stendal, universitetsadjunkt

Rasmussen, A. Hjort, forskningsstipendiat
Rasmussen, Klaus U., konsulent
Smith-Nielsen C., civilingeniør
Stoklund, Bjarne, professor
Svarre, Peter, konsulent
Sveinsdottir, Edda, afdelingsleder
Thomsen, Niels, universitetslektor
Vedel, Eline, stud. scient.
Worre, Torben, seniorstipendiat
Voss, Olfert, museumskonsulent

DIS-Danmark
118203