

INTERVIEW

SAMFUNDSVIDENSKABERNE OG DE STORE DATA

Interview med Anders Blok, Rebecca Adler-Nissen og Morten Axel Pedersen

MY MADSEN OG KATINKA SCHYBERG

I dette interview diskuterer tre samfundsvidenskabeligere forskere fra CSS – lektor i sociologi Anders Blok, professor i statskundskab Rebecca Adler-Nissen og professor i antropologi Morten Axel Pedersen – deres erfaringer med og tanker om det at arbejde med Big Data, „computational social science“, heterogene datatyper og tværfagligt samarbejde i forbindelse med projekter på Copenhagen Center for Social Data Science, SODAS.¹ Og så er CSS' nye kandidatuddannelse i social datavidenskab, med start 2020, også oppe at vende.

Interviewet er foretaget af My Madsen. Det fremstår her i samtaleformat – om end i en redigeret version.

Hvordan arbejder I med data eller digitale metoder?

Rebecca: Jeg er jo kommet ind i det her felt lidt som mange andre samfundsforskere: fordi man møder et eller andet problem, som ikke længere giver mening, når man bruger de traditionelle metoder eller teorier, vi har. Eller et empirisk fænomen, der er for stort og vigtigt til, at vi bare kan overlade det til andre. Egentlig startede det med en for *Tidsskriftet Antropologis* læsere måske meget sjov anekdote: Jeg arbejdede i Udenrigsministeriet og skulle forberede noget til udenrigsministerens møde med en ambassadør. Jeg havde lavet alle de der traditionelle talepapirer, som jeg havde printet ud – de vigtigste samtalepunkter osv. Men så skete der bare det, at hun jo allerede havde snakket med ham – hun havde allerede sms'et med ham om, hvad de skulle snakke om på mødet. På den måde var hele statens normale måde at fungere på ligesom bare blevet kørt over, fordi de engang havde udvekslet private numre. Da jeg så kom tilbage på universitetet, tænkte jeg: Der foregår altså nogle ting i det daglige, som vi slet ikke har begreb om. Nu er det her måske lidt trivielt, men det handler også om store spørgsmål. Siden har jeg så søgt to store projekter: et Carlsbergfondsprojekt, der undersøger digital disinformation fra Øst mod Vest. Her kigger vi på det her pro-Kreml-dis-

information – hvem når det, og hvem spreder det og hvordan? Det andet projekt handler om at forstå, hvordan fortrolige forhandlinger og håndteringen af offentligheden egentlig hænger sammen i dag, med fokus på sociale medier i det man – måske lidt søgt – kan kalde den digitale tidsalder, for eksempel i forbindelse med sådan noget som Brexit.

Jeg bliver lige lidt fanget af den case, du taler om – for mig er det ikke helt entydigt, hvordan sådanne personlige samtaler kan undersøges med Big Data-materiale? Hvad er sammenhængen der?

Rebecca: Hvis du er interesseret i at forstå, hvordan for eksempel Brexit forhandles både offline og online, bliver du nødt til både at have fornemmelsen af denne her Brexit-boble – altså alle de tusindvis af mennesker, der hver dag engagerer sig på Twitter og Facebook og Instagram i forskellige netværk og kommenterer eller forsøger at påvirke eller på anden måde være en del af processen – og så denne her „backstage“, som må undersøges gennem deltagerobservationer og sådan nogle lidt mere antropologisk inspirerede metoder, hvor vi for eksempel spørger eller følger og kigger på en diplomat i løbet af et par dage. Hvor tit tager han eller hun mobilen frem? Med hvilket formål osv.? Det bliver på den måde en kombination af kvalitative og kvantitative metoder, som jeg synes er nødvendigt for dette her spørgsmål, simpelthen!

Så for dig er det kombinationen af de her metoder, der kan noget nyt?

Rebecca: Ja, det tror jeg, der er rigtig meget potentiale i – også fordi mange af os på SODAS kommer med nogle ret specielle samfundsvidenskabelige eller nogle gange næsten humanistiske metoder til at forstå mening med, og hvordan folk oplever alt muligt i deres hverdag. Og når vi så møder det, der også er deres hverdag, nemlig millioner af tweets, som former den måde, man forstår virkeligheden på, tror jeg, vi kan begynde at stille nogle spændende spørgsmål.

Morten: For at give lidt baggrund for SODAS kan jeg jo lige fortælle, hvordan det startede: For efterhånden en hel del år siden kom jeg mere eller mindre tilfældigt til at snakke med David Dreyer ovre fra økonomi og Søren Kyllingsbæk, som dengang var lektor på psykologi og nu er institutleder på psykologi, om at lave et eller andet. Vores dekan udloddede på det tidspunkt nogle „seed money“ til at lave noget på tværs af institutterne. Så spurgte Søren, David og jeg, om vi ikke skulle mødes med en fyr, han kendte ude fra DTU, Sune Lehmann, fordi han var interesseret i at bruge nogle metoder, han havde udviklet til at lave analyser af netværker sammen med nogle samfundsvidenskabelige forskere. Vi blev enige om at søge pengene sammen. Det var dog lidt skævvredet i forhold til det kvalitative – de andre var meget kvantitative – så jeg tænkte: Hvorfor ikke få alle

fagene med? Så jeg spurgte Anders Blok, om han også ville være med. Og det var så starten på det. Så fik vi senere landet næsten 20 millioner kroner fra nogle interne midler på KU, som hed Sense of Excellence 2016-midler, til det projekt, der hed „Social Fabric“. Det gik ud på at uddele mobiltelefoner til omkring 1.000 førsteårsstuderende på DTU og så undersøge den population af studerende i rent ud sagt hoved og røv. Vi kaldte det „deep data“. Og det var selvfølgelig også derfor, at jeg fik med i projektforslaget, at der skulle være en etnograf, der lavede feltarbejde blandt de studerende, og som selv fik en telefon. Og det var jo så dig, My. Men der er jo også folkesundhedsfolk, økonomer, psykologer, sociologer, STS-forskere, folk fra filosofi, fra Niels Bohr Institutet og selvfølgelig fra DTU med i projektet.

Du fortæller, at det indledningsvis kom i gang på grund af netværksstudierne ude fra DTU – men hvad forestillede I jer, at det andet datamateriale skulle supplere med? Kan man tale om det som en skalering af typen af datamateriale, der var med i projektet?

Morten: Når man arbejder med sådan nogle såkaldte ustrukturerede metadata – altså de typer data, der blev logget i projektets mobiltelefoner, der fortæller, hvornår folk var online og offline, hvornår de ringede til hinanden, var på Facebook og andre sociale medier osv. – så har man at gøre med nogle typer data, som jo ikke er resultater af en hypotese, man på forhånd har besluttet, at man skal gå ud og teste a la for eksempel: Hvor mange mennesker i Odder stemmer på Dansk Folkeparti, når der er økonomisk nedtur? Det her var meget mere sådan „blue sky“-agtigt: Vi har en masse nye typer data, og der er ikke rigtig nogen, der ved præcist, hvad man kan bruge dem til, men det var helt klart, at de kunne bruges til et eller andet. Hvad sker der, hvis vi sætter dem sammen med hinanden og så oveni også bringer en masse fagligheder sammen? Vi laver for eksempel spørgeskemaer, som har den der mere traditionelle kvantitative samfundsvidenskabelige approach, og oveni lægger vi så de her etnografiske tilgange og teorier. Vi sætter alt dette her sammen, og så ser vi, hvad der kommer ud af det. Det var idéen. Og derfor var det mig meget magtpåliggende, at den person, der skulle ansættes i den etnografiske ph.d.-stilling – som endte med at være dig, My – ikke nødvendigvis skulle være en person, der på forhånd havde lavet en masse digital etnografi. Det væsentligste var, at det var en, der havde mod på og evner til at gå ud og lave sådant et fuldstændigt traditionelt antropologisk feltarbejde, hvor man virkelig nedsænker sig i en population af mennesker i et år eller længere, og så på en eller anden måde tage de data og se, om de kunne kombineres med alle de andre. Og jeg må indrømme, jeg havde ingen anelse om, hvorvidt det kunne lade sig gøre. Eller hvordan man skulle gøre.

Anders: Noget af det, som jeg synes er meget vigtigt at tale om, er, at mindst lige så meget som det handler om at kombinere samfundsvidenskab og datavidenskab, handler det også om at forsøge at gå på tværs af samfundsvidenskaberne. Og den bevægelse er mindst lige så udfordrende, om ikke mere udfordrende, som den første bevægelse.

Jeg kom ind i projektet som „sociologen“ og modsat, hvad man måske kan forestille sig, når man er et antropologisk publikum – fordi jeg har sådan en forestilling om, at antropologer konstruerer sociologer som sådan nogen, der laver det samme som dem selv, bare med tal – arbejder jeg ikke nødvendigvis så meget med tal. Så jeg kom ligesom ind som sådan en underlig kvalitativ sociolog, som ikke desto mindre jo har været med til at udforske mulighederne for at lave sociale netværksanalyser, i bred forstand, på de her typer af data.

Det andet ben, jeg måske vil nævne, som jeg også kom ind med og allerede var i gang med, kommer ud af min interesse for det, der hedder videnskabs- og teknologistudier eller STS. Man kan sige, at digitalisering og digitale teknologier laver om på sociale relationer, og så har de digitale teknologier også den egenskab, at de rematerialiserer sociale interaktioner i stort og småt, fordi de bliver opfanget og sporet i alskens platforme, på apps, i købsdata og på alle mulige typer af måder. De her digitale spor er noget, jeg har undervist i og udforsket i forhold til at forsøge at sige noget om det offentlige engagement i det, vi kalder videnskabelige kontroverser. Det gør vi sammen med folk som Bruno Latour og noget af det der STS-slæng. Og så er det også noget, jeg bruger i min egen forskning for tiden, hvor vi bruger Facebook til at forsøge at sige noget om civile gruppers engagement i bynatur. Vi er blandt andet nysgerrige på at se, hvor langt vi kan komme, så at sige, „ned“ i det nære, interaktionelle i de her grupper alene ved hjælp af de her digitale spor. Vi har bevidst valgt at lade være med at sætte dem sammen med etnografi eller interviews eller surveys. Det giver en sjov mulighed for at udforske spørgsmålet om, hvad det egentlig er for nogle kapaciteter, de her data har, når de både er store og også går helt tæt på personlige og interaktionelle dynamikker i for eksempel en byhave et sted i København, der er ved at finde ud af, hvordan de skal forhandle med kommunen, eller hvordan den og den persons ageren til sidste weekends hyggemøde var helt upassende og sådan nogle typer af ting.

Da du talte om SODAS, sagde du, at du syntes, det var vigtigt, at alle samfundsvidenskaberne var repræsenteret, og at det at arbejde på tværs af samfundsvidenskaberne var en vigtig del af projektet. Hvorfor det? Hvorfor er det ikke nok med forskelligt datamateriale, hvorfor skal der også være forskellige samfundsvidenskaber med?

Anders: Det er vigtigt for forståelsen af det, vi laver: Det former enormt meget det rum, vi tre og de andre samarbejdspartnere indgår i, og den forestilling, vi har om den uddannelse, vi også er i gang med at lave [social datavidenskab på CSS, red.]. I bredere forstand er det jo vigtigt, fordi man bliver opmærksom på, at vi i samfundsvidenskaberne trækker rundt på helt forskellige analytiske og metodiske traditioner. Vi mener for eksempel forskellige ting med de samme ord. Jeg satte mig på et tidspunkt ned for at tænke over, hvad de forskellige involverede discipliner i Social Fabric egentlig mente med et „netværk“. Det var ret interessant, for vi mente faktisk systematisk forskellige ting. Man skal huske, at „netværk“ er et interessant mødested for en hel masse discipliner. Det er jo et sted, der har en lang antropologisk tradition begyndende med, at man var ude i landsbyer og spørge folk, hvem de mødtes med, og hvem de havde forskellige typer udvekslinger med. Det bestod i, at man simpelthen satte sig ned og tog et stykke papir, og så tegnede man streger imellem. Hele den der netværksfigur er jo vedblivende en inspirationskilde for en masse antropologisk arbejde. Samtidig har den udviklet nogle matematiske dimensioner i en hel anden retning indenfor andre fag, blandt andet kvantitativ sociologi og økonomi. Og den har en historik i fysik, hvor man kan sige, at for sådan en som Sune Lehmann er det ligeegyldigt, om han studerer proteiner i en krop eller mennesker på et universitet.

Rebecca: Det slår mig, at jeg i stigende grad tænker på SODAS som et sted, der også kan føde tilbage til de gamle discipliner. Morten og Anders var begge to inde på, hvorfor de hyrede dig, My: Det var, fordi du var en dygtig antropolog i din disciplin. Det er ikke sådan, at vi hiver disciplinerne væk. Det, der lige for tiden foregår, er en transformation af, hvordan verden bliver skabt – hvordan vi laver byhaver, eller hvordan vi forhandler en fredsaftale i Ukraine – da alt er digitaliseret på alle mulige måder. Det betyder, at vi er nødt til at gå tilbage og revidere nogle af de begrebsrammer, vi allesammen render rundt med – Berger og Luckmann, Goffman, Bourdieu eller „rational man“ for den sags skyld, idet der altså foregår nogle ændringer i forholdet mellem stat og marked, mellem individ og kollektivitet, som er så grundlæggende, at vi nok skal gentænke idéer om social interaktion, som Anders er inde på. I mit felt, politologi, er der en tendens til at blive så begejstret over de her nye sociale data, at vi ligesom antager, at de i sig selv er meningsfulde i folks verden, og vi antager, hvad det er for en mening, de har, og at det må være den samme for hvert individ. Så et „like“ er et „like“ for eksempel, og så tæller vi dem. Det er det, jeg opponerer imod, blandt andet inspireret af samtaler med Morten og Anders og andre fra jeres discipliner. På den måde er der simpelthen grundlæggende noget at lære af hinanden i forhold til, hvordan vi stiller spørgsmål, og selvfølgelig i forhold til de begreber, vi bruger. Forhåbentlig vil det også føde tilbage til vores moderdiscipliner.

Morten: For mig at se er det et virkelig vigtigt spørgsmål, og det føder jo ind i, hvad SODAS egentlig er for noget. SODAS udspringer af Social Fabric-projektet. Det blev besluttet hos ledelsen på fakultetets side, at der, på grund af alt det, Rebecca lige har omtalt – de her ændringer i verden, der tilsyneladende har en irreversibel karakter, såsom digitalisering, datadrevet økonomi, politik, alle mulige ting – at der skulle sættes strategisk på det. Så der blev afsat nogle midler til at gøre dét, der egentlig var et forskningsprojekt, til et center på fakultetet. Det var så på det tidspunkt, at vi var så heldige at få Rebecca med i vores gruppe og en hel gruppe af andre folk også fra statskundskab, som var i gang med at interessere sig for nogle af de samme fænomener. Det der gør, at vi ligesom træder i tredje gear nu – først var der forskningsprojektet Social Fabric, så var der SODAS – det er den uddannelse, der er blevet godkendt af ministeriet, og som starter næste år og hedder Social Data Science. Og jeg kunne simpelthen mærke den dag, godkendelsen kom i hus, da skete der et eller andet inde i mig. Det, der er utroligt vigtigt for mig, som både udfordrer mig, men også er ufatteligt spændende, er, at der jo er tale om et helt nyt fag. Jeg kan godt lide idéen om et nyt fag, for et nyt fag har både at gøre med undervisning og forskning og med metoder – det har ligesom noget at gøre med det hele. Altså en ny faglighed. Og hvordan skaber man den? Det tror jeg, at vi har nogle potentialer for, fordi vi i hvert fald har det første, det kræver, og det er nogle dygtige forskere indenfor deres discipliner. Men vi har også nogle forskere, der stoler nok på hinanden og respekterer hinanden nok til, at de tør risikere sig selv og deres faglighed, når de sidder og taler med hinanden. Uden sådan en stemning kan man ingen vegne komme, når man skal lave et nyt fag. For ellers sidder vi bare hele tiden og passer på vores fag. Vi er i gang med at udvikle et sprog sammen. For som Anders siger, kan det godt være, at vi tror, vi taler om de samme ting, når vi for eksempel siger „netværk“, men det gør vi jo ikke. Så vi bliver nødt til at etablere et vokabularium, et begrebsapparat for, hvad vi egentlig mener, når vi siger et „netværk“ eller „kvalitativ“, „etnografisk“ eller „behaviour“. Det bliver superspændende, men det bliver selvfølgelig også en stor udfordring.

Jeg tænker på, om man kan sige, at der er to sider af det: Der er både et samfundsvidenskabeligt politisk projekt om, at det her er noget nyt, som man skal være med på for at kunne være med til at bestemme, hvordan de her ting skal defineres, og så er der også det, at data simpelthen fylder mere i folks liv. Og det leder mig til spørgsmålet om, hvad Big Data er i forhold til nogle af de andre typer data, vi også har arbejdet med, for eksempel registerdata?

Rebecca: Jeg kan fortælle, at jeg har en kollega, der har et skilt på sin dør, hvor der står: „I was into Data before it became Big“, og den, synes jeg, er meget god.

Big Data er jo på en eller anden måde blevet en „placeholder“ for, at store data er blevet lettere tilgængelige. Den nye type data er lettere tilgængeligt og tilbyder os – og Amazon – nogle muligheder, som vi slet ikke har haft tidligere for at forstå, hvordan folk agerer. Og det er jo det, samfundsvidenskaben gerne vil vide. Så på den måde ville det være mærkeligt, hvis vi ikke interesserede os for alle de nye typer digitale spor. Men en ting, vi måske også kan komme ind på, som jeg synes, er vigtigt, det er etik i forhold til dette her spørgsmål. For noget af det, vi har mødtes om, har været spørgsmålet om, hvorvidt de studerende skal kunne EU's nye persondataforordning. Og skal vi? Hvis man spørger sådan, må man svare, at det skal de og vi selvfølgelig kunne – men hvad vil det sige at „kunne“ den? Og pludselig rejser det sådan nogle lidt mere grundlæggende spørgsmål: Ville det for eksempel være nok, hvis de bare tog sådan et crash course og kunne den? Nej, det ville det jo ikke, for samfundsvidenskab er jo knap nok jura. Så vi er ude i også at stille spørgsmål som for eksempel, hvordan algoritmer former vores politiske holdninger – et typisk politologisk spørgsmål – eller kulturelle valg, for en sociolog. Så vi har en drøm om et fag, der tænker politik og etik på virkelig mange niveauer, så for eksempel spørgsmålet om persondataforordningen også bliver et spørgsmål om at følge folk, der implementerer den, og se, hvor mange grå hår i hovedet det egentlig giver folk. Jeg tænker lidt på den fase, vi er i nu, som sådan et „wild west“, og på et eller andet tidspunkt kommer samfundet for at generobre den kolonisering, der er sket af de her store mastodonter, som forsøger at kolonisere vores egne data.

Morten: I forhold til hvad Big Data er, vil jeg sige, at fundamentalt set er der jo ikke noget nyt ved Big Data som sådan. Du har haft store data i astronomien, i naturvidenskaben, i medicinen og for den sags skyld også i survey data i samfundsvidenskaben. Men det, der gør det nyt, er måske det, man kalder „social“ Big Data. Og med det menes der i særdeleshed alle de her typer af spor, der genereres ved vore digitale gøren og laden. Det er dem, der markerer en fundamental forskel fra for bare 10 år siden. Og så kan måske sige: Er det ikke bare hype med det der Big Data? Der synes jeg, at svaret både er ja og nej. Det er selvfølgelig fuldstændig rigtigt, at der allerede er et „backlash“ i gang: Vi kan se, at Magrethe Vestager i EU-Kommissionen og danske politikere fra både venstre og højre er begyndt at stille kritiske spørgsmål. Men pointen er, at de største firmaer i verden er datadrevne firmaer. De lever af at give såkaldt gratis produkter til dig. Men som en eller anden har sagt, som jeg ikke lige kan huske, hvem er: „If the commodity is free, it means *you* are the commodity.“ Det er jo det, Facebook og Amazon og sådan nogle lever af. Deres algoritmer, der er baseret på metadata, er gode nok til at forudsige med stor nok sandsynlighed, hvorvidt du vil købe en bog eller stemme på et bestemt parti, til at de kan tjene en masse penge på det. Så bare det, at det er verdens største virksomheder, viser, at sociale Big Data er

enormt effektive. De her teknikker blev brugt til at influere det amerikanske valg – måske endda på afgørende vis! Vi må simpelthen forstå, at det her, det virker.

Anders: Jeg vil også sige, at vi grundlæggende har at gøre med de her digitale spor og mængden og typen og heterogeniteten af dem, når vi snakker om Big Data – indenfor samfundsvidenskaberne. Der kan man jo lægge til, at de vil tendere til at have nogle egenskaber, som adskiller sig fra, hvordan datasæt tidligere har set ud. Der er en række interessante ting ved dem, hvoraf man kan nævne tre: for eksempel granularitet i tid og rum, det vil sige tæthed af observationer i tid og rum. Sammenlignet med en surveyundersøgelse eller registermateriale vil man få observationer, som kan være lige så granulære i rum, som vores descartianske grid nu engang er i stand til at geare os til. Så det er nogle nye typer egenskaber, som vi jo allesammen skal lære at omgås. Fysikerne kan uden tvivl nogle ting med den type data, som er interessante også for samfundsvidenskaberne bredt set. Den anden ting, man kan sige om dem, er, at mange af de her datasæt vil omhandle det, man vil kalde „hele populationer“. Det er sådan lidt et snydeord, fordi selvfølgelig er der alle mulige huller i dem, men det er mere set i forhold til, hvad det er for en forestilling om repræsentativitet, der er indbygget i de her datasæt: De er ligesom taget ud af en faktisk population. Det kan være en population af DTU-studerende, det kan være en population af Facebook-brugere på bestemte sider, det kan være dem, der shopper i Coop osv. Og det er en forskel i forhold til idéen om et statistisk udsnit, som jo har formet samfundsvidenskaberne i hundrede år og leveret forestillingen om, hvad det vil sige at repræsentere den sociale verden i kvantitativ forstand. Så der sker igen noget ret radikalt, som udfordrer alle samfundsvidenskaberne i forhold til at tænke repræsentativitet på en ny måde. Og endelig den tredje ting, som nogen har påpeget ved at kaste et STS-blik på samfundsvidenskaberne, kommer der et nyt slags subjekt frem her. Altså en ny forestilling om den sociale agent. Man kan sige, at det kvalitative interview og måske feltarbejdet har haft en tendens til at producere, hvad man kunne kalde „det reflekterende subjekt“, som forholder sig selvrefleksivt. Surveys har produceret „the opinionated borger“ – borgeren eller personen med holdningerne, der mener noget om det ene og det andet. De her nye digitale data, kunne man måske sige, har reference til sådan noget som et „transaktionelt“ subjekt. Det er det handlende subjekt, der hele tiden rager nogle relationer med ind i sine handlinger. Data bliver produceret gennem nogle tekniske infrastrukturer, og det vil sige, at handlingen også er formet af de teknologiske rammer og af den interaktionelle kontekst, hvori disse spor bliver lagt. Så du får på en eller anden måde et iboende, relationelt distribueret, handlende subjekt. Det er enormt spændende og potentielt noget, der kan have meget vidtrækkende konsekvenser, ligesom de andre metoderevolutioner i samfundsvidenskaberne har haft det.

Men måske også konsekvenser udenfor samfundsvidenskaberne? Hvilken indflydelse kan man forestille sig, at det har på almene borgere, at man for eksempel pludselig kan gå tilbage og kigge på sine handlinger? Gør det én til et mere handlende subjekt?

Rebecca: Ja, men også selvobserverende: Man spejler sig også på en ny måde og ser sig selv som noget, der både er virtuelt og ikke virtuelt. Nu har jeg børn i denne her alder, vuggestuer og børnehaver. Det er helt vildt, som der lægges digitale spor her, men det er jo digitale spor, man selv kan observere, mens de bliver til. Så da jeg skulle hente min datter Naomi i vuggestue i går for eksempel, sidder pædagogen og tager et billede af hende og siger: „Nu lægger jeg det op og ‘tagger’ dig, så du kan se, hvor sød hun ser ud.“ Og så snakker han om sin egen datter, der også godt kan lide de der digitale spor. Så de bliver altså en del af verden, selv i en situation, som vi måske tænker, ikke er digitaliseret som sådan.

Morten: Må jeg tilføje til det? Jeg tror, der er noget relativt distinkt ved Danmark i denne her forbindelse. Danmark er jo på mange parametre verdens mest digitaliserede land. Og jeg vil også kalde det måske verdens mest digitalt glade land. Min hypotese om, hvorfor det forholder sig sådan i Danmark, er, at det ikke bare skyldes den gode infrastruktur, herunder digital infrastruktur, der gør, at firmaer har testet deres produkter her, så der er sket en stor absorption af de her teknologier, men at grunden til, at Danmark er så digitaliseringsliger, også er, at vi uden at tænke over det simpelthen har taget den relation, vi har til staten, og så forskudt den ind i det digitale. Vi tror, at det digitale er en relativt rar og pålidelig infrastruktur på samme måde, som vi opfatter staten. Vi har jo en hel unik tillidsrelation til staten. Det samme gælder vores relation til det digitale. Og den udgør en meget interessant prisme, hvorigennem vi kan undersøge vores relation til staten og til hinanden og til idéer om tillid og alle mulige andre ting.

Anders: Og her kan man jo sige, at det samme gælder vores, samfundsvidenskabernes, relation til de her data, som jo selvfølgelig formes enormt meget af det forhold, at de blandt andet skabes af og tilhører monopolkapitalistiske virksomheder. Det giver alle mulige spørgsmål om adgang og distribuering af viden. Og potentielt selvfølgelig også tab af autoritet for sådan nogen som os på et universitet i et provinsland som Danmark, som kan have svært ved at hamle op med de store Silicon Valley-algoritmer osv. Så det former også vores egen relation til data.

Rebecca: To observationer i forhold til det, Anders sagde om de der måder, hvorpå data præsenterer sig på i forhold til granularitet. Der er jo også et metodisk spørgsmål i det, fordi registerforskning kom med sine „biases“, og det gør de her nye data også, og jeg oplever, at vi ligesom føler os nødsaget til at acceptere, at ja, det er rå data, og de er store, men de er jo i den grad også udvalgt. Det

er svært at komme uden om, at vi er nødt til at bero os på den struktur, som de her platforme nu har. Så en platformforståelse er i hvert fald helt nødvendig, ligegyldigt hvilken videnskab man kommer fra. Der bliver publiceret så meget lige nu, som ikke engang nævner med et ord, at de vigtigste metodiske beslutninger jo ikke er truffet af os, men af Facebook, Twitter osv.

Jeg vil gerne spørge mere ind til, hvordan det er at arbejde med de her store datasæt: Hvad kan arbejdet med Big Data og andre digitale metoder i forhold til de klassiske fag og analyser?

Rebecca: Ja, altså nu har jeg interesseret mig meget for internationalt diplomati og forhandlinger – helt grundlæggende sådan noget med krig og fred. Og i det øjeblik krig og fred medieres i så høj grad, som det gør på for eksempel sociale medier, rejser det helt vildt mange spørgsmål om alle de antagelser, vi i forskningen har gjort om, hvorfor og hvordan vi kan få fred. For eksempel at fortrolighed, langsommelighed, tillid og sådan nogle ting er nødvendige for en fredsftale. Nu kan vi se, at den bliver forhandlet i højere grad, eller i hvert fald i lige så høj grad, på Twitter. Og når vi spørger dem, der forhandler for eksempel Ukraine-aftalen, hvordan de gjorde det, ja, så taler de om håndteringen af en „offentlighed“. Det er et begreb, som samfundsvidenskaberne tror, de har forstået efterhånden, „the public“, men som man bare må sige, er så meget oppe i luften nu. Det samme gælder gamle spørgsmål om nationalitet. Vi har jo denne her metodiske nationalisme på tværs af mange af samfundsvidenskaberne, hvor vi antager, at staten er en enhed, og der rejser de sociale medier og forhandlingen af nationalitet helt nye spørgsmål. Jeg kan prøve at give et eksempel på, hvor vigtigt det er: Måden, som britiske medier rapporterer om Brexit på, og som jo direkte former, hvordan briterne danner holdninger og dermed de „polls“, der kan laves, er jo i høj grad baseret på tweets af for eksempel Donald Tusk. Når han så tweeter „There’s a special place in hell for those who didn’t have a plan for Brexit“, bliver det fortolket på en meget speciel måde i de britiske medier, hvor man ikke får hele hans tale med. Så selve den form, platformenes konkrete udformning muliggør – længden af tekst og mængden af billeder for eksempel – former løbende direkte måden, hvorpå en så vigtig beslutning – måske den allervigtigste beslutning for briterne de seneste 50 år – bliver truffet.

Morten: I forhold til antropologien og Big Data er der for mig at se tre muligheder, hvormed antropologer kunne, og til en vis grad har gjort det, relatere sig til Big Data-problematikker og muligheder: Den ene er det, jeg vil kalde antropologien om Big Data. Det er der faktisk en hel del af allerede – og rigtig gode studier. For eksempel har vi jo et projekt her på instituttet, som handler om biometriske grænseverdener, hvor man undersøger, hvordan, på ofte meget skræm-

mende og forandrende i verden måder, de her teknologier bruges i forbindelse med transnational migration. Og så er der en masse folk, der arbejder med det, jeg vil kalde „algorithmic governmentality“, altså, hvordan algoritmer instigerer en ny form for governmentality i sådan en foucauldiansk tradition; endnu en generation af sociale teknologier, hvor selvteknologien simpelthen bliver, at du kan se dine egne spor og derfor forsøger at optimere din egen handling i forhold til en eller anden subjektposition, som du forsøger at leve op til. Så er der også noget andet, som jeg selv og Anders har været med til at udvikle i forbindelse med Social Fabric-projektet, og som jeg vil kalde antropologi *med* Big Data. Det indebærer, at man på forskellige måder forsøger at kombinere kvalitative og kvantitative metoder, store data og etnografiske observationer, og simpelthen forsøger at svare mest adækvat, mest videnskabeligt på et givent spørgsmål. Men så er der en sidste form, som jeg ikke synes, findes endnu, og det er det, jeg kalder antropologi *som* store data. Det er simpelthen at gøre antropologien kvantitativ eller „computational“, så den selv bruger og genererer store data – altså siger noget med tal. [...] Lige nu er jeg mest interesseret i og mest begejstret over det, jeg kalder „computational structuralism“ – muligheden for at udleve Levi-Strauss’ drøm: en total interdisciplinær superstrukturalisme. Simpelthen lave det, jeg kalder „machine-antropologi“: lave maskiner, der kan udøve antropologisk forskning. Lidt ligesom det Anders sagde tidligere, at hvis der ikke er nogen på universiteterne, der sidder og forsøger at gøre dette her på en etisk og videnskabelig forsvarlig måde, skal firmaerne nok gøre det. Og jeg skulle hilse at sige, at det er firmaerne allerede i gang med at lave i Silicon Valley. Så vi bliver nødt til at gå ned i „backenden“ og lære at lave programmerne, lære at kode tingene på en antropologisk måde. Lære at lave fortolkninger af store datasæt, som er antropologiske fortolkninger.

Anders: Til forskel fra antropologi har politologi og sociologi allerede en kvantitativ gren. Så en version af svaret på, hvad „computational“ antropologi er, kunne være: kvantitativ sociologi. Jeg tror ikke, det er det, Morten forestiller sig, men på den måde stiller spørgsmålet sig jo interessant anderledes indenfor de forskellige fag. Derfor er jeg nok mere optaget af sådan nogle helt almene spørgsmål om, hvordan store, digitale dataformer kan omrokere det her univers af såkaldte kvalitative og såkaldte kvantitative metoder. Først og fremmest og i særdeleshed kan vi forhåbentlig indarbejde en opmærksomhed på de mange måder at kvantificere på. Altså, kvantitativ betyder jo sådan set bare, at man forsøger at tælle noget op, og noget af det, man finder ud af ved for eksempel at tale med fysikere, er, hvor mærkelige former dette har taget i samfundsvidenskaberne, og hvor meget anderledes det matematiske rum af muligheder kunne se ud. Det relancerer helt grundlæggende spørgsmål om, hvordan vi tæller, hvad vi tæller, og

hvordan vi aggregerer på kryds og tværs af tid og rum, indenfor en bevidsthed om, at det, vi netop kan nu, er at være meget tættere på den sociale interaktion i tid og rum. Det giver spørgsmålet om kvantificering nye dimensioner.

Kan man snakke om en særlig samfundsvidenskabelig tilgang til Big Data?

Rebecca: Jeg var virkelig slået af, hvor hungrende vores aftagerpanel var efter det samfundsvidenskabelige blik på det her Big Data. Og det er jo både efter sådan en overordnet linse, herunder kritik: gode gamle marxistisk inspirerede, eller i hvert fald samfundsvidenskabelige, kritiske blikke på data, som går ud på at sige, at det er kommercialisering eller kapitalisme i nye former osv. Det er der virkelig behov for – et samfundsbehov simpelthen. Og så kan man også sige noget om, hvor meget der er en våd drøm i det her for staten. Vores myndigheder hungrer efter den type kandidater, vi kan producere. Men det slår mig, hvor pauver forståelsen er af, at det er et socialt felt: Vi har „datavidenskab“, men grunden til, at vi insisterer på, at det er „social“, er, at det er mennesker, der genererer de her data. Og det er menneskelige relationer, der blandes sammen med algoritmer, der danner vores interaktioner. Så det er samspelet mellem dem, vi interesserer os for. Jeg tror, at de spørgsmål, vi kan stille, og på lidt længere sigt de svar, vi kan komme med, er sådan helt fuldstændigt afgørende for, at et land så gennemdigitaliseret som vores overhovedet kommer til at hænge sammen. For ellers er vi virkelig dårligt klædt på.

Anders: Det, synes jeg, er enormt vigtigt, det, du siger der. Man kan måske sige lidt af det samme ved i et øjeblik at tænke over, at vi har sat „social“ foran „datavidenskab“ i navnet på den nye uddannelse, og tænke over den dobbelte betydning, det har. For det er jo noget med, at „social“ både er inde i de der data, altså, det handler jo for pokker om det sociale, og det overhovedet at kunne kode dem og rense dem og fortolke dem vil kræve socialvidenskabelige færdigheder. Men „social“ handler også om data: Det er alt det, der former data, alle de etiske, juridiske, politiske, organisatoriske, kulturelle osv. vilkår og hensyn, der former, strukturerer og indkodes i data. Så „social“ ligger på begge sider af data. Og det her med at kunne bevæge sig kompetent ind og ud af datasættet, det, tror jeg nærmest, jeg vil sige, er definerende for vores forståelse af, hvad det er, vi skal lære folk i social datavidenskab-uddannelsen. Og selv skal kunne og blive bedre til.

Morten: Jeg er helt enig i det, der bliver sagt her. Man kan måske opsummere det med det, som SODAS havde lavet til et slogan: „Vi skal allesammen lære at kode, men vi skal også allesammen kunne afkode.“ Og denne her uddannelse i social datavidenskab har jo til formål, at man kan begge ting. De skal lære at kode og nogle basale programmeringsteknikker. Så skal de selvfølgelig

også lære noget etnografi, de skal lære nogle digitale metoder og nogle teorier om, hvad samfund, og politik og jura er i forhold til dem. Så skal de også lære diskussioner om forholdet mellem de forskellige typer samfundsvidenskabelig forskning. Hvad kan vi sige, og hvad kan vi ikke sige med den og den type data? Det er jo sådan noget, at samfundsvidenskabelige folk er rigtig gode til. Det er også det, de her nye kandidater skal gå ud og gøre. De skal, for det første, gå ud og arbejde med nogle folk, som har en baggrund i computervidenskab nede i det, man kalder „backenden“, hvor man laver algoritmerne. De skal selvfølgelig ikke kunne lave algoritmer lige så godt, men de skal kunne forstå det, så de kan stille relevante spørgsmål med en samfundsvidenskabelig baggrund, således at algoritmen bliver lavet på en mere, lad os kalde det, social og politisk bæredygtig måde. Så skal de være i stand til undervejs at forholde sig til, hvad vi kan bruge de her data til; hvilke databaser skal de indgå i, hvilke mennesker i en organisation, for eksempel i en kommune, skal tilgå dem og i hvilke konstellationer? Og så skal de selvfølgelig kunne formidle det. Det er i virkeligheden en fuldstændig klassisk samfundsvidenskabelig uddannelse i nogle færdigheder og nogle kritiske attituder, så man kan komme ud og begå sig i det private såvel som det offentlige, såvel forskningsmæssigt som uden for akademiet, som samfundsvidenskabelig forsker. Men at man også har viden, herunder teknisk viden, om koding og store data.

Og hvad med forskningsaspektet i det? Forestiller man sig, at den nye uddannelse også er en forskeruddannelse?

Morten: I forhold til det forskningsmæssige kan man jo sige, at det er det, vi allerede laver i SODAS – og SODAS er jo ikke præcis det samme som den nye uddannelse. Uddannelsen er en kandidatuddannelse, så forhåbentlig er der nogle kandidater derfra, der vil lave en ph.d., og jeg kunne da godt tænke mig, at nogle af de folk ville lave en ph.d. under for eksempel en SODAS-paraply. For det, der er interessant og faktisk ret nyskabende, er, at de ph.d.-studerende ikke bliver ph.d.er i et fag på det her fakultet, men i „social science“ – altså, de får en ph.d.-grad fra fakultetet. Om de så er indskrevet på institutter eller i SODAS, det er jeg sådan set fuldstændig ligeglad med, så længe der er nogen, der vil fortsætte med den eksperimenteren, som du blandt andet selv var med til starte op, My.

Indenfor antropologien har man blandt andet længe diskuteret forskerens positionering, og hvordan den spiller ind på metoderne, og hvordan metoderne spiller ind i data. Hvordan skal vi forholde os til det nu, hvor vi for eksempel indgår samarbejder på tværs af fagene, hvor folk har forskellige perspektiver på etiske spørgsmål?

Anders: Rent anekdotisk kan man sige, at det, vi har oplevet, når vi har været ude og snakke med en stor virksomhed, er, at vi får lov til at sige tre sætninger om et eller andet indholdsmæssigt, vi synes, er helt vildt spændende, og så kommer ordet „etik“ på bordet. For det er ligesom det, alle sidder og kigger på i øjeblikket. Og ingen af os har svarene, vel, for forståelserne rykker sig så hurtigt. Og, som du siger, bliver der blandet traditioner for at tænke etisk sammen på nogle måder, som er ret uoverskuelige i virkeligheden. Morten og jeg havde en lille snak den anden dag, fordi vi er i gang med at prøve, på vores egen helt amatøragtige måde, at lære os selv at kode. Og skulle vi så spørge de folk, vi „scrapet“ fra, om lov? Skulle vi give os til kende? Det ville man jo normalt gøre, hvis man kommer ud af en kvalitativ eller feltarbejds tradition. Alt andet ville være mærkeligt. Hvis man laver virtuel etnografi, ville det også være mærkeligt ikke at give sig til kende. Men mig bekendt er det absolut ikke praksis, at folk i datavidenskab giver sig til kende eller beder om lov. I hvert fald ikke fra de enkelte ejere af en side, for eksempel. Alle mulige spørgsmål dukker op hele tiden af sådan en etisk karakter.

Den kombination af viden, som I siger, at studerende på den nye uddannelse skal have en kvalificeret fornemmelse af, giver den mulighed for en øget etisk refleksion også? Måske fordi man kommer til at have indblik i forskellige niveauer af for eksempel, hvordan en algoritme bliver udviklet OG benyttet?

Anders: Helt klart. Det er meget præcist sagt. Vi kommer jo til at udbyde en palet af fag, og nogle vil være mere teknisk kodningspræget, andre vil være kvantitativ metodik af forskellig art. Noget vil være med feltarbejde, og noget vil handle om jura og etik som noget, der former datas sociale liv. Men vi har sagt til hinanden, at etiske spørgsmål hører til i alle de fag. Og det er jo udtryk for det, du påpeger: at man vil oparbejde forskellige etiske beredskaber til de forskellige steder, man kan befinde sig i relation til data. Og der bliver det vores opgave at udvikle det sprog, der kan håndtere at sætte tingene i relation til hinanden. Og måske også rydde lidt op, for det etiske sprog har vel også sådan en etisk egenskab af, at det kan vokse. Derfor er der egentlig grund til at tænke optimistisk, for det her med at sætte dataformer, forståelser af dem og evnerne til at omgås dem i centrum kan måske også gøre noget godt for bevidstheden om, hvor blikket på det sociale kommer fra.

Note

1. Morten Axel Pedersen er professor på Institut for Antropologi, Københavns Universitet. Som del af styregruppen for forskningscenter SODAS (PI) og den SODAS-baserede forskergruppe CALL har han arbejdet med kvali-kvantitative metoder, computational anthropology samt dataetik, -politik og -policy.

Anders Blok er lektor på Institut for Sociologi, Københavns Universitet. Han forsker blandt andet i, om verdens klimaforandringer giver anledning til nye risiko- og solidaritetsfællesskaber mellem mennesker, der befinder sig fjernt fra hinanden, geografisk såvel som socialt. Som del af den SODAS-baserede forskergruppe CALL har han arbejdet med og undersøgt kvali-kvantitative metoder.

Rebecca Adler-Nissen er leder af forskningsprojektet DIPLOFACE og forskningsgruppen Digital Disinformation og er blandt andet medlem af styregruppen for forskningscenter SODAS. Hun forsker i international politik, diplomati og EU's udvikling – i øjeblikket med fokus på forholdet mellem fortrolige internationale forhandlinger, offentligheden og sociale medier.

