

Kvalitative data i Dansk Data Arkiv

af Anne Sofie Fink, Dansk Data Arkiv

Inden for samfundsvidenskab har den dominerende metode til at foretage empiriske studier været at foretage kvantitativ dataindsamling. Indsamling af kvantitative data bliver dog udfordret af empirisk forskning baseret på kvalitative data indsamlet fx ved forskningsinterviews.

Det er Dansk Data Arkivs opgave at indsamle, arkivere og formidle bevaringsværdige, maskinlæsbare data produceret i forbindelse med samfundsvidenskabelige forskningsprojekter. Denne opgave løses i dag med ensidigt fokus på kvantitative data, men det er DDA's intention, at også kvalitative data skal inddrages i arkivets samling. Denne artikel vil kommentere DDA's initiativ til arkivering af kvalitative data. Da dataarkivvirksomhed er ukendt for mange, vil jeg indlede med kort at præsentere dette emne.

Arkivering af samfundsvidenskabelige forskningsdata er en opgave, som mange lande har valgt at håndtere centralt, der findes eksempelvis nationale dataarkiver i de fleste EU-medlemslande. Dataarkivvirksomhed er ukendt for mange, og jeg skal derfor indlede med lidt historisk baggrund samt præsentere rationerne bag arkivering og formidling af samfundsvidenskabelige forskningsdata. Herefter vil jeg bevæge mig over i artiklens egentlige emne nemlig 'arkivering af kvalitative data' og resummere materiale fra 6 kvalitative interviews om kvalitativt feltarbejde gennemført i foråret/sommeren 2000 med samfundsvidenskabelige forskere, der arbejder kvalitativt. Som afslutning vil jeg pege på undersøgelsens implikationer for DDA's håndtering af arkivering og formidling af kvalitative data.

Om dataarkivering

Initiativer vedr. arkivering af samfundsvidenskabelige data begyndte i USA, da anvendelse af spørgeskemaundersøgelser som dataindsamlingsmetode for alvor tog fart på universiteterne. Det første egentlige arkiv for samfundsvidenskabelige data begyndte at fungere i begyndelsen af 1960'erne ved Universitetet i Michigan. I Europa dukkede de første dataarkiver op i slutningen af årtiet i Köln, Colchester og Bergen. I Danmark oprettede man i 1973 Dansk Data Arkiv (DDA) som et forsøg med finansiering fra Statens Samfundsvidenskabelige Forskningsråd. I 1978 fik DDA permanent status placeret som institut ved Odense Universitet. I 1993 blev DDA gjort til en enhed i Statens Arkiver som Rigsarkivet, Erhvervsarkivet og landsarkiverne. DDA var fra begyndelsen tænkt bredere end dataarkiver andre steder i verden, idet arkivets virksomhed skulle rette sig mod *samfundsbeskrivende* data og ikke begrænses til data indsamlet under de samfundsvidenskabelige

fakulteter. Samlingen af datamaterialer i DDA rummer i dag både historiske og sundhedsvidenskabelige materialer ved siden af de samfundsvidenskabelige.

Hvorfor arkivere forskningsdata?

Spørgsmålet, der rejser sig, når emnet arkivering af forskningsdata præsenteres, er naturligvis ”*hvorfor bekymre sig om disse data, der sandsynligvis er udnyttet til en eller flere publikationer?*” Svaret på spørgsmålet kan fremstilles i to perspektiver: det kollektive perspektiv – de samfundsmæssige fordele, der er forbundet med dataarkivering og –formidling og det individuelle perspektiv – fordelene ved dataarkivering og genanvendelse af data for den enkelte forsker.

De kollektive fordele

Samfundsvidenskabelige undersøgelser vil afspejle den tid, de er foretaget i og derfor udgøre et billede af samfundet i nu’et, som det ud fra en historisk betragtning kan være værd at bevare. For det andet er empirisk forskning meget kostbar, og ofte er en væsentlig del af omkostningerne ved et projekt udgifterne til indsamling af data.¹ Arkivering kan ses som dokumentation af en undersøgelses gennemførelse over for diverse interessenter (finansielle kilder, respondenter, interesserede/berørte parter, offentligheden generelt etc.). Desuden berører den primære analyse af data i mange tilfælde kun et begrænset udsnit af det indsamlede datamateriale. Ved at gøre data til genstand for sekundær analyse af andre forskere er det muligt at kapitalisere yderligere på de anvendte forskningsmidler og samtidigt medvirke til en mere bredspektret anvendelse af et materiale.

Muligheden for sekundær analyse vil desuden kunne medvirke til at hæve det nationale forskningsniveau ved at give andre forskere mulighed for at verificere og ’bygge ovenpå’ oprindelige analyseresultater. Endelig betyder genanvendelse af data, at de krav, der lægges på befolkningen om at agere respondenter i forbindelse med diverse samfundsvidenskabelige undersøgelser, begrænses (Hakim, 1982).

De individuelle fordele

For den enkelte forsker indebærer arkivering af egne datamaterialer, at disse opbevares forsvarligt og ’uforældeligt’ – sidstnævnte henviser til, at data vil være tilgængelige uanset ændringer i programmer eller formater. Samtidig kan forskeren henvise til et sted, hvor dokumentationen for pub-

¹ Disse omkostninger afholdes ofte af midler fra Statens Samfundsvidenskabelige Forskningsråd (SSF), der stiller det som en betingelse, at data arkiveres.

licerede undersøgelsesresultater findes. Fra USA kendes det, at forskere i publikationer henviser til, hvor data findes opbevaret (Hakim, 1982), som man i dag kender det i forhold til andre informationskilder, der er anvendt i en publikation. En sådan åbenhed i forhold til adgang til data kan synes meget relevant i publikationer, der overvejende baserer sig på empirisk materiale. Adgang til data kan tillige forventes at kommunikere en høj grad af lødighed i det pågældende forskningsprojekt.

Forskere, der arkiverer data i DDA, kan til enhver tid få udleveret data – eksempelvis i tilfælde af, at de vil anvende materialet på ny eller ønsker at verificere oprindelige analyseresultater i tilfælde af kritik eller andet.

Fordelene ved genanvendelse af data er eksempelvis muligheden for omfattende analyse af enkelte forhold på tværs af mange undersøgelser, analyse med et andet teoretisk udgangspunkt end primærforskerens og anvendelse af andre værktøjer til analyse end primærforskeren. Et aspekt kan også være, at eksisterende datamaterialer analyseres i forhold til aktuelle samfundsmæssige debat-*emner/issues* (Hakim, 1982). Herudover findes en række aspekter af mere pragmatisk karakter – eksempelvis manglende tid, lyst og midler til at gennemføre eller deltage i dataindsamlingsprojekter.

Forskelle mellem datatyper i relation til arkivering og anvendelse

Op gennem 1960'erne og 70'erne syntes den dominerende dataindsamlingsmåde inden for samfundsvidenskab at være spørgeskemaundersøgelser. Da dette var perioden, hvor dataarkiverne blev etableret, er praksis for dataarkivering udviklet i nær sammenhæng med arkivering af kvantitative data. DDA's samling af datamaterialer rummer i dag kun enkelte kvalitative materialer.

For kvantitative data synes det oplagt, at disse må opbevares et sikkert sted, så de fremtidigt kan findes frem. Det kan være hos forskeren selv eller i et arkiv, der er specialist til opgaven. For forskeren indebærer arkivering og evt. andres anvendelse af data kun begrænset risiko, idet data betragtes som "hard, objective and rigorous" (Burgess, 1984; 3). Anderledes er det for forskeren, der indsamler data i form af forståelser baseret på udsagn og indtryk, som 'bæres med' i erindringer. Arkivering og genanvendelse af den slags data er langt fra så oplagt, og forskeren kan opfatte dette som forbundet med stor risiko, idet data, der af natur er "soft, subjective and speculative" (ibid.), må 'eksternaliseres'.

Når kvalitative data så alligevel opbevares i eksterne medier, skyldes det, at det af praktiske grunde kan være fordelagtigt, at (en del af) 'historierne' i data holdes fast. Selv om den historie, primærforskeren vil fortælle, kun ufuldstændigt findes i filer, på papir, på fotos og andre medier,

mens resten findes i forskerens erindring, så er 'historierne', som de fremstår i arkiveringsbare data, en ressource i sig selv.

En række spørgsmål rejser sig imidlertid i forbindelse med arkivering og formidling af kvalitative materialer. Som et middel til at besvare disse spørgsmål har vi i DDA valgt at gennemføre en interviewundersøgelse med fokus på barrierer over for dataarkivering set med forskernes øjne.² I det følgende vil jeg præsentere resultater af undersøgelsen i henhold til etiske og praktiske aspekter, der indvirker i forhold til arkivering af data på det individuelle plan og til slut diskutere anvendelse af arkivdata.

Det etiske...

Når Kvale (1994; 115) betegner en interviewundersøgelse som *"et moralsk foretagende"*, kan det tilskrives den rolle, som forskeren spiller i den kvalitative undersøgelsesproces. I kraft af at forskeren indgår i en personlig relation med sine respondenter, følger et etisk ansvar. Også den kvantitative forsker har naturligvis et etisk ansvar, men hans mere tilbagetrukne rolle gør det mindre påtrængende i relation til arkivering og andres anvendelse af data. Nedenstående kommentar giver en karakteristik af kvalitativt feltarbejde:

"Hver eneste gang du er ude at snakke, har du været konfronteret med nogle andre mennesker, du har siddet og haft en dyb og nær samtale med dem om måske et ligegyldigt emne, men altså du er blevet konfronteret med nogle andre mennesker, og det sætter da noget i gang i dig selv også, man er jo ikke bare ude som sådan et stykke værktøj, man er jo derude som menneske og den der menneskelige interaktion, den tror jeg er vigtig. Det gør så også det de her data, det er jo ikke kun repræsentationer af respondenter siger, det er også en repræsentation af, hvad du har oplevet, det er faktisk en del af din personlige historie som forsker. Og derfor tror jeg, jeg gætter, det kunne være derfor at man har et andet forhold til data, fordi jeg er ligeså meget selv til stede i det interview, som respondenteren er det." (ph.d.studerende, marketing)

Kommentaren præsenterer 3 nøgleaspekter (understregede) – fortrolighed – interviewet er en *"en dyb og nær samtale"*; loyalitet – *"man er jo derude som menneske"* og gensidighed – *"jeg er jo ligeså meget selv til stede"* i interviewet. I det følgende skal jeg komme ind på, hvordan disse etiske aspekter optræder og influerer spørgsmålet om arkivering og genanvendelse.

² Lidt om undersøgelsens metode: Forskerne, der deltog i undersøgelsen, var udvalgt, så gruppen havde bredde mht. forskningsgren og -felt. Fælles for de interviewede forskere var brugen af kvalitativ forskningsmetode, primært ved anvendelse af interviews som teknik til dataindsamling. Interviewene foregik semistrukturerede med udgangspunkt i en interviewguide. Båndoptagelser af interviewene blev transskriberet og derpå analyseret i henhold til temaer, der dels var prædefinerede, dels var inspireret af data.

Fortrolighed

Fortrolighed kan siges at udgøre en forudsætning for, at der gennem samtalen kan indhentes den type information/data, som den kvalitative forsker bruger. Denne fortrolighed skal endda række ud over interviewsituationen, så respondenterne har tillid til, at forskeren respekterer den også i sin analyse og rapportering.

Flere af interviewpersonerne fortæller, at de i slutningen af et interview slukker båndoptageren, hvilket fremstår som en slags symbolsk manifestation af den fortrolighed, der er skabt – respondenterne vil fortælle noget til interviewer, men ikke til båndoptageren!

Loyalitet

Fortroligheden, der karakteriserer interviewsituationen, skal ses i sammenhæng med den loyalitet, den kvalitative forsker udviser over for sine respondenter. Forskerne henviser til, at de overholder bestemte implicite etiske retningslinier ved indsamling og fortolkning af data – eksempelvis er interviewets mål at få indsigt i respondenterens 'verden', *ikke* at afsløre den.

En interviewperson peger på, at indhentning af samtykke til arkivering ved et interviews begyndelse vil være illoyalt, fordi respondenterne i løbet af interviewet vil kunne komme med uoverlagte udtalelser. Samme forsker frygter desuden, at tanken om, at data skal arkiveres, vil påvirke hende i interviewsituationen, så hendes spørgsmål vil bevæge sig på et overfladisk plan for mere eller mindre bevidst at hindre respondenterne i at komme med afsløringer.

Gensidighed

Betegnelsen gensidighed henviser til, at dataindsamlingen er foregået som interaktion mellem to parter, der i fællesskab har formet en samtale. Aflevering af et interviewresumé til de respektive respondenter synes på en gang at være en måde, hvorpå forskeren forsikrer sig om, at den enkelte respondent står ved sit bidrag og en måde at foretage en symbolsk afkobling af respondenterne som personer og dermed i højere grad kunne betragte udskrifter som 'rene' data. Måden, den kvalitative forsker håndterer gensidigheden i interviewsituationen, er således at foretage en formalisering af denne, så en del af det ansvar, forskere oplever at have over for respondenter, kan fralægges, uden at et etisk ansvar krænkes. Hermed bliver gensidighed mindre end fortrolighed og loyalitet en barriere for arkivering, idet forskeren allerede under sin undersøgelse søger at nå til en afklaring af dette.

Det går igen i interviewene, at forskerne oplever en unik relation til deres respektive respondenter. Denne relation betyder, at de vil strække sig langt for at leve op til det ansvar, de føler over for respondenterne. Denne etiske ansvarlighed kan være i konflikt med arkivering, og det er derfor en forudsætning, at arkivering kan forenes med forskerens etiske ambition om at respektere fortrolighed, loyalitet og gensidighed.

Set fra DDA's synspunkt vil det være fordelagtigt, at forskerne, i det omfang det er muligt, afklarer disse forhold vha. formalisering, som det eksempelvis fremgik i forhold til gensidighed. Endelig kan man forestille sig, at arkivering af data på længere sigt kan blive en måde at kommunikere til interessenter, at undersøgelsen er gennemført i henhold til etiske 'spilleregler'.

Det praktiske...

Det kvalitative interview vil have en løs struktur, hvilket betyder, at interviews vil forløbe meget forskelligt, hvorfor forskeren vil være mere tilfreds med ét end med et andet. Et interview, hvor forskeren eksempelvis kan se, at der er begået interviewtekniske fejl, vil mindre sandsynligt blive arkiveret. Imidlertid behøver et 'dårligt' interview ikke at være ensbetydende med lav informationsværdi, og det er derfor meget væsentligt, at forfængelighed ikke kommer til at indvirke som barriere for arkivering.

Det kan synes ligetil for forskere at overgive data til DDA i form af en kopi af de tekstfiler, der er anvendt til analyse. Denne fil vil imidlertid typisk have visse større og mindre skavanker. Teksten vil ofte være uden korrektur, så den vil indeholde både stave-, slå- og grammatikfejl. Det kan tillige være tilfældet, at data rummer oplysninger af så personlig karakter, at forskeren ikke mener, de bør indgå som data. Endelig vil et datamateriale sjældent bestå af interviewudskrifter alene, idet andre teknikker ofte er anvendt som supplement. Eksempelvis kan det være tilfældet, at forskerens noter i forbindelse med interviews er i håndskreven form. Skal disse indgå i datamaterialet, kræves det, at de gøres maskinlæsbare. Arbejdet, inden aflevering af data kan finde sted, kan derfor være betydeligt.

Ligesom etiske aspekter kan praktiske aspekter stiller sig i vejen for arkiveringstiltaget. Formodentligt vil disse dog i højere grad end de etisk betingede kunne nedbrydes af argumentation og praktisk hjælp.

Anvendelse af arkivdata

En iagttagelse ved interviewene var, at forskerne intuitivt sammenholdt arkiveringstiltaget med formidlingsopgaven. En forklaring er, at man ser grunden til arkivering som den, at andre får adgang til materialet, som det kendes fra biblioteker. Blev tiltaget adskilt i to, kunne arkivering i sig selv opfattes som mindre problematisk end andres eventuelle adgang til data.

Andres anvendelse

Mht. andres anvendelse af datamaterialer, er de interviewede forskere først og fremmest bekymrede for, om de vil kunne billige den måde, andre vælger at anvende data på. Forskerne deler en forventning om, at de, i kraft af at de selv har foretaget interviewet, vil have den 'rigtige' udlægning af dets indhold. Hertil kommer, at de synes at forvente, at andre forskeres anderledes udlægnings af data vil være fejlagtige.

Interviewene peger således på en opfattelse af, at kun primærforskeren vil finde 'sandhed' i data. Dette synes at afspejle et noget reduktionistisk syn på, hvilke 'sandheder' data rummer, der ligger langt fra den subjektive virkelighedsforståelse, som brugen af kvalitativ metode sættes i forbindelse med.

Egen anvendelse

Forskerne delte den holdning til anvendelse af arkivdata, at disse ikke kan anvendes som primærdata, dvs. data der direkte tolkes på, men kun kan anvendes som baggrundsinformation/supplerende information. En interviewperson begrundet her:

"Man skal gøre sig klart, at på basis af det, den repræsentation af data, jeg har liggende foran mig, så er der noget, jeg kan sige og noget, jeg ikke kan sige. Jeg kan ikke komme med noget konkret. Du kan godt få noget ud af et interview, som en anden har lavet [...] også noget, der er valid."
(ph.d.studerende, marketing)

Denne forventning om egen anvendelsesmåde stiller sig i kontrast til den anvendelsesmåde, der forventes af andre, der som tidligere nævnt forventes at ville anvende data uforbeholdent.

I interviewerne nævnes flere konkrete motiver til at anvende arkivdata - følgende træder frem:

- ❑ Overblik over data indsamlet om et bestemt emne

- ❑ Afsøge undersøgelsesfelt, hvor litteratur er begrænset
- ❑ Indsigt i begreb/problemstilling berørt i data, evt. kun perifert
- ❑ Sammenligning mellem egne og andres data vedr. samme emne
- ❑ Samarbejder hvor man via arkivet har adgang til gruppens data

Motiverne fremstår overvejende som tilføjelser til anvendelse af arkivdata som primære data i et forskningsprojekt, som blev omtalt i artiklens begyndelse. De interviewede forskere forventer ikke at kunne anvende andres data på lige fod med egne data. Det synes bemærkelsesværdigt, at ingen forventer at kunne genfinde den forsigtige tilgang til anvendelse af arkivdata, som de selv mener at ville have hos andre forskere!

Blandt de interviewede forskere var der generelt en opfattelse af, at data bør gemmes og en positiv indstilling til anvendelse af arkivdata. Men som det er fremgået også forbehold, og et initiativ til en central arkiverings- og formidlingsinstitution må naturligvis søge at imødekomme dette.

Implikationer

På baggrund af interviewundersøgelser skal der peges på følgende anbefalinger og tiltag for at imødekomme forskernes ønsker og behov:

Tidlig indlevering. Det vil være væsentligt, at forskerne afleverer datamaterialer så tidligt som muligt i undersøgelsesprocessen. Ud over at det vil mindske risikoen for, at data 'bliver væk', kan tidligere indlevering medvirke til, at spørgsmålet om arkivering er latent under undersøgelsesprocessen – uden at det hæmmer forskeren i sit arbejde. Hvis arkiveringsproblematikken først rejses ved afslutningen af et projekt, kan dette skabe modvilje og vil ikke mindst øge det praktiske arbejde, der vil gå forud for arkivering.

Bistand. DDA må være parat til at stille hjælp til rådighed for forskere, der ønsker at arkivere data. Hjælpen kan eksempelvis bestå i tilbud om en skabelon til et brev til respondenterne, der begrundes, hvorfor dataarkivering foretages og beder om deres accept hertil i forbindelse med det specifikke projekt. Hjælp til at gøre data klar til arkivering er et andet eksempel, der kunne gennemføres blandt andet ved at tilbyde mulighed for scanning af materiale til forskere, der har dele af data i håndskreven form.

Vejledning. Som en del af et vejledningsmateriale bør DDA udarbejde et dokument, der præciserer, hvordan det faktum, at data (muligvis) skal arkiveres, bør indvirke i løbet af undersø-

gelsesprocessen. Dokumentet skal indeholde alt fra gode råd i forhold til håndtering af etiske aspekter til anbefalinger til noteteknikker og principper for transskribering.

Opbevaring af studier som cases. På grund af det kvalitative studies beskaffenhed må det anbefales, at 'produktet', der er udarbejdet på baggrund af data, arkiveres elektronisk sammen med data og dokumentation. Dette vil give primærforskeren en yderligere mulighed for at komme til orde, og det arkiverede kvalitative studie får karakter af en *case*, idet al information, der findes edb-læsbart, arkiveres og evt. stilles til rådighed sammen.

Skel mellem adgang og anvendelse. Som det fremgår, er forskerne bekymrede for, hvorvidt deres data, hvis de tillod andre forskere at anvende dem, ville blive mistolket. Samtidig mener de ikke selv at ville anvende arkivdata som primær datakilde, men udelukkende som baggrundsinformation. En kobling af disse undersøgelsesresultater peger på, at der indføres en praksis, hvorved der skelnes skarpt mellem *adgang til at læse data* og *adgang til at publicere på baggrund af analyse af data*. Ved at introducere denne skelnen kunne man muligvis hindre en del forskere i at begrænse adgang af frygt for, at andre forskere publicerer på baggrund af data.

Privilegeret adgang. Som foreslået af en af de interviewede forskere kunne man tilbyde, at materialer arkiveres med privilegeret adgang forstået på den måde, at primærforskeren opgiver navne på bestemte kolleger, der skal have adgang til at læse i og/eller bruge materialet. Ved at tilbyde privilegeret adgang til data kunne adgangspolitik for kvalitative arkivdata modsvare den form for dataudveksling, der uformelt finder sted p.t. mellem forskere, der deler interessefelt.

Personligt kendskab. Da en væsentlig del af modviljen mod arkivering er følelsesbetinget, må en ambition fra DDA's side være at sikre en nær kontakt til forskningsmiljøer, så en del af de forskere, der udgør potentielle donorer til arkivet, har personligt kendskab til medarbejdere i DDA. På den måde sikres det, at forskerne ikke oplever det at indlevere data, som at de skal sætte deres lid til en myndighed, men til personer, de har mødt.

Dialog. Der må opretholdes en løbende dialog mellem DDA og det kvalitative forskningsmiljø, så parterne hele tiden har en idé om, hvor de hver især befinder sig i forhold til udvikling og implementering af arkiveringspraksis. En sådan dialog vil bidrage til, at forskerne vil opfatte DDA som en åben institution med interesse i et tæt samarbejde.

Afslutning

Når vi fra DDA's side ønsker at inddrage kvalitative data – p.t. begrænset til interviewtransskriptioner og maskinlæsbart baggrundsmateriale til disse – er det fordi vi ønsker, at DDA skal være et dataarkiv, der dækker hele det empiriske forskningsfelt for samfundsbeskrivende emner.

Initiativets succes vil i høj grad afhænge af, at der hos forskerne skabes motivation til at overkomme en forventet skepsis. Det er derfor afgørende, at forskerne betragter arkivering og formidling af data som en aktivitet, der er til fordel for dem i deres forskningsarbejde og ikke som noget, der er pålagt dem af en offentlig myndighed (som det rent faktisk er, såfremt et forskningsprojekt er støttet SSF). Arkivering og formidling af kvalitative data skal, som det gælder for den kvantitative forsker, være en service for den enkelte forsker, der har indsamlet sit empiriske forskningsmateriale eller ønsker adgang til andres.

Litteratur

- Alvesson, M. & Skjöldberg, K. (2000): *Reflexive Methodology – New Vistas for Qualitative Research*, London, UK: SAGE Publications Inc.
- Burgess, R. G. (1984): *In the Field – An Introduction to Field Research*, London, UK: Routledge.
- Hakim, C. (1982): *Secondary Analysis in Social Research*, London, UK: George Allen & Unwin (Publishers) Ltd.
- Kvale, S. (1994): *InterView – En introduktion til det kvalitative forskningsinterview*, København, DK: Hans Reitzels Forlag a/s
- McCracken, D. G. (1988): *The Long Interview*, Qualitative Research Methods vol. 13, Newbury Park, CA: SAGE Publications Inc.
- Rasmussen, K. B. (2000): *Datadokumentation – Metadata for samfundsvidenskabelige undersøgelser*, Gylling, DK: Odense Universitetsforlag