

Dataetik

18 dataetiske anbefalinger fra DANSK IT



Indholdsfortegnelse

Forord	3
Anbefalinger fra DANSK IT	4
Data: Hvad er det?	7
Hvad forstår vi ved begrebet dataetik?	7
Hvad er en algoritme, og hvordan fungerer den?	9
Mange aspekter af dataetik	10
Positive muligheder ved en øget brug af data	11
Tillid er en grundpille i det danske samfund	13
Privacy by design og privacy by default	14
Datakvaliteten er helt afgørende	15
Samkøring af data	16
Indsigt i algoritmer og modeller	16
Digital arv – hvad sker der med vores data, når vi dør?	17
Demokratiske processer og retssikkerhed	18
Kontrolmekanismer	19
Sammenfatning	21
Arbejdsgruppen for dataetik og DANSK IT	22
Definitioner	23
Inspiration, litteratur og kildehenvisninger	24

Forord

Data er guld. Sådan har det i de senere år lydt igen og igen fra både virksomheder og myndigheder i debatten om den digitale fremtid. Data er en yderst vigtig ressource, når vi kigger på, hvordan virksomheder bliver drevet og udviklet i dag og i fremtiden, hvordan den offentlige velfærd kan styrkes, og hvordan danskernes hverdag i det hele taget kan blive lidt lettere. Vi har over de seneste årtier været vidne til, at digitaliseringen spiller en stadig større rolle i samfundet. Data er omdrejningspunktet for denne udvikling.

Datamængderne vokser hastigt. Megabytes er blevet til gigabytes og siden terabytes, petabytes og exabytes. Og det er ikke kun datamængderne, der stiger. Nye datakilder i form af blandt andet sensorer vælter frem, og hastigheden, hvormed data kan behandles, skyder ligeledes i vejret. Med big data og kunstig intelligens som nogle af de aktuelle megatrends har vi kurs mod en helt ny verden med data som fartøjets brændstof.

Som med alle andre ressourcer og råstoffer skal man tænke sig godt om, når man jagter de mange mulige gevinster. Data kan potentielt være guld, men en skødesløs brug af data kan skade borgernes tillid til myndighederne, til virksomhederne og til hinanden. Data kan ikke kun bringe os på rette vej, men også på afveje. Netop i disse år står vi ved en korsvej, hvor vi skal træffe en række fundamentalt vigtige beslutninger om det databaserede samfund, vi ønsker at leve i. Her er dataetik kompasset, vi kan navigere efter.

DANSK IT nedsatte i første halvår 2018 en arbejdsgruppe for dataetik, hvor en række medlemmer af foreningen fik til opgave at udpege de store dataetiske udfordringer i vores samfund og komme med anbefalinger til, hvordan vi kan sikre god dataetik i fremtiden. Det er en bredt sammensat arbejdsgruppe (se oversigt bagerst), hvor hvert medlem har bidraget med viden og kommentarer ud fra sit it-faglige ståsted. Denne publikation er resultatet af en række møder i arbejdsgruppen.

Arbejdsgruppen for dataetik har været igennem mange faglige og filosofiske diskussioner om, hvad vi forstår ved begrebet dataetik, og hvad god dataetik egentlig er? Der findes ikke endegyldige svar i debatten om dataetik, men der uden tvivl fundamentale spørgsmål og debatter, vi i et demokratisk samfund aldrig må tage for let på. Samtidig bør det understreges, at dataetik er andet og mere end lovgivning. Det er ofte svære diskussioner, hvor der ikke findes lette svar.

Det er vigtigt, at man, når man læser denne publikation, har for øje, at arbejdsgruppens opgave har været at forholde sig til de dataetiske udfordringer. Der er masser af positive ting ved en øget digitalisering i vores samfund, og de fordele fokuserer DANSK IT løbende på via andre arbejdsgrupper, udvalg og fagråd.

Forhåbentlig kan de kommende sider sætte tankerne i gang hos dig og din organisation. Det er nemlig i debatten om dataetik, at vi finder nøglen til en digital fremtid, hvor vi høster digitaliseringens gevinster uden at gå på kompromis med de værdier og principper, vores demokrati og retssikkerhed bygger på. Dataetik er nøglen til en digital fremtid i balance.

Anbefalinger fra DANSK IT

Herunder følger de konkrete anbefalinger fra DANSK IT's arbejdsgruppe for dataetik. Alle anbefalinger bygger på de analyser og overvejelser, der er gjort rede for på de følgende sider i denne publikation. Anbefalingerne skal således ses som en komprimeret liste, som virksomheder, myndigheder, beslutningstagere og helt almindelige borgere kan anvende i arbejdet med dataetik.

Flere ressourcer til Datatilsynet

Datatilsynet skal sikres flere ressourcer, hvis det skal have en reel mulighed for at være borgernes vagthund over for både offentlige myndigheder og private virksomheder og kunne kontrollere de sager, der ligger inden for tilsynets ressortområde.

Styrket mandat til Datarådet

Vi anbefaler samtidig et udvidet mandat til Datarådet, der er en del af Datatilsynet. Datarådet bør have mandat til at bære sager ind i den retspolitiske debat om databehandling, der også kan rumme dataetiske dilemmaer. Dataanvendelse er i hastig udvikling. Derfor er det afgørende, at Datarådet, der har indsigt i konkrete sager, bidrager med anbefalinger til lovgivning, der sikrer en fortsat beskyttelse af individets data.

Demokratisk kontrol med dataanvendelsen

Den til enhver tid siddende regering bør udarbejde en årlig udredning af myndigheders dataanvendelse, der behandles i Folketinget. Redegørelsen skal indeholde bidrag fra alle ministerier og skal sikre en åben og demokratisk drøftelse af udviklingen i anvendelsen af data.

Et dataetisk råd med den rette sammensætning

DANSK IT støtter tankerne om etablering af et Dataetisk råd. Rådet bør kunne beskæftige sig med tekniske, etiske og juridiske dilemmaer. Sammensætningen bør derfor være bred i forhold til faglig baggrund blandt rådets medlemmer.

Mennesket skal være i centrum

Vi skal som samfund udnytte de muligheder, teknologierne giver os, men det bør altid være med udgangspunkt i mennesket og menneskets behov og de værdier, vi som samfund vil have som bærende elementer.

Privacy by design i både det offentlige og private

Der skal arbejdes med privacy by design på alle niveauer både hos offentlige myndigheder og i private virksomheder. Alle nye systemer bør udvikles med udgangspunkt i privacy by design. Her er der især brug for, at it-fagfolk bliver inddraget i arbejdet og de nødvendige overvejelser.

Computergenererede afgørelser skal kunne begrundes

I højere og højere grad vil man kunne anvende algoritmer helt eller delvist til at sagsbehandle og fremkomme med endelig afgørelser. Borgeren bør altid kunne få indsigt i, hvilke af borgerens data der ligger til grund for afgørelsen. Ligeledes skal afgørelsen kunne begrundes. Det gælder både fra en offentlig myndighed og fra eksempelvis ens forsikringsselskab eller pensionsselskab.

Behandling af samkørte data

De medarbejdere, der arbejder med sammensatte data, bør lære, hvordan de omgås dem kritisk og tager højde for, hvor de kommer fra og hvilken sammenhæng, de er skabt i. De, der udvikler algoritmer eller sætter parametrene for selvlærende systemer, bør på samme måde forholde sig kritisk til hvilke data, der kan og bør indgå og med hvilken valør.

Borgeren ejer egne data

Det bør være et gennemgående princip i både det offentlige og private, at man som individ selv kan beslutte, hvilke data om ens person, der indhentes og opbevares, og hvordan data anvendes. Man bør altid kunne få indsigt i, hvor data, der vedrører ens person, stammer fra, til hvilket formål de blev indsamlet, samt hvem der har tilgået ens data. Herunder bør muligheden for at kunne sige fra også sikres. Konsekvensen kan eksempelvis være en mere besværlig sagsgang, fordi en offentlig myndighed ikke har de nødvendige data som udgangspunkt - men det bør altid være individets ret at kunne træffe det valg.

Sikring og beskyttelse af vores digitale arv

Individets ret til at bestemme over egne data skal beskyttes, også når personen er død. Man har ikke pt. mulighed for at beslutte, om de data, der genereres, må bruges til andre formål efter ens død. Der er heller ikke besluttet, om de efterladte kan bestemme det. Anbefalingen er derfor, at der bliver skabt klarhed over, hvordan man sikrer og beskytter sin digitale arv.

Den digitale dannelse skal styrkes

Der skal være et vedblivende fokus på at sikre og vedligeholde danskernes digitale dannelse, da denne sammen med digitale kompetencer er en forudsætning for, at man kan agere bevidst som beslutningsmyndig borger i det digitale samfund. Digital dannelse og digitale kompetencer kan styrkes via et kontinuerligt samarbejde mellem uddannelsesinstitutioner, myndigheder, virksomheder, interesseorganisationer og civilsamfundets mange foreninger.

Virksomheder skal være bevidste om egen indsats og rolle

I dag er data af meget stor værdi for private virksomheder, og kunder afgiver ofte data for en service eller et produkt lige såvel, som de betaler med penge. Dette bør afspejles i virksomhedernes årlige rapporteringer såvel som i deres politikker. Dataetisk adfærd bør ses som en konkurrenceparameter, men man bør også være opmærksom på, at intentionen om at tage dataetikken seriøst i nogle tilfælde kan betyde, at man må vinke farvel til en potentiel indtægt, hvis en given løsning ikke kan realiseres uden at gå på kompromis med dataetikken.

Internationalt samarbejde

Dataetiske anbefalinger er som udgangspunkt lige så relevante for små virksomheder som for internationale selskaber, men i forhold til store, internationale selskaber skal der mere til end nationale anbefalinger og indsatser. Kontrol og regulering af globale selskaber kræver international prioritering og vilje til at handle. Det er en vigtig opgave at understøtte for den danske regering, EU og andre internationale samarbejdspartnere.

Der skal etableres whistleblowerordninger

Virksomheder bør etablere en dataetisk whistleblowerordning. Dette skal give medarbejdere mulighed for anonymt at kunne rapportere om problematisk og uansvarlig anvendelse af data.

Ét sæt grunddata frem for de samme data mange steder

Arbejdsgruppen for dataetik anbefaler, at man såvel hos offentlige myndigheder som i private virksomheder tilstræber at have ét sæt grunddata om personer, der kan anvendes forskellige steder i organisationen til forskellige løsninger – frem for at man har de samme data liggende mange steder og skal holde dem opdaterede der. Det vil mindske risikoen for forældede data og ukorrekte data, og det vil gøre det lettere at få indsigt i, hvilke data en given virksomhed eller myndighed opbevarer om ens person, hvordan de anvendes og af hvem.

Et dataetisk løfte

Når man som it-professionel arbejder med persondata, har man en særlig forpligtelse til at gøre det på en ansvarlig måde og med de rette intentioner ud fra et dataetisk perspektiv. DANSK IT mener, at dette vil kunne styrkes ved at formulere et dataetisk løfte for dem, der arbejder med data og algoritmer. Det dataetiske løfte skal ses som en pendant til lægernes lægeløfte.

En dataetisk mærkningsordning

Der bør etableres en dataetisk mærkningsordning ud fra nogle klare principper, som virksomheder og myndigheder kan tage udgangspunkt i. Det vil gøre det langt lettere at være forbruger og borger, fordi man ved hjælp af mærkningsordningen har noget at træffe sine valg ud fra.

Viden om algoritmer og data på tværs af alle uddannelser

Dataetikken er langt fra kun en sag for it-folk. Det er derimod noget, som alle faggrupper skal kunne arbejde med. Derfor er det nødvendigt, at man på tværs af alle uddannelser giver de studerende en grundlæggende viden om, hvad data og algoritmer egentlig er, og hvorledes dataetik er en del af det job og de arbejdsopgaver, som de studerende på sigt skal varetage.

Data: Hvad er det?

Teknologi spiller en stadig vigtigere rolle i vores personlige liv og i vores samfund. Grundlæggende rettigheder for den enkelte afhænger ikke kun af lovrammer og sociale normer, men også af de teknologier, samfundet og den enkelte råder over. Det er nødvendigt, at udformningen og brugen af teknologi tager højde for behovet for at respektere den enkeltes rettigheder.

Ordet data stammer fra latin og er flertal af datum, hvilket kan oversættes til "noget, der er givet". Data kan betyde at give noget videre. At give data videre og være optaget af at bearbejde data må kunne tilskrives menneskelig nysgerrighed og virketrang. Vi har altid indsamlet, vurderet, fastholdt, gemt og anvendt data. Det, vi som mennesker har set og oplevet, er blevet brugt til at formidle viden og udvikle vores samfund og vores levemuligheder. Vi har videregivet data via helleristninger, symbolske skulpturer og vægmalerier. Vi har fastholdt traditioner og identitet via mønstre og former i klædedragter og vævemønstre. Vi har kommunikeret data via tegn og senere bogstaver, mangfoldiggjort via udvikling af trykkeri og kunst.

Mennesket har igennem tiderne opfundet mere og mere effektive måder at indsamle, vurdere, fastholde, gemme og anvende data på. Meget har dog ændret sig i de seneste 50 år i takt med digitaliseringen. Anvendelse af data og overvejelser om hvilken betydning anvendelsen kan få for vores liv og prioriteringer stiller nye krav til etisk ageren og omtanke. Overvejelser om dette vil blive uddybet i det følgende samtidig med, at dataetiske udfordringer og dilemmaer beskrives.

Dataetik handler i høj grad om transparens i forhold til adgang til data og brug af data og bias både i forbindelse med indsamling og anvendelse af data. Dataetik kan både være formuleret i handlinger og processer og som særlige etiske holdninger med et samlet dataetisk formål. Dataetikken udfordres og udvikles konstant i civilsamfundets samspil mellem:



I civilsamfundet kan borgeren optræde både som civilperson og som producent, myndighed og lignende på samme tid. Det stiller krav til dannelse og forståelse af viden, teknologi og rammer for anvendelsesmuligheder. Samtidig er det essentielt, at vi er bevidste om, at der i nogle situationer kan være forskellige roller og interesser alt efter, om man agerer som borger, stat eller/og marked.

Hvad forstår vi ved begrebet dataetik?

I takt med den omsiggribende digitalisering af stort set alle aspekter af menneskers liv opstår der enorme datamængder om og mellem mennesker. Allerede de tidlige, sporadiske erfaringer, vi nu kan gøre os med mulighederne i at anvende disse data – ofte til andre formål, end de egentlig er skabt til – vidner om et enormt potentiale for påvirkning af menneskers liv, individuelt og kollektivt. Denne påvirkning kan være af både positiv og negativ karakter, og mens vi i DANSK IT ofte er med til at fremhæve de mange positive muligheder, der skabes med digitaliseringen, mener vi også, at det er helt afgørende at italesætte de dilemmaer og udfordringer, der også følger med udviklingen.

På det individuelle plan ser man allerede nu ændrede muligheder for fx kreditvurdering og sundhedsydelser, og på det kollektive plan er selve demokratiet under pres med hidtil usete muligheder for manipulation. Vi har også set, hvordan såkaldt selvlærende kunstig intelligens kan forfalde til racisme eller kønsstereotyper.

Data er som udgangspunkt neutrale: De er ikke gode eller onde i sig selv, men de kan bruges til både onde og gode ting. Der er derfor behov for, at vi som demokratisk samfund begynder at diskutere, hvad vi vil med data, og hvordan vi undgår ondskabsfuld eller uhensigtsmæssig anvendelse af data. Det er en vanskelig og kompleks diskussion, som meget hurtigt bliver teknisk og indviklet – og på den måde en diskussion forbeholdt de særligt indviede. Det er vigtigt, at data og datas muligheder diskuteres på et plan, hvor man kan være med, også selv om man ikke kender forskel på bits og bytes. Det er derfor afgørende, at danskernes digitale dannelse styrkes.

Det er vigtigt, at man forstår, at vi blandt andet har brug for dataetik, fordi der er tale om spørgsmål og dilemmaer, vi ikke altid kan lovgive os ud af. Det er også essentielt, at der ikke er tale om diskussioner, hvor svaret er enten ja eller nej, men hvor der ofte vil være tale om kompromiser. Det kan eksempelvis være ud fra devisen: "Det skal skabe mest mulig lykke til flest mulige med mindst mulig skade."

Det er vigtigt, at datas anvendelse diskuteres på et etisk grundlag, og at vi accepterer, at vi har noget, vi kan kalde dataetik. Etik kan hurtigt blive til en indviklet filosofisk kategori, men i denne sammenhæng kan vi godt stille os tilfredse med at konstatere, at vores civilisation bygger på det etiske fundament, at intet er vigtigere end mennesket. Det enkelte menneske må aldrig reduceres til kun at være et middel, der er underordnet andre mål. Det kan derfor godt være, at man fx kan opnå noget eftertragtellesværdigt ved, at data om alle familier i Gladsaxe skal være tilgængelige for kommunen for, at man kan spotte udsatte familier. Men man bør spørge sig selv, om dette eftertragtellesværdige mål er så vigtigt, at det dermed er i orden at indsamle data om alle familier i Gladsaxe - også det store flertal, der er ikke-udsatte familier. Det er ikke nogen nem diskussion, men det er vigtigt, at vi tager diskussionen på et oplyst grundlag.

Problemerne med data og dataetik minder meget om problemerne med overvågning. Virksomheder, myndigheder og udviklere, som beskæftiger sig med en dataanvendelse, som også indebærer en form for overvågning, bør overveje både kort- og langsigtede etiske konsekvenser heraf. Konsekvenserne på den korte bane vedrører eksempelvis, at man rykker grænserne for, hvad der er tilladeligt at "snage i". Det kan være alvorligt nok, men på længere sigt bliver det måske endnu værre. Her risikerer man nemlig at pille ved vores fælles etiske fundament.

Man hører tit bemærkninger i retning af, at "jeg er ikke bange for overvågning, for jeg har jo ikke gjort noget galt". Her må man blot overveje, hvad overvågningen måske kan gøre ved os. Når man fx overholder hastighedsbegrænsningerne i trafikken, kan det begrundes på forskellige måder. Man kan overholde regler om trafik, fordi man ved, at det er det rigtige at gøre. Men der er også den mulighed, at man blot overholder reglerne, fordi man pga. overvågning ved, at man automatisk idømmes en straf, hvis man bryder dem. Det er to forskellige måder at agere på: Den gode borger overholder reglerne, fordi han ved, at det er det rigtige at gøre – og ikke blot fordi han er bange for straf.

Problemet er, at hvis den automatiske straf pga. overvågning er en sikker sag, ved vi ikke længere med sikkerhed, om vi overholder reglerne, fordi vi er bange for at få en straf, eller fordi vi ved, at det er det rigtige at gøre. Overvågning – hvad enten det nu er med videokameraer eller med data – truer på den måde med at erodere vores fælles, etiske fundament.

Som alle andre etiske spørgsmål vedrører spørgsmålet om dataetik grundlæggende, hvad der er det gode og det rigtige for det enkelte menneske og for fællesskabet. Det vil altid være en afvejning af forskellige hensyn, og det vil ofte være dilemmafyldt. Men det betyder ikke, at alt kan være lige rigtigt eller lige forkert. Vi må som samfund og som individer tage stilling og etablere nogle normer, der af de fleste mennesker i de fleste situationer opleves som rigtige og retfærdige. Sådanne normer ændrer sig over tid, og de bør derfor med mellemrum tages op til overvejelse og debat. Der findes ingen evigtgyldige svar i debatten om dataetik, men der bør findes svar, som har en høj grad af permanens og som kun ændres langsomt.

Hvad er en algoritme, og hvordan fungerer den?

En algoritme er kort fortalt en beskrivelse af, hvordan man løser et problem af en bestemt type. Algoritmen leverer en løsning eller et resultat på problemet.

I vort digitaliserede samfund er vi alle udsat for computeres anvendelse af algoritmer til at træffe afgørelser hver dag, lige fra den simple beregning af moms på rundstykkerne hos bageren, til de lidt mere komplekse beregninger af vores skat, eller om vi er berettigede til fx boligstøtte og i så fald hvor meget.

Når man går fra en algoritme, som generelt kan beregne moms, og til en helt konkret moms i Danmark, hvor momssatsen er 25%, så taler man ofte om en model, altså den konkrete udmøntning af algoritmen. De fleste af os accepterer computerens afgørelse. Vi kan i større eller mindre grad gennemskue modellerne og har for det meste tillid til dem, der har ansvaret for at lave programmerne. Det er fint, for modeller er for det meste nyttige og hjælper os med at være effektive.

Nogle gange møder vi algoritmer og modeller uden at vide det. Fx i webshoppen, hvor vi ofte præsenteres for budskabet om, at "andre, der købte X, købte også Y og Z". Dette kan være både hjælpsomt og/eller irriterende, men det er oftest acceptabelt. Men når vi har søgt efter fx en speedbåd via en søgemaskine, så undrer vi os, når vi samme dag andre steder på nettet "tilfældigvis" mødes af reklamer om præcis dette – og føler os overvåget. Der er algoritmer og modeller på spil, og de næres af data, som vi sikkert - men uden fuldt at gennemskue det - har sagt ja til at dele.

Udviklingen har givet os muligheden for at udvikle eller videreudvikle på algoritmer/modeller ved hjælp af data. Eller nærmere bestemt statistisk analyse. Denne kombination kaldes machine learning og betyder, at computeren lærer at genkende mønstre ud fra eksempler og data, der er velkendte (dvs. træningsdatasæt) og dermed opbygge modeller. Hvis der er en tilstrækkelig mængde af data og kvalitet i data, vil modellerne blive konsistente og stabile og kunne bruges til at forudsige det sandsynlige resultat af eksempler, hvor resultatet ikke er kendt på forhånd. Dette er selvfølgelig utroligt kraftfuldt og rummer mange positive muligheder, særligt på områder, hvor antallet af variable og kompleksiteten er for voldsom til, at vi kan udlede en model på anden vis.

Disse teknologier er et af de mest kraftfulde teknologiske spring de seneste år, hvor teknologien har givet mulighed for at gennemføre de bagvedliggende komplicerede matematiske beregninger, der i teorien har været kendt i flere årtier, men først nu kan gennemføres i praksis. De positive muligheder for at udnytte teknologierne til "noget godt" er enorme: Der er potentiale til at redde liv, forbedre livskvalitet, bekæmpe sult osv. via forudsigelse af sygdomme, prioritering af behandling, optimering af gødning i forhold til høstudbytte og miljøpåvirkning osv.

Med store muligheder følger ofte risiko for tilsvarende store faldgruber - således også her. Hvis fx der i træningsdatasættet er en skævhed (bias) i data, så vil modellen også blive skæv. Og det kan gå helt galt, hvis man fx vil bygge en model til at hjælpe domstole med at dømme i sager og så fodrer modellen med udfaldet i en masse konkrete sager fra USA, hvor risikoen beviseligt er større for at blive dømt, hvis man er af afro-amerikansk oprindelse. Ja, så skaber man en racistisk model.

"Machines get biased because the training data they're fed may not be fully representative of what you're trying to teach them."

- Guru Banavar, IBM Chief Science Officer for Cognitive Computing

"AI can be used for social good. But it can also be used for other types of social impact in which one man's good is another man's evil. We must remain aware of that."

- James Hendler, Director of the Institute for Data Exploration and Applications, Rensselaer Polytechnic Institute

Teknologierne kan også misbruges på andre måder. Den samme model, som fx kan forudsige risiko for sygdom, så man som borger kan tage skridt til at forbygge konsekvenserne, kan også bruges af et forsikringselskab til at nægte samme borger en livsforsikring eller kræve forhøjelse af præmierne.

Algoritmer, modeller, machine learning og kunstig intelligens er kommet for at blive og vil få stigende betydning de kommende år. Og det skal de områder også have; gevinsterne for den enkelte borger, for virksomhederne og for samfundet er alt for store til at kunne ignoreres. Efterhånden som modellerne i stigende grad bliver uigennemskuelige for de fleste pga. kompleksiteten, må vi derfor sikre os, at det, vi lærer dem via træningsdata, styringen af hvem der har adgang, og hvad vi bruger dem til, foregår med etisk ubeklækkelige mål og med høje faglige standarder, så vi fortsat kan have tillid til det, der foregår.

Mange aspekter af dataetik

"La guerre! C'est une chose trop grave pour la confier à des militaires."

- Georges Clemenceau

Denne kendte bemærkning (krig er for alvorlig en sag at overlade til generaler) anvendes ofte til at antyde, at dette eller hint har så vidtrækkende konsekvenser, at det ikke er nok at stole på såkaldte fagfolks opfattelse af, hvad der er rigtigt og forkert. Pointen passer udmærket på spørgsmålet om dataetik.

Konsekvenserne af big data, machine learning og kunstig intelligens er nemlig så vidtrækkende, at det er et anliggende for os alle. Vi er så at sige i samme båd, uanset om vi er adel, borgere, præster eller bønder.

Herunder identificeres både tekniske, demokratiske, filosofiske og juridiske aspekter af problemstillingerne.

De tekniske aspekter: En lang række datasikkerhedsmæssige problemstillinger er primært af teknisk art: Hvordan sikres det, at uvedkommende ikke får adgang til data? En lang række formelle problemstillinger er tilsvarende af bl.a. teknisk art: Hvordan sikres det, at algoritmerne regner, som de skal? Hvordan dokumenteres algoritmer, som er selvlærende? Hvordan sikrer man en høj nok datakvalitet?

De demokratiske aspekter: Med store datamængder og kunstig intelligens kan grænsen mellem fakta og fup nemt udviskes. Det bliver stadig vanskeligere at verificere, om en udtalelse, et billede eller en filmstump er ægte. Den demokratiske samtale, som er et fundament under de vestlige samfundssystemer,

bygger på et grundlag af tillid og verificerbarhed. Dette grundlag er sat under pres, og vores demokratiske institutioner må finde løsninger herpå.

De filosofiske aspekter: Upassende omgang med data er forbudt. Det regulerer bl.a. GDPR. Men *hvorfor* er det forbudt? Vi forstår alle umiddelbart, hvorfor det er forbudt at køre for hurtigt i bil i bymæssig bebyggelse. Det kan nemlig volde stor skade på mennesker og ting, hvis det går galt. Men hvad med data? Det er her, filosofisk indsigt er vigtig. I det omfang, at big data, AI og machine learning påvirker vores identitet, vores kultur og vores humanitet, kan filosoffer hjælpe os med at forstå, hvordan den teknologiske udvikling kan både skade og styrke vores samfund.

De juridiske aspekter: Big data, AI og machine learning læner sig op ad en række komplekse juridiske problemstillinger. Selvlærende algoritmer, der "træffer beslutninger", som viser sig at være forkerte, skadevoldende eller uretfærdige: Hvem har ansvaret for det? Selvlærende algoritmer, som skaber nyt indhold eller modificerer eksisterende værker: Hvem har ophavsretten, og hvem har ansvaret, hvis andres intellektuelle rettigheder krænkes?

En sammenhængende og fyldestgørende vurdering af eventuelle konsekvenser af en given anvendelse af store datamængder, kunstig intelligens og machine learning kan derfor ikke overlades til en bestemt faggruppe eller indfaldsvinkel. En samtidig belysning fra mindst de fire indfaldsvinkler, som nævnes ovenfor, må være ønskelig. Det er i den forbindelse vigtigt at notere, at retten til at anlægge de perspektiver, som kan anføres under de forskellige aspekter, ikke kun tilhører eksempelvis jurister, filosoffer, politologer eller ingeniører. Alle har lov til at blande sig – men alle må derfor også finde sig i, at nogen kan have mere indsigt end andre.

Positive muligheder ved en øget brug af data

Globalisering og digitalisering har "elimineret" tid og sted således, at information – eller data – er tilgængeligt globalt og typisk i løbet af nogle få sekunder. Denne udvikling giver kolossale muligheder for værdiskabelse og øget velfærd.

Danmark og danskerne er sandsynligvis positioneret allerbedst i verden for at kunne skabe værdi ud fra dette datapotential, bl.a. på grund af vores store tillid, et allerede rigt datagrundlag og gode kompetencer og fortrolighed med it/digitale platforme. Vi skal "bare" udnytte disse muligheder proaktivt.

Ved enhver opfindelse (eksempelvis elektricitet eller bilen) eller brug af nye ressourcer (som olie, gas eller data) overskygger kritik og bekymringer ofte de mange muligheder. Men det er her, at Danmark har muligheden for at skabe den dataetiske samfundsmodel og proaktivt skabe værdi samtidig med, at dataetiske principper tilsikrer samfundsborgerens stigende tillid, retssikkerhed og bedre velfærd. Dette vil hurtigt kunne blive et forbillede for mange andre lande og give eksportmuligheder for danske virksomheder og samfundet generelt.

I dag bidrager den enkelte borger til vores fælles velfærd ved at betale skat og afgifter eller ved at uddanne sig til i fremtiden at kunne bidrage, men her har vi med data yderligere muligheden for, at alle borgere kan bidrage. For med data kan vi alle bidrage (bevidst eller ubevidst) med vores data og dermed gøres samfundets datagrundlag endnu mere værdifuldt for vores fælles velfærd. Nogle konkrete eksempler på dette er:

Data om vores færden (GPS-tracking):

- Giver mulighed for bedre (og billigere) planlægning af ny infrastruktur som veje og nye offentlige transportruter.
- Bedre, billigere og hurtigere prioritering af infrastruktureparationer.
- Dynamisk (næste realtids) justering af lysregulering afhængig af trafiksituationen netop nu og deraf mindre spildtid i trafikken, mindre forurening og lavere transportomkostninger.
- Ambulance, politi og brandvæsen kan komme hurtigere frem.

Overvågningskameraer:

- Giver mulighed for øget sikkerhed og højere retssikkerhed. F.eks. kunne videoovervågning på landets stationer analysere adfærdsmønstre og i tilfælde af "mulig lommetyv" automatisk udsende et højttalerudkald om "pas på lommetyv".
- Eller et mere tragisk adfærdsmønster som "mulig selvmordskandidat" kan få tilkaldt hjælp og undgå en tragisk hændelse samt en togførers og mulige øjenvidners behov for krisehjælp og potentielle langtidssygemeldinger.

Samkøring af eksisterende data:

- Giver også mulighed for øget sikkerhed og beskyttelse af den enkeltes ejendom. F.eks. hurtig rapportering og lokalisering af en stjålet bil eller mere avancerede analyser, der giver politiet mulighed for patruljering der, hvor risiko for indbrud den pågældende dag/time er størst.
- Eller noget mere uskyldigt som automatisk advisering om fornyelse af et pas, der udløber. Specielt op til de typiske ferieperioder er det en god service til borgeren og sandsynligvis bedre for medarbejderne i borgerservice, da vi kan reducere eller måske helt undgå en stressfyldt periode med meget ekstra arbejde (endda lige op til medarbejdernes egen ferie).
- Med sundhedsdata er der mulighed for hurtigere helbredelse, og i specielle tilfælde kan sjældne sygdomme hurtigere diagnosticeres og behandles.

Data kan selvfølgelig misbruges, ligesom mange værktøjer, vi kan købe i ethvert byggemarked, kan misbruges. Derfor skal vi (hvor det ikke allerede er tilfældet) have samme retssikkerhed for, at de personer, der bruger værktøjer – eller data – til noget ulovligt, bliver retsforfulgt.

Mange af disse muligheder betyder ikke, at staten – eller de brugere, som staten stiller data til rådighed for – får direkte adgang til data om os som enkeltpersoner. For princippet om "privacy by design and default" skal Danmark være verdensmester i. Det handler ikke kun om den tekniske udførelse, men også om en høj lægmandsforståelse hos alle borgere, så der er en almen forståelse af og forventning om, at man bidrager til samfundet med sine data, fordi det skaber værdi – og at man kan vide sig sikker, fordi data anonymiseres.

Tillid er en grundpille i det danske samfund

Danskernes tillid til hinanden, til det politiske system og til de offentlige institutioner er i særklasse. Tilliden placerer det danske samfund i en kategori, hvor kun få andre nationer kan være med. Denne tillid - også til de institutioner, som forvalter magten over os, og til de institutioner, som på godt og ondt griber ind i vores liv - er noget, som de fleste opfatter som et absolut gode. Undertiden tales der også om, at det er den "naturressource", vores lille land besidder, hvor andre har kul, jern eller olie. Det høje tillidsniveau gør os effektive og gode til tillidsfuldt samarbejde. På den måde bliver det høje danske tillidsniveau set som en konkurrencemæssig fordel i en globaliseret og digitaliseret verden.

Uanset om man nu lægger vægt på det formålstjenstlige eller mere ser på det etiske aspekt, er den gensidige tillid et gode, som vi utvivlsomt bør værne om og om muligt endda styrke. Vi må nok indse, at hvor rodfæstet og solid denne tillid end kan forekomme os at være, kan den mistrives og visne, hvis der ikke passes på den. Der skal både vandes og gødes, og ukrudt skal holdes på afstand.

Vigtigt er det især, at vi undgår lemfældig omgang med (andre menneskers) data. Hvis myndigheder og virksomheder tager vare på danskernes data på en måde, som danskerne ikke kan være tjente med, vil det uden tvivl erodere denne tillid. Hver gang en offentlig institution eller en virksomhed bliver grebet i at sjuske med sikkerheden eller endnu værre i at misbruge data, som folk i tillid har overladt til dem, gnaver det i tillidens rod. Hver gang en myndighed eller en virksomhed i jagt på øget effektivitet eller profit sætter hensynet hertil over hensynet til det enkelte menneske og reducerer det fra at være et medmenneske til blot at være et middel, piller der ved forudsætningerne for den gensidige tillid os danskere imellem.

Samtidig indebærer digitalisering, automatisering og centralisering, at antallet af fysiske møder mellem borger og myndighed/virksomhed reduceres væsentligt. Borgeren møder oftere enten en anonymiseret stemme eller slet og ret ikke noget menneske overhovedet. Spørgsmålet er, om der er en mindstegrænse for, hvor ofte man skal have den personlige kontakt, for at tilliden til de abstrakte systemer kan holdes i live? Om der er særlige situationer, hvor den personlige kontakt har særlig stor betydning for tilliden? Eller om der er særlige måder at have kontakt på, fx pr. telefon eller pr. video, der kan bidrage til at bekræfte tilliden? Det ved vi ikke, og det er et problem. Der er i det hele taget alt for lidt viden om konsekvenserne for tillid, retfærdighed og sammenhængskraft ved de omfattende ændringer, man i disse år gennemfører i den måde, hvorpå borgerne møder og interagerer med systemet (myndigheder og virksomheder) på.

Når data anvendes, og interaktioner digitaliseres/automatiseres, skal der derfor tænkes langsigtet. Man bør altid spørge sig selv, hvilke positive og negative konsekvenser, der er af nye tiltag med eksempelvis automatisering. Et vigtigt spørgsmål bør her være, om man risikerer at reducere det enkelte menneske til et middel eller et objekt, så man går fra at være et medmenneske til at være et datasæt? Det er etisk problematisk og kan påvirke borgerens oplevelse af at kunne træde i karakter som individ og af at blive set, hørt, respekteret og taget alvorligt. I det lange løb kan det risikere at underminere vores danske tillid. Det betyder ikke, at man skal lade være med at anvende data, men det betyder, at man skal anvende data på en ordentlig måde. Det er værd at minde om, at dataanvendelse rummer mange positive muligheder – og også kan bidrage til samfundets sammenhængskraft og tillidsniveau. Det er via anvendelse af data, at samfundsnedbrydende ting som skattesnyd og selskabstømning etc. forebygges og opklares.

Tillid skabes ved ordentlig opførsel. I dataetisk sammenhæng indebærer dette gennemsigtighed, klare aftaler, tydelighed, forståelighed og respekt for det enkelte menneskes integritet.

Privacy by design og privacy by default

Med EU's databeskyttelsesforordnings ikrafttrædelse den 25. maj 2018 blev databeskyttelse ved design (privacy by design) og som standard (privacy by default) en juridisk forpligtelse.

Det generelle princip om privacy by design og privacy by default omfatter en etisk dimension i overensstemmelse med principperne og værdierne i EU's charter om grundlæggende rettigheder og de specifikke juridiske forpligtelser i forordningen.

Privacy by design er en tilgang, der sikrer, at virksomheden indarbejder databeskyttelsen som en integreret del af virksomhedens forretningsprocesser, værdikæde og produktionscyklus. Alt dette skal ske lige fra produktionsfasen til, at produktet havner hos slutbrugeren. Kravet om databeskyttelse gennem design indebærer, at du som dataansvarlig har en generel overvejelses- og håndteringsforpligtelse.

Privacy by default betyder, at produkter allerede fra start er indstillet til at sikre den højeste persondatabeskyttelse. Videre skal man sikre sig, at personoplysninger om en bruger kun bliver opbevaret så længe som nødvendigt for at få leveret et produkt eller en service. Man skal også sikre sig, at personoplysninger ikke uden den registreredes indgriben stilles til rådighed for et ubegrænset antal personer.

Efter ordlyden af forordningen skal standardindstillinger forstås bredt, således at det omfatter både tekniske og organisatoriske foranstaltninger. Der er i forordningen også en pligt for den dataansvarlige til at sikre, at når f.eks. et softwareprogram, en onlinetjeneste, et it-system eller lignende anvendes til at behandle personoplysninger, skal de indstillingsmuligheder, som systemet indeholder, som standard indstilles på en måde, der understøtter forordningens krav i artikel 25, stk. 2, om bl.a. formålsspecifik behandling af personoplysninger.

Formålet med begge principper er at minimere mængden af data, der automatisk deles, og sørge for, at forbrugerne som udgangspunkt selv har ret til at bestemme, hvor meget af deres data, der skal deles mv. Derfor er det nødvendigt, at ovenstående principper følges hver eneste gang, man skal udvikle løsninger, der behandler personoplysninger. Videre kan kravet om databeskyttelse gennem standardindstillinger (artikel 25, stk. 2) ses som et påkrævet supplement til kravet om databeskyttelse gennem design (artikel 25, stk. 1).

Der er behov for at fremme den nyeste teknologi og brug af privatlivsforbedrende løsninger. Der skal udvikles privacy-bevidste produkter og strukturer ud fra retningslinjer for ikke manipulerende adfærdsdesign og med indbygget samtykkehåndtering. Design- og driftsprocessen skal overveje hele livscyklussen for en tjeneste eller et produkt, fra den første planlægning til service og frem til bortskaffelse af produktet.

Databeskyttelse gennem design skabes på basis af principper som **proportionalitet**, hvor mål og metode skal være i balance, **beskyttelse og kontrol** ved at have gennemsigtighed og ansvarlighed i forbindelse med processer og **inddragelse og involvering** ved at være åben og transparent i forhold til produkt og flow.

Databeskyttelse gennem design kræver, at når mål og metode fastlægges, skal man vælge tekniske og organisatoriske foranstaltninger for at sikre en effektiv implementering af de grundlæggende databeskyttelsesprincipper. I designfasen laves samtidig en analyse og vurdering af de sociale og etiske konsekvenser af teknologien.

Foranstaltninger, der kan indbygges i virksomhedernes drift og udgøre databeskyttelse gennem design, kan eksempelvis være:

- Tekniske foranstaltninger og sikring af infrastrukturen mod uautoriseret indtrængen
- Kryptering af data i transit eller hvile
- Pseudonymisering og anonymisering samt løbende opdatering af software
- Organisatoriske foranstaltninger som principper om minimering af persondatabehandlingen
- Implementering af it-sikkerhedspolitik samt opfølgende træning af medarbejdere
- Anvende designmetoder og forretningsmodeller med udgangspunkt i privatlivsteknologi
- Databeskyttelse gennem standardindstillinger, eksempelvis dynamisk rettighedsdeling
- Slettet og sendt post slettes automatisk fra e-mail konto efter 30 dage

Datakvaliteten er helt afgørende

Data er værdifulde, og værdifulde data rummer muligheder, som kan transformeres til eksempelvis penge, tilfredse borgere, nye indsigter og nye muligheder. Det er de muligheder, der ligger til grund for den megen hype, der er om big data, kunstig intelligens og machine learning mv.

Når så virkeligheden indfinder sig, opdager de fleste, at mange af disse vidunderlige data ikke er så anvendelige og værdifulde endda. De viser sig ofte at være forældede, dårligt formaterede og med datamodeller, der passer til et eller andet specifikt behov, som var der på det tidspunkt, hvor modellerne blev lavet. Man opdager som oftest også, at data er dårligt dokumenterede, har uigennemskuelige feltnavne og savner sammenhæng i nøglerne. Man finder ud af, at felter, som var beregnet til ét formål, undervejs er blevet brugt til et eller flere andre formål af brugere, som syntes, det var lettere end at overbevise it-afdelingen eller leverandøren om at ændre i systemet.

Det er indlysende, at anvendelse af sådanne data kun bør ske med store forbehold og med stor forsigtighed. Det er også indlysende, at hvis man ydermere begynder at kombinere data fra flere forskellige kilder, skabt på forskellige tidspunkter og med forskellige formål for øje, bliver det endnu mere usikkert.

Et eksempel fra den virkelige verden kan illustrere, at data ikke altid er, hvad de giver sig ud for at være:

I en dansk kommune er tryghedsbesøg inden for ældreplejen et kendt og udbredt fænomen. Et tryghedsbesøg er et besøg, hjemmeplejen aflægger hos en medborger, som egentlig ikke har behov for nogen særlig eller praktisk bistand på dét tidspunkt, men som klart trives bedst med ikke at føle sig ensom og forladt. Der kan for kommunen være en udmærket business case i sådan et tryghedsbesøg, eftersom den pågældende borger til gengæld eksempelvis undlader at bruge sit nødkald, typisk uden for normal arbejdstid, eller på anden måde medfører ekstra opgaver for hjemmeplejen. Imidlertid findes ydelsen "tryghedsbesøg" desværre ikke i kommunens ydelseskatalog. Det kan med andre ord ikke lade sig gøre at bevilge et tryghedsbesøg. Personalet skal dog stadig vælge en ydelse, når de registrerer deres besøg. De kalder det derfor bare noget andet. Forestiller man sig nu, at nogle kollegaer, fx i økonomiafdelingen, lidt for ukritisk baserer sig på de data, der kommer ud af ovenstående lille udpluk, siger det sig selv, at resultaterne meget vel kan være grundigt misvisende – for under hvilke kategorier er tryghedsbesøget nu blevet registreret?

I nogle sammenhænge kan vi måske godt acceptere, at der er fejl og usikkerhed, hvis vi til gengæld får nogle nye muligheder. Når det drejer sig om sundhed, demokrati eller økonomi, er det dog bedst at stille nogle krav. Eksempelvis ville det være godt, hvis den usikkerhed, de mange fejlmuligheder bidrager med i analysen af data, bliver gjort synlig og målelig.

Samkøring af data

Før man begyndte at dele data, blev data typisk til og anvendt og fortolket i den faglige og organisatoriske kontekst, hvor data var skabt. Ofte af de samme mennesker, der havde skabt data, eller deres nære kolleger. Når man på dataform beskrev et menneskes eller en families situation og problemer, ville dem, der efterfølgende brugte disse data, fordi de havde den samme forståelsesramme og måske havde mødt de mennesker, det drejede sig om, kunne tilføje kontekst, sætte ansigt og krop på og implicit forstå, hvad der lå bag data – herunder også begrænsningerne.

Når de samme data i dag anvendes i en helt anden kontekst og på et andet tidspunkt, evt. sammen med andre data og af medarbejdere med andre fagligheder, er der en risiko for, at mennesket bliver = data. Data forstås og fortolkes måske på en anden måde, end de blev i den kontekst, de blev skabt i. Indgår de sammenstillede data i automatiserede processer, ligger tolkningen blot et andet sted – hos dem, der laver algoritmerne, eller skjult i de bias, der er i de aggregerede data.

Det er uundgåeligt, at der vil være et konteksttab og en risiko for fejltolkning, når oplysninger bruges på tværs af tid og sted – det kan enhver historiker tale med om. Men hvor historikerne har deres kildekritiske metode, er dette ofte fraværende i administrative processer. Sådan behøver det heldigvis ikke at være. De medarbejdere, der arbejder med sammensatte data, bør lære, hvordan de omgås dem kritisk og tager højde for, hvor de kommer fra og hvilken sammenhæng, de er skabt i. De, der udvikler algoritmer eller sætter parametrene for selvlærende systemer, bør på samme måde forholde sig kritisk til hvilke data, der kan og bør indgå og med hvilken valør. Måske skal data kasseres, fordi valøren er for usikker, eller udtages til manuel tolkning (af kildekritiske medarbejdere).

Risikoen ved ikke at gøre dette er ikke kun, at der træffes forkerte afgørelser, men også at borgerne kan føle sig uretfærdigt behandlede af et uigennemsigtigt og umenneskeligt system. Og det kan svække tilliden til systemet, underminere dets sociale legitimitet og mindske borgernes vilje til at efterleve regler og afgørelser af sig selv.

Indsigt i algoritmer og modeller

I bogen *Weapons of math destruction* advarer den amerikanske matematiker Cathy O'Neil om, hvordan myndigheders og andres voksende brug af algoritmer kan forstærke uligheder og cementere magtpositioner. Mennesket giver data til teknologierne, men hvad kræver det af mennesket at forstå sammenhænge og eventuelle misbrug? Selvlærende algoritmer, "decision making algorithms", "cognitive computing" eller "machine learning" er alle begreber, det kræver viden at tage stilling til og forstå konsekvenser af.

Cathy O'Neil påpeger, at algoritmer er tilbøjelige til at forstærke de eksisterende magtstrukturer, fordi algoritmerne kan bruges til at automatisere behandlingen af den svageste og mindst velstående del af befolkningen. Computersystemerne er sorte bokse, og det er kun et fåtal, der overhovedet har den tekniske indsigt og viden til at forstå, hvordan algoritmerne når frem til deres konklusioner, når de f.eks. laver kreditvurderinger eller frasorterer jobansøgninger.

For nylig efterlyste EU-Kommissionen i en rapport om etik og kunstig intelligens, at algoritmer skal være *explainable*, dvs. at man skal kunne inspicere en algoritme og gennemskue, hvordan den når frem til sine konklusioner. Rapporten argumenterede for, at vi alle har en ret til forklaring af formål og anvendelse af en teknologi, fremfor bare en ret til underretning, hvis der sker noget.

Det handler om at gøre det muligt at forklare og argumentere for formålet med en algoritme, da algoritmer langt fra er neutrale og nøgterne, som man skulle tro. Som udgangspunkt skal en algoritme instrueres i, hvilke mål, den skal opnå, og hvilke parametre, den skal måle og analysere. Dermed indlejrer den menneskelige programmør uvægerligt en række værdier og mulige skævheder i algoritmen, som typisk vil afspejle og understøtte interesserne hos den instans eller myndighed, der driver computersystemet. Beslutninger taget på baggrund af dataprofiler og algoritmer er ikke nødvendigvis nøjagtige og fri for partiskhed eller skøn. Vi bliver heller ikke nødvendigvis spurgt efter samtykke, når vores data bliver behandlet og genbrugt kontinuerligt i intelligente dataprocesser. Der er ikke et klart retsgrundlag gældende for dataprofilering, og vi kender ikke altid ræsonnementerne, der leder frem til bestemte beslutninger og handlinger på baggrund af dataprofilerne. Dermed ikke sagt, at alting var perfekt, før digitaliseringen slog igennem, for den menneskelige vurdering, der tidligere har været omdrejningspunktet for eksempelvis analysen af en borgers behov, rummer også problemstillinger.

Når en algoritme forskelsbehandler, er det svært at opdage – og tilsvarende svært at klage over. I andre tilfælde ved programmørerne måske ikke engang præcist, hvordan algoritmerne tager beslutninger. Med machine learning kan computeren selv tage ved lære og opbygge erfaringer, efterhånden som den analyserer stadig flere data, og derfor kan det i praksis være umuligt for mennesker at se, hvad det er for antagelser og principper, den arbejder ud fra. Dette kan åbne uanede muligheder men også skabe etiske dilemmaer.

Målet må derfor være, at man skal kunne inspicere algoritmer og vurdere, om deres beslutninger er tilbøjelige til at diskriminere eller på andre måder har skadelige bivirkninger for dem, der bliver vurderet af systemet. Vurderingen foreskriver et minimumskrav for, om en algoritme kan siges at fungere på en fair og forsvarlig måde. I fremtiden vil det muligvis være et helt almindeligt krav fra investorer eller myndigheder, at virksomheder kan dokumentere, at deres algoritmer tager beslutninger på en rimelig og etisk forsvarlig måde. Det vil aktionærene forlange.

Revision og mærkning af algoritmer vil blive et vigtigt signal, som kan bruges som en konkurrencefordel af de virksomheder, der er certificerede, på samme måde som fairtrade-, økologi- og energimærkning fungerer i dag.

Digital arv – hvad sker der med vores data, når vi dør?

Arv er ikke længere blot mursten, værdipapirer og malerier. Med vores øgede digitale tilstedeværelse, efterlader vi også en lang række værdifulde ting på nettet. Vi efterlader en digital arv i alle de spor, som der efterlades på computere eller i skyen. Digital arv er alt fra postmapper, fotos, dokumenter, kontrakter, digitalt underskrevne dokumenter til f.eks. digitale værdier og gevinster og forudbetalte abonnementer og profiler på onlineplatforme.

Der er ingen automatisk nedlukning på medier og platforme, ligesom der heller ikke er automatisk udbetaling af indeståender ved dødsfald. Personlige data forsvinder ikke fra nettet, de forbliver i cyberspace. Digitalt kan vi leve videre for evigt.

Post mortem data er et nationalt anliggende, også efter 2018 og GDPR. Det er således op til national lovgivning at beslutte på hvilken måde, man ønsker at beskytte afdødes ret til privatlivets fred. Man har i Danmark vedtaget, at databeskyttelsesforordningen finder anvendelse på oplysninger om afdøde personer i 10 år efter personens død. Der er ikke uddybet yderligere i forhold til post mortem data. Der er således

ingen national praksis, ligesom der heller ikke er nogle internationale regler eller en certificering af dataetiske systemer for at undgå krænkelse af den afdødes privatliv.

Der er ikke taget stilling til, hvordan individets rettigheder, herunder "the right to be forgotten", skal balancere i forhold til samfundets interesser, det være sig økonomiske, praktiske, sensitive eller sentimentale interesser. Samfundet kan f.eks. have interesse i afdødes sundhedsdata. Man har ikke pt. mulighed for at beslutte, om de data, der genereres online, må bruges til andre formål efter ens død. Der er heller ikke besluttet, om de efterladte kan bestemme det.

Det er endnu ikke klart, hvilke konsekvenser det kan have for den enkelte, de efterladte og/eller for samfundet, at dette område ikke er yderligere reguleret. Det er lige nu op til den enkelte at tage stilling til sin digitale arv, ligesom man tager stilling til sin fysiske arv. Man skal beslutte, hvad der skal ske med ens oplysninger og informationer på nettet. Der findes ingen automatik for, hvem der skal råde over ens digitale dødsbo. Man må derfor selv udvælge en person, der kan få adgang til ens konti i tilfælde af dødsfald. De dataetiske overvejelser er ikke overstået efter ens død, de fortsætter.

Et eksempel på et dilemma: Staten ønsker at forebygge og behandle lidelser, der er belastende både menneskeligt og økonomisk - såsom psykiske lidelser. Borgerne har en lang række sundhedsdata liggende på borger.dk. Der er samtidigt voksende og indbringende marked for at udvikle medicin og behandlingsmetoder for stat såvel som private virksomheder. Staten ønsker at videregive uidentificerbare oplysninger til både private og offentlige institutioner, der har et formål inden for forebyggelse og behandling.

Oplysningerne er både på levende og afdøde, herunder oplysninger vedr. dødsårsager såsom selvmord. Arvelighed bliver et fokusområde: Der er måske genetiske spor, der - hvis de bliver identificeret tidligt - kan bidrage til en effektiv forebyggelse af menneskelige lidelser og tunge behandlingsbudgetter. Samkøring af registre er muligt, hvis man må benytte identificerbare data.

Hvem skal give tilladelse til anvendelse af identificerbare data? Skal en familie med et tragisk selvmord som livsbagage, man har brugt år på at bearbejde, tvinges til at tage stilling til adgang til en afdød fars sundhedsdata til gavn for forskning og udvikling? Og kan familien være sikker på, at data, som nu kan henføres til dem selv som personer, ikke bliver brugt til andre formål, eksempelvis i forbindelse med forsikringer eller lignende?

Demokratiske processer og retssikkerhed

Der er uanede muligheder i at anvende data aktivt i både det offentlige og private. Gevinsterne skal naturligvis høstes, for det er sådan, vi udvikler vores samfund. Men netop fordi der er så mange potentielle gevinster, skal vi være meget opmærksomme på, *hvordan* vi arbejder med data på en etisk forsvarlig måde. Her er de demokratiske processer og retssikkerheden afgørende.

Data kan anvendes i rigtig mange sammenhænge. Med den offentlige sektors øgede digitalisering kan data eksempelvis bruges som forudsigende data eller til en helt automatiseret sagsbehandling. Et eksempel på sådanne systemer er eksempelvis forudsigende politiarbejde - det vil sige politiindsatser baseret på kunstig intelligens-forudsigelser af, om en person vil begå kriminalitet.

I udlandet ses en stigende tendens, hvor disse systemer tages i brug af blandt andet politiet, og også i Danmark er det for nyligt blevet aktuelt, da Rigspolitiet i foråret 2018 introducerede brugen af systemet POL-INTEL, hvor computeralgoritmer, kunstig intelligens og data fra blandt andet sociale medier bringes i

spil i efterforskningsarbejdet. POL-INTEL skulle oprindeligt også have en forudsigende funktionalitet, men Rigspolitiet har valgt ikke at benytte denne funktion på grund af den lave kriminalitetsrate, der er i Danmark. Det interessante her er dog, at et system som dette igangsættes uden større folkelig debat eller stillingtagen til de retssikkerhedsmæssige konsekvenser.

Det er afgørende, at der værnes om borgernes retssikkerhed i den digitale verden. Når mekanismer, der er af særlig indgribende karakter for den enkelte borger, indføres uden for meget stillingtagen til særligt de retssikkerhedsmæssige konsekvenser for den enkelte borger, affødes deraf tilsvarende en trussel mod gennemsigtigheden, retssikkerheden og ikke mindst tilliden til samfundet. For slet ikke at nævne truslen mod borgerens fundamentale frihedsrettigheder. Tillid går begge veje, hvorfor der ved indførelse af sådanne mekanismer tilsvarende bør indføres demokratiske processer og effektiv demokratisk kontrol.

Det er også værd at bemærke, at der er en vis fare ved at lægge for meget vægt på en algoritme og sætte sin sunde fornuft i baggrunden. Det bør ske med den største forsigtighed grundet algoritmers/datas kvalitet, herunder at systemer/data kan være fejlbehæftede, og såfremt de er det, hvem opdager det så og hvordan?

Det fremgår af GDPR, at *Data processing should be designed to serve mankind*¹ og ikke omvendt. Mistillid til systemet vil kunne give borgeren incitament til mindre deling af (korrekte og ikke mindst følsomme) data med offentlige myndigheder.

Forudsigende data og til dels også automatiserede afgørelser giver anledning til etiske og ikke mindst juridiske problemstillinger, som skal overvejes grundigt, inden de bliver en del af borgerens hverdag og vores samfund. Er det først sat i værk uden en grundig stillingtagen til disse problemstillinger, kan det have konsekvenser for tilliden og ikke mindst retssikkerheden for den enkelte borger.

Kontrolmekanismer

Vi står med en række grundlæggende spørgsmål, hvor svarene har stor betydning for fremtidens digitale Danmark: Hvem kontrollerer dataetikken i både det private og det offentlige, og hvem har mandat til at udtale sig om og træffe afgørelser i sager med større samfundsmæssige, dataetiske problemstillinger?

Datatilsynet: I Danmark er det Datatilsynet, der består af et råd, Datarådet, og et sekretariat, der er den myndighed, der fører tilsyn med, at reglerne i persondataloven overholdes. Der er tale om et kontrolorgan, der skal holde øje med, om virksomhederne og det offentlige overholder love og regler.

Datarådet træffer primært afgørelse i sager af principiel karakter og fastsætter selv sin forretningsorden². Datarådet behandler og træffer afgørelse i sager:

- 1) af principiel karakter
- 2) af betydelig almindelig interesse eller med betydelige følger for en offentlig myndighed eller privat virksomhed m.v.
- 3) der af andre særlige grunde findes at burde afgøres af rådet.
- 4) som et rådsmedlem ønsker optaget til rådsbehandling

Der er ikke tale om et organ, der direkte kontrollerer eller alene forholder sig til dataetiske problemstillinger.

Det Ethiske Råd: I Danmark findes der i dag "Det Ethiske Råd", men det blev i sin tid skabt for at skabe debat om og rådgive regeringen i etiske spørgsmål vedrørende sundhedsvæsenet og nye bio- og genteknologier, der berører menneskers liv, vores natur, miljø og fødevarer. I praksis er Det Ethiske Råd arbejdsmæssigt begrænset til at tage sig af netop sundhedsområdet og ikke det dataetiske område.

Dataetisk råd: Regeringen besluttede i efteråret 2018, at der skal nedsættes et dataetisk råd som en del af reformen Digital service i verdensklasse. DANSK IT bakker op om den idé og kan tilføje, at det er afgørende, at man sørger for, at det dataetiske råd bliver et bredt sammensat råd med mange forskellige profiler. Således skal rådet både kunne beskæftige sig med tekniske, etiske og juridiske dilemmaer samt træde i karakter der, hvor der i øvrigt måtte være et behov. Det dataetiske råds anbefalinger må ikke blot forblive anbefalinger, men skal også følges op af konkret handling andre steder i systemet.

Rådet for digital sikkerhed: Vi har også Rådet for digital sikkerhed, der hører under Erhvervsstyrelsen. Rådet er en uafhængig medlemsorganisation med højt specialiseret viden og erfaring om informationssikkerhed og privatlivsbeskyttelse. Rådet bringer denne viden i spil for at skabe en kvalificeret debat og en tryk digitalisering via bl.a. www.digitalsikkerhed.dk, men det er ikke et organ, der decideret kontrollerer eller tager stilling til dataetiske problemstillinger.

Regeringens dataetiske ekspertråd: I marts 2018 udnævnte daværende erhvervsminister Brian Mikkelsen en ekspertgruppe i dataetik. Ekspertgruppen skal ultimo 2018 komme med anbefalinger til regeringen om, hvordan danske virksomheder kan gøre ansvarlig dataanvendelse til en konkurrencefordel. Regeringens nuværende ekspertgruppe skal primært fokusere på, hvordan virksomheder kan håndtere de etiske udfordringer ved brug af data, men debatten er lige så vigtig for hele den offentlige sektor, så de teknologiske muligheder kan komme borgerne til gode på en dataetisk ansvarlig måde. Desuden er ekspertgruppen kun nedsat for 2018, men i de kommende år vil vi også løbende opleve, at data blive anvendt på nye måder.

Virksomhederne selv: Private virksomheder har i vid udstrækning også selv mulighed for at etablere forskellige former for interne kontrolmekanismer, der kan være med til at sikre et vedblivende fokus på dataetik på tværs af organisationen. Man kan eksempelvis etablere særlige samarbejdsudvalg, have konkrete politikker og også have dataetikken som en integreret del af CSR. På den måde er det meget lig, hvis en fysisk butik eksempelvis vælger at fjerne tobaksvarer ved kassen for på den måde at være med til at påvirke debatten om salg af tobak til unge.

Sammenfatning

DANSK IT's arbejdsgruppe for dataetik har siden foråret 2018 holdt en række møder, hvor omdrejningspunktet har været opgaven med at udpege de store dataetiske udfordringer i vores samfund og komme med anbefalinger til, hvordan vi kan sikre god dataetik i fremtiden.

Helt bevidst valgte DANSK IT fra starten, at det skulle være en bredt sammensat arbejdsgruppe. Da der sjældent findes rigtige og forkerte svar, når det kommer til dataetik, er det afgørende, at debatten om dataetik er nuanceret, og at alle synspunkter bliver hørt. Hvert medlem af arbejdsgruppen har således bidraget med viden og synspunkter ud fra sit it-faglige ståsted.

Arbejdet har ført til de dataetiske anbefalinger fra DANSK IT. Forhåbentlig vil anbefalingerne både blive anvendt som inspiration til konkrete handlinger hos myndigheder og virksomheder og som et udgangspunkt for at styrke debatten om dataetik.

Hvis du har kommentarer eller idéer efter at have læst denne publikation, er du altid velkommen til at kontakte DANSK IT.

DANSK IT
Bredgade 25 A
1260 København K

Kontaktperson:
Kim Stensdal, chef for kommunikation og viden.
Email: ks@dit.dk
Telefon: +45 22794373

Arbejdsgruppen for dataetik

Allan Bager, business development manager, Public Sector, IBM.

Alexander Holten Steffen, partner, Right People Group.

Amina Bazai, jurist.

Bjørn Borup, områdedirektør & CIO, IDA.

Esben Wolf, programleder, AU Uddannelse – Studiesystemer.

Gitte Mathiasson, COO, hoej.dk.

Henrik Friis, it-driftschef, Spæncom.

Lars Kristian Larsen, IT service management-konsulent, BusinessNow.

Lisbeth Neve, informationsarkitekt, Velliv.

Stig Jacob Drejer, master expert, Business Process Management.

Søren Skaarup, senior consultant, Rambøll Management Consulting.

Tobias Fonsmark, selvstændig, Semaphor.

Fra DANSK IT's sekretariat:

Rikke Hvilshøj, administrerende direktør.

Kim Stensdal, chef for kommunikation og viden.

Rasmus Hammer-Jakobsen, politisk konsulent.

Om DANSK IT

DANSK IT er en uafhængig interesseorganisation og et fællesskab for it-professionelle. Her kan du sætte fokus på din karriere gennem fortsat kompetenceudvikling og deltagelse i Danmarks største it-faglige netværk. DANSK IT arbejder for at fremme og understøtte it, hvor dette skaber værdi for samfundet og den enkelte. At samle, styrke og udvikle it-brugere og it-professionelles kompetencer og faglighed. Og på et uafhængigt grundlag varetage samfundets og medlemmernes it-interesser.

Definitioner

Data er information, der behandles og lagres digitalt. Her har vi tre grundlæggende datatyper: heltal, flydende tal (decimaltal) og tegn (fx bogstaver, mellemrum og specialtegn). Data kan opsamles én gang eller kontinuerligt, såsom motionsdata og søvndata. Data genereres gennem søgninger, handlinger, uploads, sensorer, internet of things m.v.

Datastrukturer er sammensatte datatyper, dannet ud fra grundlæggende datatyper.

Metadata eller "data om data" er fællesbetegnelse for mange forskellige typer af struktureret information, der bruges til at beskrive, administrere og genfinde samlinger af data.

Big data angiver både en størrelse af datamængder, et databehandlingsformål samt de teknikker og teknologier, der gør data-håndtering i en ekstrem skala økonomisk opnåeligt.

Pseudonymisering kan defineres som behandling af data på en sådan måde, at data ikke længere kan henføres til en bestemt person uden brug af supplerende oplysninger. Supplerende oplysninger skal opbevares separat og være underlagt tekniske og organisatoriske foranstaltninger for, at data ikke kan henføres til identificerbare personer.

Anonymisering af personoplysninger betyder, at ingen fysiske personer kan identificeres ud fra oplysningerne eller i kombination med andre oplysninger. Anonymiseringen skal være uigenkaldelig.

En personoplysning er enhver form for information, der kan henføres til en bestemt person, også selv om personen kun kan identificeres, hvis oplysningen kombineres med andre oplysninger. Personoplysninger kan for eksempel være personnumre, registreringsnumre, et billede, et fingeraftryk, en stemme, lægejournaler eller biologisk materiale. Man siger, at oplysningen er "personhenførbare".

Tre typer personoplysninger:

- Almindelige oplysninger kan være identifikationsoplysninger som navn og adresse eller oplysninger om økonomi, skat, gæld, væsentlige sociale problemer, andre rent private forhold, sygedage, tjenstlige forhold, familieforhold, bolig, bil, eksamen, ansøgning, CV, ansættelsesdato og -stilling, arbejdsområde og arbejdstelefon.
- Særlige kategorier af oplysninger, er følsomme oplysninger om race og etnisk oprindelse, politisk overbevisning, religiøs eller filosofisk overbevisning, fagforeningsmæssige tilhørsforhold, genetiske data, biometriske data med henblik på entydig identifikation, helbredsoplysninger, seksuelle forhold eller seksuel orientering.
- Oplysninger om straffedomme og lovovertrædelser eller tilknyttede sikkerhedsforanstaltninger.

Dataportabilitet giver en person mulighed for at overføre og videreanvende egne oplysninger til egne formål og på tværs af forskellige tjenester. Man har ret til at kunne modtage egne personoplysninger i et struktureret, almindeligt anvendt og maskinlæsbart format. I forbindelse med dataportabilitet skal det være nemt for personen at flytte, kopiere eller transmittere personoplysninger fra ét it-miljø til et andet samt enkelt selv at administrere egne personoplysninger, for eksempelvis at tjene penge på data, donere dem til forskning eller bruge dem til gavn for andet. Fælles standarder og formater på tværs af sektorer og brancher kan være nødvendige for at kunne leve op til kravene med hensyn til retten til dataportabilitet.

Inspiration, litteratur og kildehenvisninger

Tænkehandletanken DataEthics.eu:

”Dataetik er en konstruktiv, handlingsorienteret og menneske-centrisk tilgang, hvor man kigger på alle processer fra et interdisciplinært perspektiv, der inkluderer teknologi, lovgivning samt sociale og kulturelle aspekter.”

Europarådets Bioetikkonvention af 1997, hvor menneskets forrang fastslås i artikel 2: ”Menneskets interesser og velfærd skal have forrang frem for interesser, der alene vedrører samfundet og videnskaben.”

EU’s charter om grundlæggende rettigheder af 2000, hvori det i artikel 8 slås fast, at ”Enhver har ret til beskyttelse af personoplysninger, der vedrører ham/hende”, og at ”Disse oplysninger skal behandles rimeligt, til udtrykkeligt angivne formål og på grundlag af de berørte personers samtykke eller på andet ved lov fastsat grundlag. Enhver har ret til adgang til indsamlede oplysninger, der vedrører ham/hende, og til berigtigelse heraf.”

Ethics Advisory Group report 2018 – European Data Protection Supervisor.

Eksponeret – Grænser for privatliv i en digital tid, Gads Forlag, 2018.

Weapons of Math Destruction, Cathy O’neil, 2016.

Dataetik - den nye konkurrencefordel, Hasselbalch og Tranberg.

How to fix the future, Andrew Keen, 2018.

Algoritmer er nutidens masseødelæggelsesvåben, Peter Hesseldahl, 24. juni 2018, Mandag Morgen.

¹ <https://www.i-scoop.eu/gdprrecital/gdpr-recital-4/>

² <http://www.justitsministeriet.dk/ministeriet/raad/dataraadet>