Dataetisk Råd

DATAETISKE PERSPEKTIVER

AI på socialområdet



Dataetisk Råd

København 2026

Dataetisk Råd er en offentlig myndighed, som rådgiver om etiske udfordringer ved nye digitale teknologier. Læs mere på **dataetiskraad.dk**

Projektet bag denne rapport har i rådet været forankret i en referencegruppe bestående af:

- Anette Høyrup
- Birgitte Arent Eiriksson
- Johan Busse
- Martin Østergaard
- Rikke Frank Jørgensen

Projektet er udarbejdet med sekretariatsbistand fra projektleder Francisca Nordfalk.

Indhold

Formandens forord 4

Kapitel 1: Om rådsprojektet **5**

AI på socialområdet 5

Dataetisk råds tilgang til projektet 6

Dataetisk Råds principper 7

Dataetik og jura på socialområdet 8

Kapitel 2: Dataetisk analyse af anvendelsen af AI på socialområdet **9**

Fem centrale fund fra praksis 10

De seks brugsscenarier 11

Brugsscenarie 1: Informationsassistenter til medarbejdere 12

Brugsscenarie 2: Digitale receptionister 16

Brugsscenarie 3: Automatisk referatskrivning 20

Brugsscenarie 4: Personlige AI-agenter 25

Brugsscenarie 5: Intelligente beslutningsstøttende værktøjer 30

Brugsscenarie 6: Uofficiel anvendelse af AI 35

Kapitel 3: Anbefalinger **39**

Dataetisk Råds seks anbefalinger til anvendelsen af AI på socialområdet 40

Om Dataetisk Råd 41

Formandens forord

Anvendelsen af AI (kunstig intelligens) er allerede nu en del af den offentlige sektors hverdag. Også på socialområdet er forventningerne til teknologien store, både når det gælder understøttelse af medarbejdere, effektivisering af arbejdsgange og forbedret adgang til information og service for borgerne. Samtidig rejser anvendelsen af AI-teknologier på socialområdet nogle grundlæggende dataetiske spørgsmål om, hvordan vi som samfund ønsker at møde og støtte de mennesker, der er i kontakt med socialforvaltningen.

Socialområdet er præget af komplekse faglige vurderinger, betydelige sociale forskelle og behovet for at tage individuelle og menneskelige hensyn. Når AI bliver en del af dette arbejde, giver det i særlig grad anledning til overvejelser om værdier som retssikkerhed, værdighed og ansvarlighed, og hvordan vi sikrer, at anvendelsen af AI-teknologier sker på en måde, der styrker tilliden til socialforvaltningen og til de offentlige myndigheder.

Dataetisk Råd har derfor i 2026 gennemført et rådsprojekt om de dataetiske overvejelser ved implementering af AI på socialområdet. Formålet har været at skabe et fælles grundlag for debat og refleksion over de muligheder og faldgruber, der følger med teknologiens udbredelse.

I dette projekt har rådet arbejdet praksisnært med udgangspunkt i seks forskellige anvendelser af AI på socialområdet. Gennem dataetiske analyser har vi undersøgt, hvilke hensyn og principper der kan blive styrket, og hvilke der kan blive udfordret, når AI bliver en del af socialforvaltningens arbejde. Det har samtidig været vigtigt for os at inddrage perspektiver fra de aktører, som arbejder med, eller berøres af AI på socialområdet. Deres erfaringer, perspektiver og kritiske refleksioner har været afgørende for projektets analyser og anbefalinger.

Dataetisk Råds ambition med projekter er ikke at give entydige svar på, om AI bør anvendes eller ej. Svaret på det, afhænger helt af om lovgiver og myndigheder får set sig for, tænkt sig om og håndteret de potentielle skadevirkninger af den konkrete anvendelse af teknologien. Ambitionen er at kvalificere samtalen om, hvordan teknologien kan anvendes ansvarligt. Det afgørende er, at vi løbende drøfter, hvilke værdier der skal være styrende for anvendelsen af AI-teknologier i mødet mellem borgere og samfundet. På vegne af Dataetisk Råd håber jeg, at rapporten kan bidrage med inspiration, refleksion og konkrete perspektiver til alle, der arbejder med eller interesserer sig for AI på socialområdet.

Tak til alle, der har bidraget til projektet. Jeres erfaringer og indsigter har været afgørende.



Johan Busse, formand for Dataetisk Råd.

Kapitel 1
Om rådsprojektet

AI på socialområdet

I denne rapport behandler Dataetisk Råd emnet "AI på socialområdet". Det giver anledning til at klargøre, hvilken forståelse af "AI" og "socialområdet", rådet har arbejdet ud fra.

Dataetisk Råd har tidligere skrevet mere fyldestgørende om AI og generativ AI i rapporten "Generativ AI og offentlige myndigheder".[1] Nærværende rapport om anvendelsen af AI på socialområdet tager udgangspunkt i samme forståelse af AI og generativ AI:

AI (kunstig intelligens)

Et computersystem, der løser en kompleks opgave, det ellers ville kræve menneskelig intelligens at løse.

Generativ AI

En gren af AI, der omhandler design og implementering af algoritmer, som kan generere nyt indhold. Dette indhold kan variere fra tekst, billeder, musik, til video, der ofte er af en sådan kvalitet, at det kunne antages at være skabt af mennesker.

Generativ AI adskiller sig fra traditionel AI, der typisk fokuserer på at analysere og fortolke data, ved i stedet at skabe nyt og originalt indhold baseret på de data, den er trænet på.

I kapitel 2 gennemgås seks forskellige former for anvendelse af AI på socialområdet. I dette nævnes også, hvilken form for AI-teknologi, anvendelsen bygger på. Da vores fokus ikke er på selve teknologien, men på effekterne af anvendelse, skrives betegnelsen "AI" og "AI-teknologi" uanset hvilken form for AI der anvendes.

Der findes ikke en fælles national definition for, hvad der falder ind under "socialområdet". Forståelsen af socialområdet for dette projekt beskrives derfor nærmere i følgende afsnit.

[1] Dataetisk Råd: "Generativ AI og offentlige myndigheder: Pålidelighed og dataetik" (2025). Tilgængelighed: dataetisk-raads-rapport-om-generativ-ai.pdf

Dataetisk råds tilgang til projektet

Med dette rådsprojekt og denne rapport ønsker Dataetisk Råd at kvalificere samtalen om de mange (dataetiske) overvejelser, der opstår, når AI bliver anvendt på socialområdet. I dette projekt har Dataetisk Råd således arbejdet ud fra en overordnet forståelse af socialområdet som "*Den del af samfundet der beskæftiger sig med social forsorg og socialt betingede økonomiske ydelser*".[2] I arbejdet med de dataetiske overvejelser på socialområdet har den kommunale socialforvaltning været rådets praktiske indgang til og forståelse af socialområdet. Her arbejder rådet ud fra en forståelse af socialforvaltningen som "*Kommunal administrativ enhed som forvalter og træffer afgørelse i sager, der hører ind under det socialpolitiske område*".[3]

Dataetisk Råd taler således ind i en større samtale om, hvordan vi som samfund ønsker at anvende AI i kontakten med samfundets borgere. Anvendelsen af AI på socialområdet bliver både omtalt som et område; med høj fart på udviklingen (det går for stærkt) og et område med et stort uforløst potentiale (det går for langsomt).

I regeringsgrundlaget fra juni 2026 ønsker regeringen en "strategisk og balanceret tilgang til kunstig intelligens",[4] hvilket rådet med denne rapport ønsker at bidrage til. Ligeledes har Regeringen, KL og Danske Regioner sat en retning for anvendelsen af AI i den offentlige sektor med tre storskalaprojekter, som forventeligt også vil have en effekt på socialområdet. Samtidig er socialområdet meget borgernært, hvilket ikke altid harmonerer med anvendelsen af AI-teknologier.[5] Anvendelse af AI-teknologier på netop socialområdet bør derfor især give anledning til nogle etiske overvejelser.

I dette rådsprojekt om AI på socialområdet, har rådet valgt en praksisnær tilgang, som inddrager flere perspektiver. Projektet tager udgangspunkt i seks brugsscenarier, som eksemplificerer anvendelsen af AI på socialområdet. Hver af de seks brugsscenarier præsenteres med en dataetisk analyse der gennemgår både argumenter for og imod anvendelsen. Hver anvendelse, eller brugsscenarie afsluttes med fem dataetiske råd til en ansvarlig anvendelse.

Samtidig er rådet ikke eksperter på socialområdet. Derfor er vores analyser og overvejelser kvalificeret gennem samtaler med relevante aktører på området. Her har vi hørt og drøftet, hvordan anvendelsen af AI på socialområdet giver anledning til mange overvejelser ud fra

[2] Se også "socialsektor" i Den Danske Ordbog: socialsektor — Den Danske Ordbog

[3] Se "socialforvaltningen" i Den Danske Ordbog: socialforvaltning — Den Danske Ordbog

[4] Se punkt 12 i " Det politiske grundlag for firkløverregeringen"
Det politiske grundlag for firkløverregeringen - Statsministeriet

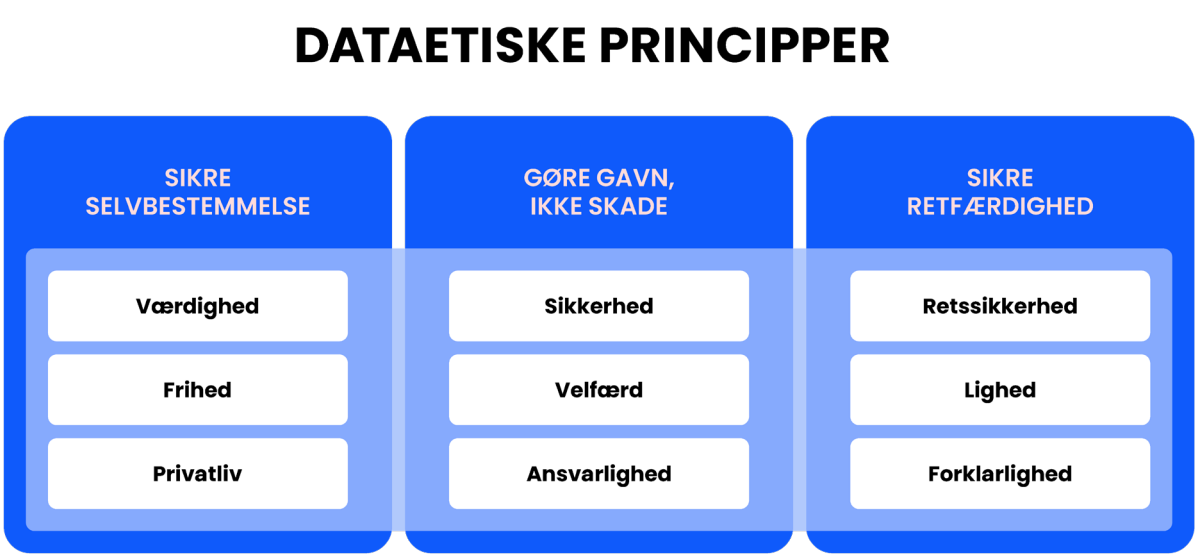
[5] Se fx "Borgerforslag kræver ret til menneskelig sagsbehandling - TV 2"

de forskellige aktørers perspektiver.

Derfor vil vi gerne takke Aarhus Kommune, SAND - De hjemløses Organisation, Socialrådgiverforeningen, Danske Handicaporganisationer og KOMBIT for at bidrage til vores projekt med relevante og uundværlige perspektiver og overvejelser.

Til sidste i denne rapport præsenterer Dataetisk Råds seks anbefalinger til anvendelsen af AI på socialområdet.

Dataetisk Råds principper



Dataetisk Råd har formuleret tolv principper for ansvarlig brug af data og teknologier – herunder AI. Her arbejder rådet med tre overordnede og ni underordnede dataetiske principper. Når man arbejder med dataetik, bliver det ofte lettere at navigere i dilemmaer og afvejninger, hvis man har principperne som reference. Tilsammen dækker de et bredt udsnit af de hensyn, som typisk er relevante i en dataetisk problemstilling.

I denne rapport præsenterer vi en nærmere analyse af de principper, som især har vist sig at være relevante i de brugsscenarier, som projektet tager udgangspunkt i. Her kan Dataetisk Råds principper danne grundlag for en velovervejet drøftelse af de etiske udfordringer og overvejelser, der vil opstå, når AI-teknologier bliver anvendt på socialområdet. I analysen er de anvendte dataetiske principper markeret med *kursiv*.

Man kan læse mere om alle de enkelte principper og om, hvordan Dataetisk Råd arbejder med dem, på Dataetisk Råds hjemmeside www.dataetiskraad.dk

Dataetik og jura på socialområdet

Dataetik og jura har flere overlappende felter, da begge fagområder beskæftiger sig med ansvarlig håndtering af data og teknologier. Men hvor juraen omhandler den lovgivning og de regler, der regulerer behandlingen af data, fokuserer dataetikken i højere grad på principper og værdier, der bør være styrende for tilgangen til data og teknologier. Man kan altså godt overholde alle juridiske regler og stadig gøre sig relevante overvejelser om de dataetiske implikationer af et AI-projekt. Omvendt kan man ofte også anvende dataetikken til at forstå en juridisk lovgivning. Lidt populært kan man sige, at teknologien svarer på, om man 'kan', at juraen svarer på, om man 'må', og at dataetikken svarer på, om man så 'bør'. I rådets arbejde med at belyse dataetiske principper og hensyn arbejder vi altid ud fra en forventning om, at gældende lovgivning overholdes. Derfra ser vi på de dataetiske perspektiver. Uafhængigt af, hvilket område AI-teknologier anvendes på, er det essentielt at gøre sig bekendt med især Databeskyttelsesforordningen og AI Act:

Databeskyttelsesforordningen (GDPR) er en EU-forordning, der trådte i kraft i 2016, og som beskytter persondata og privatlivets fred for borgere i EU. Den er i Danmark suppleret med databeskyttelsesloven. Tilsammen fastsætter de regler for behandling af personoplysninger og sikrer borgernes rettigheder til egne data.

AI Act er en EU-forordning, der regulerer brugen af AI og trådte i kraft i 2025. Forordningen sigter mod at fremme ansvarlig udvikling og anvendelse af AI. Forordningen introducerer et ensartet rammeværk og klassificerer AI-systemer efter deres potentielle risici. En række nærmere bestemte former for AI-praksis kategoriseres som forbudte inden for EU's grænser, mens andre kategoriseres som højrisiko-AI og derfor er underlagt særlige krav.

Derudover er socialområdet omfattet af flere andre regelsæt, som også sætter rammerne for anvendelsen af AI-teknologier, herunder **serviceloven, barnets lov, forvaltningsloven, retssikkerhedsloven, offentlighedsloven** og **menneskerettighederne**.

Kapitel 2
Dataetisk analyse af anvendelsen af AI på socialområdet

I dette kapitel gennemgår vi de dataetiske analyser for anvendelsen af AI på socialområdet. Det dækker indledningsvist over fem centrale fund, som går på tværs af vores møder med aktører på området og derefter en dataetisk analyse af hver af de seks brugsscenarier i projektet.

Fem centrale fund fra praksis

I vores møder med nogle af de aktører, som bliver berørt af anvendelsen af AI på socialområdet, fik vi mange værdifulde perspektiver og indblik i, hvordan AI-teknologier opfattes, og hvilke overvejelser de forskellige aktører finder relevante og væsentlige. Selvom flere af deres bidrag var specifikke for den enkelte aktør, var der også nogle tværgående perspektiver.



Fem centrale fund om AI på socialområdet



1. AI-teknologi er baseret på statistik, gennemsnit og normer. Omvendt befinder borgere i socialforvaltningen sig ofte i komplekse livssituationer, som vanskeligt passer ind i generelle kategorier og standardløsninger.



2. Datakvaliteten på socialområdet er ofte lav og der mangler fælles standarder for registreringer og ensartet data.



3. Det sociale arbejde er ofte relationelt og personligt og kan ikke nødvendigvis omsættes til data.



4. Borgernes teknologiske ressourcer kan være høje eller lave og tilgængelige eller fraværende på flere måder.



5. Ansatte på socialområdet har forskellige teknologiske forudsætninger og det vil kræve uddannelse og rammer for ansvarlig anvendelse af AI.

Disse fund samt andre input fra vores møder med aktørerne har formet vores analyser og rådets endelige anbefalinger. Samtidig mener rådet, at de fem fund er relevante for alle, der arbejder med AI på socialområdet, og vi håber, at de kan være med til at sætte en ramme for samtalerne omkring udviklingen og anvendelsen af AI-teknologier.

De seks brugsscenarier

I det følgende præsenteres og analyseres hver brugsscenarie enkeltvis. I analysen belyses anvendelsen af AI på socialområdet især ud fra, hvilke påvirkninger den vil have på det enkelte menneske, herunder borgeren og medarbejderen, og på samfundet generelt. Alle de dataetiske perspektiver er belyst ud fra en forudsætning om, at teknologien generelt virker, som den er intenderet.

For udvælgelsen af brugsscenarier har rådet opsat fire kriterier:

1. AI-teknologien skal anvendes på socialområdet (men ikke nødvendigvis udelukkende der)
2. Anvendelsen skal være lovlig inden for gældende dansk og europæisk lovgivning
3. Anvendelserne skal være klart afgrænsede i forhold til hinanden
4. Anvendelsen skal være relevant for praksis i dag eller i den nærmeste fremtid



Samlet har rådet derfra valgt seks forskellige brugsscenarier. Analysen af hver brugsscenarie afsluttes med fem dataetiske råd til en ansvarlig anvendelse af den specifikke AI-teknologi. Det er således muligt at læse de dataetiske perspektiver på de enkelte brugsscenarier. I det afsluttende kapitel 3, giver Dataetisk Råd deres seks anbefalinger til anvendelsen af AI på socialområdet på tværs af de seks brugsscenarier.



Brugsscenarie 1
Informationsassistenter til medarbejdere

Ansatte på socialområdet skal kunne navigere i store mængder af lovtekst, retningslinjer og interne dokumenter. Derfor anvendes digitale AI-informationsassistenter til hurtigt at finde de relevante oplysninger i afgrænsede kilder. Teknologien fungerer således, at medarbejderen stiller spørgsmål i et naturligt sprog og informationsassistenten leverer et samlet overblik, som kan understøtte den videre opgaveløsning. Det kan eksempelvis være at finde relevant information på intranettet eller søge efter relevante lovtekster. Formålet er at spare tid på informationssøgning og sikre medarbejderen et hurtigt overblik over relevant viden.

Hvem anvender teknologien: Anvendes af medarbejdere i socialforvaltningen som interne værktøjer, særligt inden for administration og sagsbehandling, og bruges primært til informationssøgning og opgavestøtte.

Hvilken AI-teknologi: Løsningerne er baseret på generativ AI, som kan besvare spørgsmål ved at trække på relevante datakilder. Teknologien kombineres typisk med Retrieval Augmented Generation (RAG), hvor systemet søger i interne, forhåndsgodkendte dokumenter og anvender disse som grundlag for svarene. Dette skal reducere fejlkilder og sikre, at sagsbehandlingens informationer baseres på eksisterende og godkendt materiale. AI-informationsassistenterne skal derfor fungere som støtte til informationssøgning og vidensoverblik.

Eksempler:

- Bertha i Vejle kommune som kan bruges til intranettet og korrekturlæsning.
- Kommunernes Sygedagpengesystem (KSD) assisterer i gældende lovgivning og systembrug.
- HolgerBæk i Holbæk kommune med information, retningslinjer, procedurer og øvrige ressourcer tilgængelige på intranettet.

Dataetiske overvejelser ved anvendelsen af informations-assistenter til medarbejdere

Anvendelsen af AI-baserede informationsassistenter rummer et stort potentiale til at effektivisere informationssøgning og give medarbejdere et hurtigt og samlet overblik over relevant viden. Samtidig giver anvendelsen af AI informationsassistenter anledning til nogle dataetiske overvejelser. I det følgende gennemgås de dataetiske argumenter der taler for og imod anvendelsen af informationsassistenter til medarbejdere på det sociale område.

Argumenter og hensyn, som taler for anvendelsen af informationsassistenter til medarbejdere

Hvis informationsassistenten er baseret på afgrænsede og forhåndsgodkendte dokumenter, og anvender tydelige kildehenvisninger, kan det give medarbejderen en mulighed for nemmere at henvise til, hvilken information eller hvilket regelsæt, de har anvendt. AI-assistenten kan derved bidrage til at gøre komplekse informationssøgninger mere tilgængelige og forståelige for medarbejdere, og det kan derfor øge gennemsigtigheden og *forklarligheden* i sagsbehandlingen.

Det kan ligeledes øge *retssikkerheden*, hvis relevant lovgivning og retningslinjer er lettere tilgængelige for medarbejderne. Hurtigere adgang til korrekt regelgrundlag kan mindske risikoen for menneskelige fejl, som i sagsbehandlingen kan udgøre en risiko for borgeres retslige rettigheder.

Hvis informationsassistenter til medarbejdere bliver implementeret ens og bredt, kan det reducere variation i sagsbehandling på socialområdet. Ensartet adgang til lovgivning, retningslinjer og interne procedurer kan mindske betydningen af personlige og lokale vidensforskelle og kan potentielt understøtte mere *lighed* og *retfærdighed* for borgerne i sagsbehandling på tværs af medarbejdere, enheder og kommuner.

AI-baserede informationsassistenter kan altså *gøre gavn* for både borgere og medarbejdere i socialforvaltningen. Mere effektiv adgang til korrekt viden kan bidrage til en mere velfungerende socialforvaltning og dermed gavne samfundets borgere samlet set.

Argumenter og hensyn, som taler imod anvendelsen af informationsassistenter til medarbejdere

Omvendt udfordrer anvendelsen af informationsassistenter til medarbejdere også nogle centrale dataetiske principper.

På socialområdet har sagsbehandlingen ofte stor betydning for borgerens livssituation, derfor kan fejl, udeladelser eller misvisende opsummeringer få alvorlige konsekvenser. Selvom informationsassistenten ikke træffer afgørelser, kan den indirekte påvirke sagsbehandlingen ved at fremhæve bestemte fortolkninger eller opsummeringer af information, regelsæt eller lovgivning. *Retssikkerheden* afhænger derfor også af, om medarbejderen reelt er i stand til at gennemskue og efterprøve assistentens svar. Hvis det er uklart, hvordan AI-assistenten er kommet frem til et svar, kan det blive vanskeligere at kontrollere, om lovgrundlaget er korrekt fortolket og anvendt.

For at teknologien kan anvendes forsvarligt, skal det være muligt for medarbejderen at forstå, hvor informationsassistentens svar kommer fra, herunder hvilke dokumenter der er anvendt, og hvilke antagelser der ligger til grund for svaret. Det skal sikre *forklarlighed* i sagsbehandlingen for både medarbejderen og borgeren. Hvis medarbejderen ikke har indsigt i systemets virke og begrænsninger, herunder risiko for udeladelser, forældet viden eller forenklede fortolkninger, kan det svække den faglige dømmekraft og øge risikoen for ukritisk brug. Det kan således blive vanskeligt for medarbejdere at forklare, og for borgeren at forstå, hvilken rolle AI har spillet i informationsgrundlaget for sagsbehandlingen. Det kan være skadeligt for både borgerens sag og for den generelle gennemsigtighed i den offentlige opgaveløsning.

Nye teknologier bliver sjældent implementeret ensartet alle steder samtidig. Informationsassistenter vil derfor anvendes forskelligt på socialområdet og det kan føre til variationer i sagsbehandlingen på tværs af borgere, enheder og kommuner. Hvis informationsassistenter får indirekte betydning for, hvordan borgere behandles på socialområdet, kan det potentielt påvirke *lighedsprincippet*. Samtidig er der en risiko for, at AI teknologien reproducerer eller forstærker eksisterende bias. Hvis bestemte perspektiver eller målgrupper systematisk vægtes højere end andre, kan det have betydning for, hvordan borgernes situation forstås og håndteres i praksis. Konsekvensen kan være, at visse borgere, eller bestemte typer af sager, indirekte behandles anderledes, hvilket kan føre til diskrimination og dermed udfordre principperne om *lighed* og *retfærdighed*.

Hvis det er muligt for medarbejdere at indtaste borgerdata, som informationsassistenten kommer til at behandle, kan det gå ud over borgeres ret til *privatliv*. Socialområdet håndterer ofte meget følsomme oplysninger om borgerne, og selv begrænset eller indirekte brug af sådanne data kan indebære væsentlige privatlivsrisici. Hvis borgerdata anvendes som input til systemet, eksempelvis i forbindelse med opsummering eller søgning relateret til en

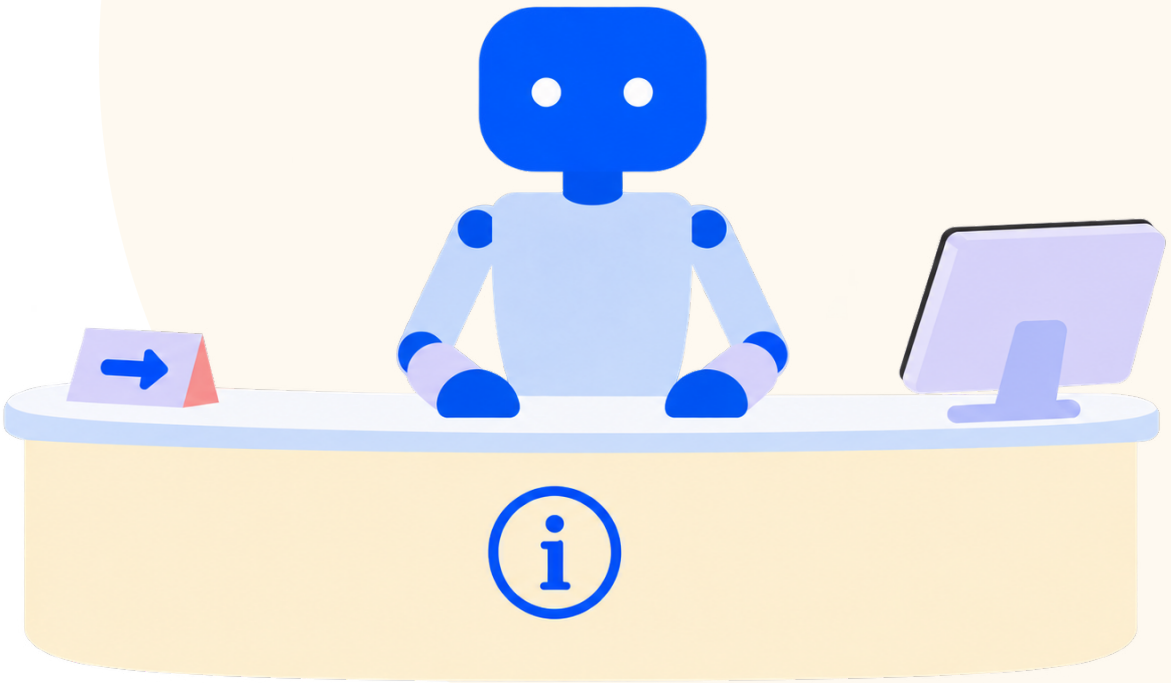
konkret sag, kan det udgøre en risiko, hvor borgeroplysninger behandles i systemer, som ikke er beregnet til det. Samtidig kan der være *sikkerhedsrisiko* for utilsigtet videregivelse eller genbrug af oplysninger, hvis systemets afgrænsning og adgangsstyring ikke er tilstrækkelig udviklet.

Anvendelsen af AI baserede informationsassistenter kan altså få betydning for borgerens sag, selvom systemet formelt kun fungerer som et støtteværktøj. For borgeren er det afgørende, at anvendelsen af teknologien ikke skaber uklarhed om, hvem der træffer beslutninger, og hvem der kan stilles til *ansvar* for disse beslutninger. *Ansvarlighed* forudsætter, at myndigheden har klare procedurer for brugen af informationsassistenter, og kan redegøre for deres rolle i sagsbehandlingen. Ved systematisk anvendelse af AI, kan der opstå et misforhold mellem det formelle ansvar hos mennesker og den reelle indflydelse, som systemet har haft på resultatet. Anvendelsen af AI kan altså skabe uklarhed om ansvarskæden, især hvis der opstår fejl i sagsbehandlingen.

Samlet set kan anvendelsen af AI-informationsassistenter til medarbejderen altså både *gøre gavn* og risikere at *gøre skade* for både borgere og medarbejdere i socialforvaltningen.

Fem dataetiske råd til en ansvarlig anvendelse af informationsassistenter:

1. Informationsassistenten skal være baseret på afgrænsede og forhåndsgodkendte dokumenter og anvende tydelige kildehenvisninger.
2. Der skal være organisatoriske rammer, som sikrer, at medarbejderne er tilstrækkeligt uddannede til at anvende og vurdere systemets svar.
3. Teknologien skal søges at blive implementeret ens på tværs af enheder og kommuner og med fælles vejledning i, hvordan informationsassistenten må indgå i sagsforberedelsen.
4. Systemets afgrænsning og adgangsstyring skal være klar og med præcise regler for, om eller hvordan borgerdata må indgå i systemet.
5. Der skal være tydeligt, hvem der har ansvar for informationsassistenten og hvem der kan stilles til ansvar for de beslutninger, som teknologien er en del af.



Brugsscenarie 2
Digitale receptionister

Borgere i socialforvaltningen har ofte behov for hjælp til at finde relevant information, navigere i kommunale tilbud og få svar på standardiserede spørgsmål. Derfor håndterer medarbejderne et stort antal gentagne henvendelser på tværs af kanaler og fysisk fremmøde. En løsning på dette er anvendelse af AI baserede digitale receptionister til mødet med borgeren. Disse findes allerede i flere kommuner i forskellige formater, herunder som online chatbots og fysiske skærme.

Hvem anvender den: Anvendes af borgere i mødet med kommunen. Anvendes især online, men findes også tilgængelig ved fysisk fremmøde eller ved telefonisk henvendelse. De digitale receptionister bliver brugt til at søge hjælp eller information og ved spørgsmål om service og selvbetjening.

Hvilken AI-teknologi: Digitale receptionister anvender generativ AI kombineret med teknologier, der genkender borgernes spørgsmål og matcher dem med eksisterende kommunal information. Mange af løsningerne bygger på tekst-til-tekst, tale-til-tale, tale-til-tekst og tekst-til-tale, så borgere kan interagere både skriftligt og mundtligt. Teknologien gør det muligt at give borgerrettede svar baseret på godkendte data. I enkelte tilfælde indgår også visuel "avatar"-teknologi, hvor en digital figur repræsenterer receptionisten.

Eksempler:

- Trothilde i Jobcenter Aarhus, som hjælper borgere med at finde vej, besvare spørgsmål og give praktisk information.
- MUNI - Fælles kommunal chatbot er en AI-drevet chat- og voicebot, der hjælper borgere med kommunale services via hjemmesider og telefon.
- Kommune Kiri bruges i borgerservice i flere kommuner til at hjælpe borgere med deres henvendelser ved hjælp af en AI-drevet chat- og voicebot, som man både kan skrive og tale med.

Dataetiske overvejelser ved anvendelsen af digitale receptionister

Borgeres møde med socialområdet kan være afgørende for deres tilværelse. Derfor skal borgeren her mødes på en måde, som er meningsfuld for dem og som gør det muligt for ansatte på socialområdet at bruge deres ressourcer bedst muligt. I det følgende gennemgås de dataetiske argumenter, der taler for og imod anvendelsen af digitale receptionister på det sociale område.

Argumenter og hensyn, som taler for anvendelsen af digitale receptionister

Kontakt med det sociale område kan have karakter af både generel informationssøgning og spørgsmål af mere personlig karakter. I begge tilfælde kan der være borgere, som oplever det mere trygt eller potentielt mindre stigmatiserende at stille spørgsmål til en teknologi frem for et menneske. Derfor kan digitale receptionister understøtte borgerens oplevelse af *værdighed, privatliv* og kontrol over egen situation.

Samtidig kan digitale receptioner, især online, også øge *ligheden* i adgangen til information på socialområdet. Både ved at sikre, at borgere, der har svært ved at møde fysisk op eller anvende åbningstider, kan få svar døgnet rundt og uden ventetid. Tale-til-tekst eller tekst-til-tale kan samtidig forbedre adgangen til information for nogle borgere med forskellige behov og forudsætninger. Mere ensartede svar på standardiserede spørgsmål kan også reducere vilkårlig forskelsbehandling mellem borgere, som henvender sig på forskellige tidspunkter eller til forskellige medarbejdere.

Digitale receptionister kan derfor *gøre gavn* for både borgere og medarbejdere. Øget tilgængelighed kan gavne borgere med akutte simple informationsbehov. Medarbejdere kan frigøres fra rutineprægede spørgsmål og i stedet frigive tid til at fokusere på mere komplekse og borgernære opgaver, hvor deres faglige kompetencer skaber størst værdi.

Argumenter og hensyn, som taler imod anvendelsen af digitale receptionister

Det er helt afgørende, at det er tydeligt for borgeren, at det i mødet med den digitale receptionist netop interagerer med en AI-teknologi, uanset hvordan teknologien kommer til udtryk. Det skal altså sikres, at der er gennemsigtighed og *forklarlighed* overfor borgeren. Ellers kan borgeren fejlagtigt opfatte den information de får, som individuelle vurderinger eller vejledning fra et menneske, som kan tage højde for borgerens individuelle situation eller kontekst. Hvis den digitale receptionist gives et meget menneskelignende udtryk, kan der opstå en risiko for, at borgere tillægger den digitale receptionist egenskaber som forståelse, empati eller dømmekraft, som AI-teknologier ikke besidder. Anvendelsen af visuelle avatarer rejser her særlige forklarlighedsudfordringer, også i fremtiden. Der bør altid anvendes udtryk, som tydeligt signalerer, at borgeren kommunikerer med en teknologisk løsning og ikke et menneske.

Det kan også ske, at borgere deler personlige og følsomme oplysninger i interaktionen med den digitale receptionist. Det kan udfordre borgerens ret til *privatliv*. Der er en risiko for, at borgere deler mere viden om dem selv og deres situation, end der er nødvendigt for at opnå den information, som den digitale receptionist kan give. Ligeledes kan borgeren have vanskeligt ved at gennemskue, hvordan deres oplysninger behandles, opbevares og eventuelt anvendes videre, særligt hvis dialogen opleves som uformel eller menneskelig.

Ligesom anvendelsen af digitale receptionister kan øge *ligheden*, kan det også udfordre *lighedsprincippet*. De digitale kompetencer og ressourcer kan variere meget mellem borgere i socialforvaltningen, hvilket også gælder for forudsætningerne for at anvende chat og talebaserede løsninger. Hvis digitale receptionister i praksis bliver den primære indgang til kommunen, kan det føre til ulighed i adgang til hjælp og information. Derudover er der risiko for bias i AI teknologiens svar, som kan opstå gennem skævheder i det informationsgrundlag systemet bygger på, eller i måden borgerens input fortolkes på. Hvis bestemte grupper, sprogformer eller problemstillinger systematisk håndteres dårligere, kan det udfordre *lighedsprincippet*.

Derudover kan borgere i sårbare eller komplekse livssituationer opleve nedsat *velfærd*, hvis deres behov ikke bliver identificeret eller forstået tilstrækkeligt af en digital receptionist. Det kan føre til frustration, afmagt eller utryghed, hvis digital kontakt erstatter menneskelig kontakt i situationer, hvor borgeren har behov for støtte, omsorg eller nuanceret vejledning.

De digitale receptionister repræsenterer socialforvaltningen i mødet med borgeren og de skal derfor kunne sikre *ansvarlighed* i mødet. Der kan opstå uklarhed om, hvem der har ansvaret, herunder om det ligger hos kommunen, leverandøren, systemudvikleren eller specifikke medarbejdere. Borgere kan fejlagtigt opfatte den digitale receptionist som en autoritativ og ansvarshavende repræsentant for kommunen. Det skal derfor være tydeligt for både medarbejdere og borgere, hvem der har ansvaret for de svar og den vejledning,

systemet giver. Det skaber behov for viden og klare rammer for tilsyn, kontrol og løbende kvalitetssikring af teknologien. For borgeren er det afgørende, at anvendelsen af en digital receptionist ikke skaber uklarhed om, hvem der kan stilles til ansvar, eller hvordan de kommer videre, hvis de ikke får den nødvendige hjælp.

Samlet set kan digitale receptionister derfor både *gøre gavn* og *gøre skade*. Borgere kan lide skade, hvis mødet med den digitale receptionist er uklart eller giver forkerte, misvisende eller ufuldstændige oplysninger, som fører til uhensigtsmæssige handlinger eller manglende adgang til hjælp. En overdreven offentlig brug af digitale receptionister kan også reducere den menneskelige kontakt i borgerservice, hvilket kan være skadeligt for udsatte borgere.

Fem dataetiske råd til en ansvarlig anvendelse af digitale receptionister:

1. Det skal være helt tydeligt for borgeren, at de i mødet med den digitale receptionist netop interagerer med en AI-teknologi, uanset hvordan teknologien kommer til udtryk.
2. Teknologien i den digitale receptionist skal have klare rammer for databehandling og begrænsende tekniske samt organisatoriske foranstaltninger for uhensigtsmæssig deling af data fra borgeren.
3. Den digitale receptionist må ikke erstatte menneskelig kontakt, og det skal være tydeligt for borgeren, hvornår og hvordan de skal søge yderligere hjælp.
4. Der skal være et alternativ til den digitale receptionist, således at alle borgere, uanset teknologiske kompetencer og ressourcer, har lige adgang til hjælp og information.
5. Det skal tydeligt sikres, hvem der har det organisatoriske og faglige ansvar for teknologien og dens konsekvenser.

[6] Se fx: Fleisig, Eve, et al. "Linguistic bias in ChatGPT: Language models reinforce dialect discrimination." Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024.



Brugsscenarie 3
Automatisk referatskrivning

Dokumentation fra møder i socialforvaltningen er ofte et krav, men kan være tidskrævende og variere i kvalitet og struktur, hvilket kan påvirke både effektivitet og ensartethed i sagsbehandlingen. Derfor anvendes AI til automatiske referater i interne møder i socialforvaltningen og i mødet med borgere i socialforvaltningen. Den mest udbredte teknologi til dette i det kommunale er "RoboRef".

Hvem anvender den: Anvendes af medarbejdere i socialforvaltningen ved udarbejdelse af referater eller socialfaglige noter.

Hvilken AI-teknologi: Automatisk referatskrivning bygger på en kombination af tale-til-tekst-teknologi, som transskriberer lydoptagelser til tekst, og generativ AI, der omformer den transskriberede tekst til et struktureret udkast. Systemet anvender Natural Language Processing (NLP) til at analysere indholdet, identificere sekvenser og tematiske sammenhænge samt tilpasse udformningen til den valgte mødetype.

Eksempel: Læs mere om RoboRef her: RoboRef | Offentlig AI og her: Automatisk AI-generede notater og referater

Dataetiske overvejelser ved anvendelsen af automatisk referatskrivning

Referater og sagsnotater spiller en central rolle på socialområdet. De kan udgøre en væsentlig del af sagens dokumentation og indgå i afgørelsesgrundlaget. Derfor rejser automatiseret referatskrivning en række dataetiske spørgsmål, som adskiller sig fra de primært informationsbaserede AI løsninger. Ved anvendelsen af automatisk referatskrivning kan mange af de samme dataetiske principper påvirkes både positivt og negativt, alt efter hvordan teknologien anvendes. Det er derfor vigtigt, at teknologien anvendes dataetisk ansvarligt, hvis potentialet i teknologien skal udnyttes uden, at det gør mere skade end gavn. I det følgende gennemgås de dataetiske argumenter, der taler for og imod anvendelsen af automatiske referater til medarbejdere på det sociale område.

Argumenter og hensyn, som taler for anvendelsen af automatisk referatskrivning

Hvis anvendelsen af automatisk referatskrivning giver et systematisk og struktureret referat, kan det gøre det mere gennemsigtigt for borgeren og medarbejderen, hvad der blev sagt på mødet, og hvilket grundlag der efterfølgende arbejdes videre på. Det kan derfor øge *forklarligheden* hos både borger, sagsbehandler og myndighed. Især i de tilfælde, hvor borgeren inddrages i, hvordan referatet skrives, og eventuelt kan have medindflydelse på formatet, kan et AI referat gøre det lettere for borgeren at forstå sagens indhold, følge med i forløbet og gøre indsigelser. Standardiserede processer for transskribering og referatskrivning kan gøre det lettere for socialforvaltningen at beskrive og dokumentere, hvordan oplysninger fra møder behandles og hvilke oplysninger der bliver sagsfremmende. *Forklarligheden* forudsætter dog, at det tydeliggøres, at referatet er et AI genereret udkast, og at det skal kvalificeres og godkendes af sagsbehandleren.

De automatiske referater har samtidig potentialet til styrke borgerens *retssikkerhed*, hvis det giver anledning til mere fuldstændig og ensartet dokumentation. Referaterne kan blive mindre afhængige af den enkelte medarbejders hukommelse, notatteknik og subjektive udvælgelse af oplysninger, hvilket kan bidrage til mere ensartet sagsbehandling. Ligeledes kan teknologien skabe mere kontrolleret dokumentation af oplysninger fra møder end uformelle noter, hvilket kan understøtte en mere systematisk håndtering af personoplysninger og bedre sikre borgerens *privatlivsrettigheder*. Automatiseret dokumentation kan gøre det lettere at etablere faste procedurer for, hvilke oplysninger der skal dokumenteres, og hvordan de behandles. Samtidig kan automatiseret dokumentation understøtte mere standardiserede arbejdsgange og dermed reducere nogle typer af fejl og kvalitetsvariation i dokumentationsarbejdet. Teknologien kan gøre det lettere systematisk at kontrollere, kvalitetssikre og overvåge dokumentationsprocesser, hvilket også kan være med til at

øge *sikkerheden* for borgerne og for forvaltningen. Endelig kan logning af processen gøre det lettere at undersøge, hvordan et bestemt referat er fremkommet og tydeliggøre, hvilke oplysninger der lå til grund for det endelige referat og sagsvurdering. Det kan styrke mulighederne for efterfølgende at gennemgå samtaler og referater, hvis der er tvivl om *ansvarlighed* og ansvarsplacering.

Hvis anvendelsen af automatisk referatskrivning giver anledning til mere ensartede referater og socialfaglige noter, kan det bidrage til, at borgere i højere grad behandles lige på baggrund af samme type dokumentation, uanset sagsbehandler, hvilket kan mindske utilsigtet forskelsbehandling mellem borgere. Mere standardiseret dokumentation på tværs af medarbejdere og enheder kan også understøtte mere ensartet sagsbehandling og dermed større *lighed* mellem borgere på tværs af sagsbehandlere og kommuner.

Automatisk referatskrivning kan også bidrage positivt til samtalesituationen og borgerens oplevelse af *værdighed*. Hvis automatisk referatskrivning gør, at medarbejdere kan bruge mindre tid på at skrive noter og mere tid på den direkte dialog med borgeren, kan det styrke oplevelsen af at blive set og hørt. Hvis teknologien anvendes ansvarligt, kan den understøtte en offentlig sektor, hvor mennesker mødes med større opmærksomhed og mindre administrativt fokus under samtalen. For at understøtte borgerens *værdighed* bør det være klart for borgeren, hvornår og hvorfor samtalen optages, og hvordan optagelsen og referatet vil blive anvendt. Borgeren bør opleve, at teknologien anvendes med respekt for deres situation og ikke som et kontrol eller overvågningsværktøj.

Automatisk referatskrivning kan frigøre tid fra dokumentation til borgerkontakt, hvilket kan være til *gavn* for både borger og medarbejder. Samtidig har automatisk referatskrivning potentialet til at øge værdien af både at skrive referatet og det at have et referat i sagsbehandlingen efterfølgende. Frigivelse af ressourcer på borgerrettede aktiviteter frem for administrativ dokumentation, kan på sigt øge den samlede nytte af velfærdsindsatsen. Hurtigere og mere effektiv dokumentation kan understøtte en mere effektiv offentlig sektor og dermed skabe værdi for flere borgere.

Argumenter og hensyn, som taler imod anvendelsen af automatisk referatskrivning

Forklarlighed kan i denne use case både udfordres og styrkes. Udfordringen består især i, at det kan være vanskeligt for både medarbejdere og borgere at forstå, hvordan teknologien går fra lydoptagelse til færdigt referat, herunder hvilke oplysninger der er udvalgt, sammenfattet eller udeladt. Hvis medarbejderen ikke forstår, hvordan systemet fungerer, kan de heller ikke meningsfuldt anvende det til sagsbehandling eller forklare det til borgerne. Hvis det ikke kommunikeres tydeligt, at AI anvendes, eller hvordan forskellige oplysnin-

ger behandles i processen, svækkes *forklarligheden* og man kan ende med referater, som mister sin værdi, fordi de ikke repræsenterer det medarbejderen eller borgeren har oplevet i mødet.

Det kan samtidig påvirke borgernes *retssikkerhed*. *Retssikkerhed* er et helt centralt hensyn i denne use case, da referater og socialfaglige noter ofte er et afgørende dokument i borgerens sag. AI-genererede referater kan indeholde fejl, udeladelser eller misfortolkninger af det sagte. Spørgsmålet om, hvem der gennemgår og godkender et AI genereret referat, er derfor afgørende. Hvis fejl, udeladelser eller misforståelser i referatet ikke opdages, kan det få potentielt alvorlige konsekvenser for borgerens sag. Samtidig kan der opstå situationer, hvor borgeren deler oplysninger, som en medarbejder normalt ikke ville vurdere som relevante at notere, men som automatisk kommer med i referatet. Det kan føre til, at unødvendige oplysninger indgår i sagsbehandlingen uden en faglig vurdering af relevans og proportionalitet.

På samme måde kan det digitale referat inkludere personlige oplysninger, som ikke er nødvendige for sagens formål, hvilket kan føre til en mere omfattende registrering af borgerens *privatliv*. Lydoptagelser, transskriptioner og AI-genererede referater kan medføre, at flere oplysninger indsamles og lagres, end hvis medarbejderen alene havde udarbejdet et traditionelt referat. Samtidig er der med AI-genereret referatskrivning en større risiko for negativ påvirkning af *sikkerheden*. Hvis automatisk referatskrivning anvendes bredt og generelt, kan socialforvaltningens (og samfundets) dokumentationspraksis blive mere afhængig af tekniske systemers stabilitet og robusthed, hvilket øger sårbarheden over for nedbrud, driftsforstyrrelser, cyberangreb eller andre sikkerhedshændelser. Det kan få store konsekvenser for både borgere og samfund, når AI-teknologier håndterer store mængder dokumentation fra socialområdet.

Ud over de mere teknologiske udfordringer, som automatisk referatskrivning kan give anledning til, kan det også negativt påvirke borgerens oplevelse af *værdighed* i mødet med socialområdet. Automatisk referatskrivning anvendes direkte i samtalesituationer mellem borger og socialfaglig medarbejder og det kan opleves *uværdigt*, hvis en personlig samtale reduceres til et standardiseret AI-genereret referat, hvor nuancer og personlige perspektiver går tabt. Fejlfortolkninger eller upræcise formuleringer kan medføre, at borgeren fremstilles på en måde, som vedkommende ikke kan genkende eller oplever som stigmatiserende. Optagelse af samtaler kan have en effekt på relationen, tilliden og oplevelsen af fortrolighed. Samtidig kan nogle borgere føle sig overvåget eller tilbageholdende i deres fortælling, når samtalen optages.

Kvaliteten af de automatiske transskriptioner med tale til tekst teknologier kan være dårlige til at opfange eller korrekt gengive bestemte kommunikationsformer, dialekter, sproglige udfordringer eller atypiske måder at formulere sig på. Det er især relevant for nogle af de borgere, som ofte har møder med socialforvaltningen og som kan udtrykke sig anderledes

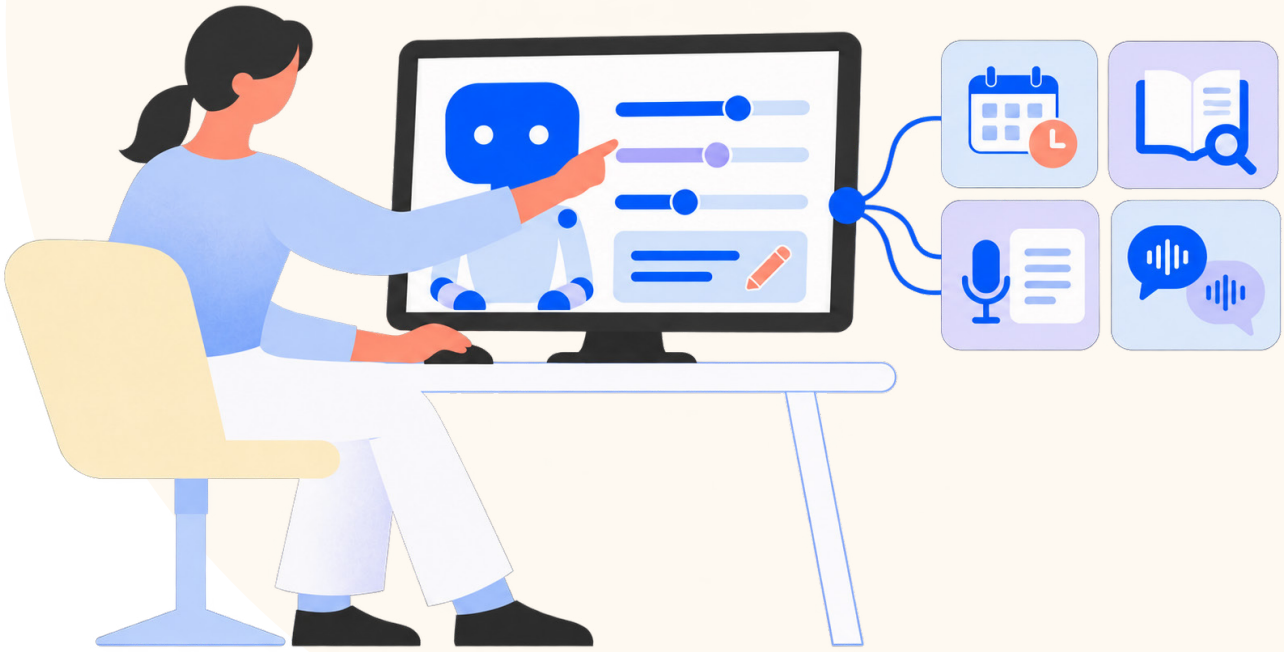
end normen. Der er derfor risiko for, at bestemte borgere systematisk får mindre præcise eller mere mangelfulde referater end andre, hvilket kan påvirke *ligheden* i systemet. Hvis sådanne forskelle ikke opdages og korrigeres, kan det føre til forskelsbehandling.

Endelig må anvendelsen af automatiske referater ikke skabe uklarhed om, hvem der har *ansvaret* for dokumentets indhold og format. Ved høj vægtning af AI-genererede referater, kan det blive vanskeligere at placere ansvar for fejlfortolkninger eller udeladelser af væsentlige oplysninger. Samtidig kan det potentielt blive svært at vurdere, hvilke dele af referatet der stammer fra den oprindelige samtale, og hvilke dele der er resultat af AI-teknologiens bearbejdning. Der kan opstå ansvarsudvanding, hvis fejl forklares med henvisning til teknologien frem for til identificerbare personer eller organisationer. Det er derfor essentielt, at det er tydeligt for både borger, sagsbehandler og myndighed, hvem der har godkendt og kvalitetssikret det AI-genererede referat.

Automatisk referatskrivning har altså potentialet til både at gøre meget *gavn*, men også at gøre en del *skade*. Referaterne kan være afgørende i sagsbehandlingen og konsekvenserne kan være langvarige for borgerens trivsel og livssituation. Derfor må anvendelsen af AI-genererede referater ikke ske på bekostning af referatets kvalitet eller relationen mellem borger, sagsbehandler og socialforvaltningen.

Fem dataetiske råd til en ansvarlig anvendelse af automatisk referatskrivning:

1. Det skal være helt tydeligt og forklarligt overfor både medarbejderen og borgeren, hvorvidt og hvordan der anvendes AI til automatisk referatskrivning.
2. Det skal sikres, at alle referater gennemgås og godkendes af sagsbehandleren, så fejl og unødvendige oplysninger opdages og håndteres.
3. Det skal sikres, at der er teknisk opmærksomhed på systemets begrænsninger, og klare organisatoriske krav om ensartet kvalitetssikring, herunder håndtering af mange måder at kommunikere og udtrykke sig på.
4. Det er afgørende, at medarbejderen og socialforvaltningen klart kan redegøre for, hvem der bærer det endelige ansvar for referatet og dets konsekvenser.
5. Teknologien skal anvendes med respekt for borgeren, og for relationen mellem borger og sagsbehandler.



Brugsscenarie 4
Personlige AI-agenter

AI-agenter bliver skræddersyet til den enkelte medarbejder med den information eller viden, de har mest brug for. Altså en form for specialiserede chatbots, som hjælper kommunens medarbejdere i deres specifikke arbejdsprocesser. Der er endnu ikke så mange konkrete eksempler på personlige AI-agenter på socialområdet, men det er taget med som brugsscenarie, da det med rimelighed kan forventes at være en teknologi, som vil blive implementeret endnu mere i den nærmeste fremtid.

Hvem anvender teknologien: Anvendes af medarbejdere i de funktioner, som de finder nødvendige. De personlige AI-agenter kan eksempelvis anvendes til informationssøgning, mødereferater, tolkeassistent og beslutningsstøtte baseret på eksisterende lovgivning og retningslinjer. I modsætning til AI-assistenter, kan AI-agenter tilpasse sig, lære af erfaringer og håndtere komplekse opgaver på tværs af systemer og afdelinger.

Hvilken AI-teknologi: Da de er skræddersyet til den enkelte, vil teknologien afhænge af den specifikke AI-agent. Teknologien bygger på generativ AI, som anvendes til at forstå spørgsmål og formulere svar i et naturligt sprog. Den kan kombineres med Retrieval Augmented Generation (RAG), hvor systemet søger i interne og forhåndsgodkendte dokumenter for at generere svar baseret på eksisterende materiale. Løsningerne kan desuden omfatte tale-til-tekst og tekst-til-tale, der muliggør både skriftlig og stemmebaseret interaktion.

Derudover kan Natural Language Processing (NLP) bruges til at analysere og strukturere tekst, identificere sekvenser og udforme udkast til referater, notater eller analyser baseret på transskriberet indhold.

Eksempel: Promte Chat, som er en platform, hvor man kan bygge sine egne digitale assistenter: Promte Chat | Offentlig AI

Dataetiske overvejelser ved anvendelsen af personlige AI-agenter

AI agenter adskiller sig fra AI-assistenter ved, at assistenter er reaktive værktøjer, der udfører opgaver på brugerens anmodning, mens AI agenter arbejder proaktivt og mere selvstændigt. AI agenter kan planlægge handlinger, igangsætte processer og håndtere opgaver helt eller delvist uden løbende menneskelig indblanding. I modsætning til mere afgrænsede AI assistenter kan AI agenter altså have en højere grad af autonomi og interaktion, og potentielt tilgå flere datakilder og systemer.[7,8] Det giver anledning til nogle dataetiske overvejelser. Særligt når agenten får en tæt integration i medarbejderens daglige praksis på et område, hvor afgørelser kan have vidtrækkende konsekvenser for borgerne. I det følgende gennemgås de dataetiske argumenter, der taler for og imod anvendelsen af personlige AI-agenter til medarbejdere på det sociale område.

Argumenter og hensyn, som taler for anvendelsen af personlige AI agenter

AI-agenten kan hjælpe med at sammenfatte komplekse regler, procedurer og informationer på en måde, der kan gøre det lettere for medarbejdere at forstå det informationsgrundlag, der anvendes i deres arbejde. Hvis forvaltningen etablerer klare rammer for agenternes anvendelse, datakilder og funktioner, kan det styrke gennemsigtigheden og *forklarligheden* af, hvordan data anvendes på socialområdet.

Samtidig kan personlige AI-agenter hjælpe medarbejdere ved at støtte dem i deres administrative og analyserende arbejde. Både ved hurtigt at finde relevant materiale og ved at producere sagsstøttende viden. Det kan understøtte, at borgernes rettigheder og lovkrav i højere grad bliver identificeret og efterlevet i den konkrete sag. AI-agenter kan hjælpe med at sikre, at relevante juridiske forhold inddrages i sagsbehandlingen, fordi medarbejdere hurtigere kan få adgang til den nødvendige viden og dermed være med til at sikre en bedre *retssikkerhed* på det sociale område.

Hvis AI-agenten samtidig reducerer tidsforbruget på administrative opgaver, kan medarbejderen bruge mere tid på den menneskelige relation, dialog og anerkendelse af borgerens situation, hvilket kan understøtte oplevelsen af *værdighed*. AI-agenten kan hjælpe

[7] AI-agenter: Hvad er de og hvordan fungerer de - helt simpelt? - DI
[8] https://www.ibm.com/think/topics/ai-agents-vs-ai-assistants

medarbejderen med at huske relevante oplysninger og tidligere aftaler, så borgeren i mindre grad skal gentage belastende eller personlige oplysninger. En *værdig* anvendelse forudsætter dog, at teknologien understøtter, frem for erstatter, den menneskelige faglighed, erfaring og relation. Ligeledes skal det tydeliggøres over for borgeren, hvordan teknologi og menneskelig sagsbehandler arbejder sammen.

Derudover kan AI-agenter understøtte medarbejdere i komplekse sager og dermed bidrage til mere konsistente vurderinger, uanset medarbejderens erfaring, lokalkendskab eller adgang til viden, hvilket kan fremme en mere *lige* og *retfærdig* behandling af borgere. Hvis teknologien frigør tid fra administrative opgaver, kan medarbejderne få bedre mulighed for at tage individuelle hensyn til borgernes forskellige behov og livssituationer.

Der er altså et potentiale for, at AI-agenter kan *gøre gavn* ved implementeringen i socialforvaltningen. Ved at reducere tiden brugt på administrative opgaver, kan teknologien frigive mere tid til borgerkontakt og fagligt arbejde, hvilket kan gavne den enkelte borgers sagsbehandling og oplevelse i socialforvaltningen. Bedre adgang til viden og beslutningsstøtte kan understøtte højere kvalitet i den offentlige sagsbehandling og dermed gavne borgere bredt. Endelig kan AI-agenter hjælpe medarbejdere med at håndtere stigende kompleksitet og informationsmængder, hvilket kan styrke kvaliteten af den offentlige service.

Argumenter og hensyn, som taler imod anvendelsen af personlige AI agenter

Ligesom personlige agenter kan øge *forklarligheden*, kan de ofte også opleves nærmest *uforklarlige*. Personlige AI agenter anvender ofte komplekse generative AI modeller, der kan kombinere information fra mange kilder og interagere dynamisk med brugeren. For medarbejderen kan det være vanskeligt at gennemskue, hvilken information AI-agenten baserer sine svar på, hvorfor agenten foreslår bestemte handlinger, prioriterer information på en given måde eller udelader nogle forhold. Borgere kan have svært ved at forstå, hvordan deres personoplysninger behandles og påvirker de forslag, som AI-agenten genererer, særligt hvis databehandlingen foregår på tværs af flere systemer. Hvis forvaltningen eller leverandøren ikke har fuld indsigt i modellernes interne processer, kan det være vanskeligt at forklare over for borgere, medarbejdere og tilsynsmyndigheder, hvordan (person)data konkret behandles og omsættes til svar. Manglende gennemsigtighed og *forklarlighed* kan også føre til en øget risiko for automation bias, hvor medarbejderen tillægger agentens output større vægt end egen faglig vurdering. Dette er særligt problematisk på socialområdet, hvor faglig dømmekraft, helhedsvurderinger og individuelle menneskelige hensyn er helt centrale.

Selvom personlige AI agenter ikke er tiltænkt som beslutningstagere, kan de indirekte få betydning for borgerens *retssikkerhed*, når agentens svar og forslag indgår i sagsbehandlingen. AI-agenten kan give fejlagtige eller mangelfulde oplysninger om lovgivning, praksis

eller rettigheder, hvilket kan føre til afgørelser eller handlinger, der svækker borgerens retssikkerhed. Samtidig kan risikoen for automation bias øge sandsynligheden for, at medarbejderen læner sig mere ind i teknologien end i egen faglighed.[9] Hvis medarbejdere lægger for stor vægt på AI-agentens anbefalinger, kan der opstå risiko for, at juridiske vurderinger i praksis påvirkes af et system, som ikke selv kan bære et juridisk ansvar. *Retssikkerhed* forudsætter derfor, at myndigheden kan redegøre for, hvordan og i hvilket omfang, AI agenter anvendes, at det altid er tydeligt, hvilke informationer der har dannet grundlag, samt hvordan AI-agentens svar indgår som input i borgerens sag. Hvis AI-agenten systematisk formidler fejlfortolkninger af regler eller praksis, kan retssikkerhedsproblemer blive et samfundsproblem. En stigende afhængighed af AI-baseret beslutningsstøtte kan svække den juridiske og faglige kontrol, som er central for at beskytte borgernes grundlæggende rettigheder.

Personlige AI agenter kan potentielt få adgang til store mængder data afhængigt af indstillingerne. Det gælder både interne dokumenter, mails og i nogle tilfælde borgeroplysninger. Det rejser væsentlige spørgsmål om *sikkerhed* og *privatliv*, især i form af datatilgængelighed, dataminimering og adgangsstyring. Hvis agenten kan tilgå data på tværs af systemer, øges risikoen for utilsigtet brug eller videregivelse af følsomme oplysninger. Samtidig kan AI-agenten skabe incitament til at samle og anvende flere oplysninger om borgere end det, der er nødvendigt for den konkrete opgave. Derudover kan teknologien bidrage til en udvikling, hvor det bliver lettere og mere almindeligt at udlede detaljerede indsigter om borgere på baggrund af store mængder data, hvilket kan udfordre samfundets normer for *privatliv*. Hvis agenten integreres med mange interne systemer, kan konsekvenserne af fejl, nedbrud eller kompromittering blive større end for mere afgrænsede løsninger.

De personlige AI-agenter kan også påvirke dataetiske principper som *retfærdighed* og *lighed*. Teknologiens forslag kan være baseret på normative mønstre, som ikke nødvendigvis er rimelige i forhold til den konkrete borgers situation, og dermed risikere at fastholde eksisterende skævheder. Borgere i socialforvaltningen kan have individuelle forhold og særlige behov, som risikerer at kunne få mindre opmærksomhed, hvilket kan føre til mindre *lige* eller *retfærdige* løsninger for nogle borgere. Ligeledes kan borgere med komplekse eller atypiske problemstillinger risikere at blive dårligere understøttet, hvis AI-agenten primært er optimeret til mere typiske sagsforløb.

Anvendelsen af personlige AI-agenter rejser også spørgsmål om medarbejderens og borgerens *værdighed*. For sagsbehandleren kan en AI agent opleves som en form for digital "skygge". Det rejser spørgsmål om faglig autonomi og om muligheden for enten ikke at anvende teknologien eller for at kunne sige systemet imod. Nogle medarbejdere vil kunne opleve det som uværdigt, at deres faglighed og menneskelige erfaringer bliver "reduceret" til noget en computer kan løse. Således kan AI-agenter ændre socialfagligheden og reducere fælles faglig refleksion, kollegial sparring og faglig pluralitet. For borgeren kan *værdig-*

[9] Se fx: Strauß, Stefan. "Deep automation bias: How to tackle a wicked problem of AI?." Big Data and Cognitive Computing 5.2 (2021): 18.

heden udfordres, hvis teknologien indirekte får en større rolle i vurderingen af deres sag og især hvis borgeren oplever, at deres livssituation og menneskelige og sociale udfordringer primært mødes af en teknologi med et fokus på effektiv sagsbehandling. Hvis teknologien gradvist overtager centrale dele af de professionelle arbejdsprocesser, kan det svække betydningen af relationel forståelse, empati og menneskelig dømmekraft, som ofte er centrale for at behandle borgere med *værdighed*.

Personlige AI agenter kan potentielt udføre flere og mere komplekse opgaver end mere afgrænsede AI værktøjer, og kan i praksis komme til at fungere som en slags digital forlængelse af medarbejderen. Dette rejser spørgsmålet om *ansvarlighed* overfor agentens handlinger, vurderinger og forslag. Særligt problematisk bliver det, hvis agenten i praksis får karakter af at være selvkørende, eksempelvis ved automatisk at indsamle information, prioritere opgaver eller foreslå konkrete faglige vurderinger. Således kan medarbejdere blive i tvivl om de ansvarsmæssige konsekvenser af at anvende en personlig AI-agent. Hvis roller og beføjelser ikke er klart defineret, kan der opstå uklarhed om ansvarsplaceringen, både internt i organisationen og i forhold til borgeren.

Generelt har AI-agenter et potentiale til både at *gøre gavn* og til *at skade* både borgerens og medarbejderens situation. Personaliseringen af agenten kan medføre, at medarbejderen får stor tillid til teknologien og overser relevante alternative perspektiver eller handlemuligheder. Hvis teknologiens evner overvurderes eller reducerer den menneskelige kontakt, kan kvaliteten af opgaveløsningen forringes til *skade* for borgere og samfund. Samtidig kan en generel afhængighed af AI-støtte på længere sigt svække medarbejdernes faglige kompetencer. Dermed kan vigtig viden og erfaring gå tabt i socialforvaltningen.

Fem dataetiske råd til en ansvarlig anvendelse af personlige AI agenter:

1. AI-agentens funktioner skal være gennemsigtige, og det skal være forståeligt, hvilket datagrundlag og hvilke antagelser der ligger bag agentens svar og forslag. Ligeledes skal medarbejderen uddannes og oplæres grundigt i anvendelsen af en AI-agent.
2. Der skal være skarpe tekniske og organisatoriske rammer for, hvilke data AI-agenten må tilgå og til hvilke formål.
3. AI-agenten og systemet bag skal være sat op til at håndtere borgernes komplekse og ofte mere atypiske problemstillinger.
4. Der skal være klare rammer for ansvar for agenten, samt tydelige procedurer for tilsyn og kvalitetssikring.
5. Det skal sikres, at teknologien understøtter, frem for erstatter, den menneskelige faglighed, erfaring og relation, og at det tydeliggøres overfor borgeren, hvordan teknologi og menneskelig sagsbehandler arbejder sammen.



Brugsscenarie 5
Intelligente beslutningsstøttende værktøjer

Denne case findes lige nu kun som pilot-projektet "MaiLife" som en AI-teknologi, der understøtter sagsbehandlere på det specialiserede børne- og ungeområde i kommuner. Systemet analyserer relevant data, inklusiv sundhedsdata, og identificerer mønstre og risikofaktorer ved mistrivsel hos børn og unge. Den er taget med som brugsscenarie, da det med rimelighed kan forventes, at være en teknologi, som kan blive mere udbredt i den nærmeste fremtid.

Hvem anvender den: Anvendes som pilotprojekt af medarbejdere på det specialiserede børne- og ungeområde, herunder sagsbehandlere og fagpersoner, som arbejder med komplekse vurderinger og har behov for dataunderstøttet indsigt. De kan få teknologiens hjælp til at analysere data for at identificere mønstre, risikofaktorer og foreslå fagligt funderede handlinger. MaiLife skal fungere som et intelligent beslutningsstøttende værktøj, ikke som beslutningstager.

Hvilken AI-teknologi: MaiLife anvender avancerede AI-teknologier til at analysere store sundheds- og sagsdatasæt for at identificere mønstre og risikofaktorer. Systemet bearbejder komplekse datakombinationer på tværs af fx psykisk sundhed, skolegang og familiedynamik, og det kan generere analyser og dokumentation automatisk. MaiLife gør desuden brug af generativ AI til at udforme datadrevne anbefalinger og faglige analyser samt Natural Language Processing (NLP) baserede teknikker, der strukturerer og behandler tekstbaseret materiale.

Eksempel: Læs mere om MaiLife (Genair8) her: MaiLife (Genair8) | Offentlig AI og her: Om Mai Life - Genair8

Dataetiske overvejelser ved anvendelsen af intelligente beslutningsstøttende værktøjer

Anvendelsen af intelligente beslutningsstøttende værktøjer analyserer store mængder data med henblik på at identificere mønstre, risikofaktorer og relevante fokusområder i komplekse sager. Selvom teknologien ikke er tiltænkt som en selvstændig beslutningstager, befinder use casen sig tættere på reel myndighedsudøvelse. Det gør de dataetiske overvejelser særligt væsentlige. I det følgende gennemgås de dataetiske argumenter der taler for og imod anvendelsen af intelligente beslutningsstøttende værktøjer på det sociale område.

Argumenter og hensyn, som taler for anvendelsen af intelligente beslutningsstøttende værktøjer

Anvendelsen af intelligente beslutningsstøttende værktøjer kan potentielt have positive effekter for *retssikkerheden*, hvis teknologien kan være med til at identificere relevante mønstre eller informationer, som ellers kunne blive overset i komplekse sager med stort datagrundlag. Teknologien kan bidrage til, at relevante forhold dokumenteres og inddrages mere ensartet i vurderinger og indsatser på tværs af kommuner. Her skal man dog være opmærksom på, at de relevante datapunkter kan bestå af både mere objektive informationer og mere socialt producerede registreringer. Det vil kræve et højt niveau af datakvalitet på tværs, samt at teknologien kan skelne, hvordan disse datapunkter skal fortolkes.

Hvis intelligente beslutningsstøttende værktøjer kan hjælpe medarbejdere med at identificere behov, risikofaktorer og relevante indsatser mere systematisk, kan de bidrage til, at borgere får den støtte, de faktisk har behov for. Dette kan understøtte mere *retfærdige* sociale vilkår for udsatte borgere. Beslutningsstøttende værktøjer kan styrke samfundets mulighed for at identificere sociale problemer tidligere og målrette indsatser mod borgere med størst behov. Derved kan teknologien understøtte, at begrænsede velfærdsressourcer anvendes på et mere informationsbaseret grundlag, så indsatser i højere grad når de grupper, der har størst behov for dem. På samme måde kan teknologien bidrage til at reducere betydningen af forskelle i medarbejdernes erfaring, viden og kapacitet og dermed bidrage til mere *lige* adgang til relevante indsatser og støtte.

Ved at analysere store mængder data kan intelligente beslutningsstøttende værktøjer *gøre gavn*, ved at understøtte medarbejdere i komplekse sager og bidrage til et bedre beslutningsgrundlag. Teknologien kan hjælpe med at identificere ellers usete forhold, og dermed potentielt gavne sagsbehandlingen, ligesom den kan bidrage til mere effektiv udnyttelse af faglige og økonomiske ressourcer ved at målrette indsatser mod de områder, hvor behovene er størst.

Argumenter og hensyn, som taler imod anvendelsen af intelligente beslutningsstøttende værktøjer

Da avancerede AI modeller ofte fungerer som komplekse og delvist uigennemsigtige systemer, er *forklarlighed* en væsentlig udfordring ved anvendelsen af intelligente beslutningsstøttende værktøjer. For medarbejdere kan det være vanskeligt at forstå, hvordan systemet når frem til sine analyser, og hvilke faktorer der vægtes i vurderingen. Manglende forklarlighed kan svække medarbejderens mulighed for at efterprøve og udfordre teknologien og dermed øge risikoen for ukritisk anvendelse og automation bias. Ligeledes kan borgere have svært ved at gennemskue, hvordan forskellige oplysninger er blevet vægtet og kombineret i analysen. For at understøtte forklarlighed skal teknologien tydeligt kunne redegøre for informations- eller datagrundlaget, de væsentligste overvejelser der er taget med, samt de forhold, der taler for og imod en given vurdering.

Anvendelsen af intelligente beslutningsstøttende værktøjer kan udgøre en væsentlig risiko, hvis systemets output i praksis får karakter af at være beslutningstagende, frem for beslutningsstøttende. Borgere kan komme til at blive vurderet på baggrund af statistiske sandsynligheder og risikoprofiler frem for konkrete forhold i deres sag. Det kan udfordre grundlæggende retssikkerhedsprincipper om individuel vurdering. For at sikre *retssikkerheden* er det afgørende, at AI systemet alene anvendes til at pege på relevante opmærksomhedspunkter og ikke til at konkludere eller anbefale endelige løsninger. Samtidig skal det være muligt for medarbejderen at kunne redegøre for, hvad der taler for og imod en given beslutning, og hvilke elementer der indgår i vurderingen. Hvis beslutningsstøtten i praksis får stor betydning for myndighedens handlinger, kan det blive uklart, i hvilket omfang afgørelser og indsatser reelt bygger på menneskelige vurderinger.

AI modellerne ved intelligente beslutningsstøttende værktøjer bygger på eksisterende data, som kan indeholde skævheder og bias. Hvis datagrundlaget er ufuldstændigt, skævt eller af varierende kvalitet, kan systemet reproducere eller forstærke eksisterende forskelle i behandlingen af borgere, og de dataetiske principper om *lighed* og *retfærdighed* bliver derfor nemt udfordret. Risikoen for bias er særlig høj i komplekse sociale data, hvor der ofte er mange forskellige datakilder, og hvor en social forståelse af konteksten kan være afgørende. Borgere fra grupper, som er underrepræsenteret eller atypiske i datagrundlaget, kan risikere mindre præcise analyser og dermed dårligere støtte end andre borgere. Risikovurderinger baseret på intelligente beslutningsstøttende værktøjer kan medføre, at borgere mødes ud fra sandsynligheder og prognoser frem for deres aktuelle situation og individuelle behov.

Anvendelsen af intelligente beslutningsstøttende værktøjer kan også påvirke dataetiske principper om *værdighed* i mødet mellem borger og socialforvaltning. Teknologien indebærer ofte behandling af omfattende og følsomme oplysninger om borgeres livssituation,

herunder sundhed, familieforhold og sociale problemer. Samtidig sammenstiller teknologien oplysninger fra flere livsområder, hvilket kan give myndighederne indsigt i dele af borgerens liv, som ellers ville være spredt på forskellige systemer og sammenhænge. Anvendelsen af datadrevne risikovurderinger kan bidrage til en kultur, hvor mennesker i stigende grad betragtes som datasæt, risikoprofiler eller analyseobjekter frem for som personer med iboende værdi og *værdighed*. Muligheden for at udlede nye oplysninger om borgere gennem avancerede analyser kan betyde, at der skabes viden om personer, som de ikke selv har valgt at dele. Dette stiller høje krav til beskyttelse af *privatliv* og proportionalitet i dataanvendelsen. Risikoen er, at individuelle nuancer og social kontekst går tabt i analysen, hvilket kan påvirke oplevelsen af at blive set som et helt menneske.

Endelig er *ansvarlighed* også et helt centralt dataetisk princip for anvendelsen af intelligent beslutningsstøttende værktøjer. Anvendelsen af AI må ikke skabe uklarhed om, hvem der træffer beslutninger, og hvem der kan stilles til ansvar. Det kan være uklart for borgeren, hvilken rolle AI-værktøjet har spillet i en konkret vurdering eller indsats, og samtidig hvem der reelt er ansvarlig for konsekvenserne. Medarbejdere kan blive tilbøjelige til at lægge stor vægt på systemets analyser, hvilket kan skabe usikkerhed om, hvordan ansvaret kan placeres. Der kan opstå risiko for "ansvarsdiffusion", hvis flere aktører er involveret i udvikling, drift og anvendelse, uden at nogen har et klart ansvar for konsekvenserne.

En teknologi med så avanceret AI som intelligente beslutningsstøttende værktøjer, har potentialet til at *gøre gavn*, men samtidig også en risiko for at forvolde meget *skade*. Givet teknologiens høje niveau af kompleksitet, kræver det både grundige overvejelser og klare rammer for anvendelsen. Hvis der kommer en overdreven tillid til de meget komplekse datarevne analyser, kan det bidrage til at fortrænge faglige perspektiver, relationel viden og lokalkendskab, hvilket kan reducere kvaliteten af hjælpen til borgerne i socialområdet.

Fem dataetiske råd til en ansvarlig anvendelse af intelligente beslutningsstøttende værktøjer:

1. Det skal defineres og rammesættes hvornår der er tale om et beslutnings*støttende* værktøj og et beslutnings*tagende* værktøj.
2. Anvendelse forudsætter, at der er et betydeligt arbejde med datakvalitet på tværs af relevante sektorer og en opmærksomhed på, hvordan forskellige grupper af borgere eller sager påvirkes af dataanvendelsen i teknologien.
3. Der skal opbygges kompetencer blandt medarbejdere på socialområdet til at forstå og anvende systemet både kritisk og hensigtsmæssigt, således at der er gennemsigtighed omkring teknologiens anvendelse, også mellem medarbejder og borger.
4. Anvendelsen på socialområdet kræver, at teknologien anvendes som supplement til en helhedsorienteret vurdering med en stærk forankring i et erfaringsbaseret og socialfagligt skøn.
5. Det skal altid være tydeligt, hvem der har ansvaret for sagsbehandlingen og de endelige afgørelser. Ansvarlighed forudsætter både ledelsesmæssig stillingtagen, klare retningslinjer og løbende kontrol med, hvordan systemet anvendes i praksis.



Brugsscenarie 6
Uofficiel anvendelse af AI

Ud over de formelt implementerede AI-teknologier, som er initieret eller rammesat af kommunerne, foregår der forventeligt også en del uofficiel anvendelse af AI blandt medarbejdere – også på socialområdet. Anvendelsen sker typisk uden for organisationens systemer og rammer, og kan ske via private værktøjer eller enheder.

Her tænkes både på medarbejdere, der anvender Microsofts CoPilot, og på brugen af fx ChatGPT, Claude eller Mistral til at understøtte opgaveløsningen. I nogle tilfælde vil anvendelsen ikke kunne spores, da den kan ske via private adgange, uden logning eller uden for organisationens it-systemer.

Hvem anvender teknologien: Anvendes af medarbejdere baseret på deres egen vurdering af, hvornår AI-teknologien kan være til gavn for dem i deres arbejde. Hvordan teknologien anvendes, og hvilken data der bliver delt med teknologien er uvist og afhængig af den enkelte situation.

Hvilken AI-teknologi: Teknologier som Copilot, ChatGPT og Claude er generativ AI og bygger på store sprogmodeller, der kan forstå og generere tekst i et sprog tilpasset modtageren. De kombinerer dybe neurale netværk, NLP-teknikker, prompt-styring, og ofte RAG-funktioner, som henter dokumentbaseret viden for at skabe mere præcise og kontekstuelle svar.

Eksempler: Der er ikke nogen officielle kilder på, hvor ofte medarbejderne bruger AI i deres arbejde uden om de officielle kanaler og rammer.

Dataetiske overvejelser ved uofficiel anvendelse af AI

Formålet med at anvende AI for den enkelte medarbejder er ofte at effektivisere arbejdet og få bedre støtte i opgaveløsningen. Men hvis anvendelser sker uden organisatorisk forankring, risikerer man, at hverken databehandling, sikkerhed eller rammerne for anvendelsen er afklaret. Denne use case er altså ikke et bevidst implementeret tiltag, men snarere en praksis, der opstår i takt med teknologiens udbredelse. Dette gør use casen særligt dataetisk udfordrende, da der ikke vil være viden om, hvordan og hvornår teknologien bliver anvendt, hvilken data der deles og hvordan teknologiens svar integreres ind i den egentlige sagsbehandling. I det følgende gennemgås de dataetiske argumenter der taler for og imod uofficiel AI-anvendelse til medarbejdere på det sociale område.

Argumenter og hensyn, som taler for anvendelsen af uofficiel anvendelse af AI

Selvom denne måde at anvende AI primært påvirker de dataetiske værdier og hensyn negativt, er der stadig en mulighed for, at anvendelsen kan *gøre gavn*. Herunder at medarbejdere kan opleve mindre arbejdspres og administrativ belastning, hvis AI bruges som støtteværktøj. Ligeledes kan hurtigere hjælp til informationssøgning, tekstproduktion og opgaveløsning, frigøre tid til kerneopgaver og kontakt med borgere.

Argumenter og hensyn, som taler imod anvendelsen af uofficiel anvendelse af AI

Retssikkerheden vil blive udfordret, hvis AI anvendes uden dokumentation, sporbarhed og kvalitetssikring. Hvis en medarbejder eksempelvis anvender et eksternt AI værktøj til at formulere dele af et notat eller en vurdering, kan det være uklart, hvilke dele af materialet der er AI genereret, og hvilket grundlag det bygger på. Dette kan gøre det vanskeligt for både borgeren og myndigheden at gennemskue, hvordan en sag er blevet behandlet, og hvilke overvejelser der ligger bag. Manglende dokumentation og transparens kan dermed svække borgerens mulighed for indsigt, indsigelse og klage. Ligeledes kan det svække medarbejderens egen mulighed for at fremvise sagsgangen, og manglende dokumentation af AI-anvendelsen kan gøre det vanskeligere at efterprøve en sag. En udbredt praksis med uofficiel AI-anvendelse kan gøre det vanskeligere at sikre, at forvaltningsmæssige garantier og juridiske rettigheder håndhæves ensartet og dokumenterbart. Udbredt uofficiel brug af AI kan udgøre en risiko for at opretholde *retssikkerheden* på et bredt samfundsniveau.

Der er også en risiko for alvorlige konsekvenser for retten til *privatliv*. Hvis medarbejdere indtaster personoplysninger eller følsomme oplysninger om borgere i AI-tjenester, som ikke er godkendt til formålet, risikeres det, at eksterne aktører kan få adgang til oplysningerne. Herved kan man miste kontrol over, hvem der får adgang til oplysninger om deres personlige forhold. Hvis medarbejdere deler personoplysninger eller følsomme oplysninger med

AI-værktøjer uden for de godkendte systemer, kan det udfordre borgernes ret til *selvbestemmelse* over egne data og livshistorie. Da anvendelsen kan ske via private konti eller private enheder og uden for organisationens logning, kan det være vanskeligt at vide, hvor oplysningerne befinder sig, og hvordan de behandles. Udbredt uofficiel anvendelse af AI kan svække den offentlige sektors evne til at sikre systematisk beskyttelse af borgernes personoplysninger.

Ligeledes kan anvendelse af AI uden for kommunens godkendte systemer øge risikoen for, at *sikkerheden* bliver kompromitteret og manglende logning og overvågning kan gøre det vanskeligere at opdage og håndtere sikkerhedshændelser. Uofficiel anvendelse kan skabe et mindre robust og mere fragmenteret sikkerhedsbillede i den offentlige sektor, fordi anvendelsen sker uden fælles kontrolmekanismer. Manglende overblik over, hvilke AI-tjenester der anvendes, kan gøre det vanskeligere at gennemføre effektive sikkerhedsforanstaltninger og risikostyring. Hvis sikkerhedshændelser opstår som følge af uofficiel anvendelse, kan det svække tilliden til offentlige myndigheders evne til at beskytte informationer og drive stabile digitale løsninger.

I tilfælde, hvor AI anvendes via private konti eller uden for organisationens godkendte systemer, kan det være vanskeligt at dokumentere, hvem der har anvendt teknologien, og hvordan den er blevet anvendt. Derved kan der opstå uklarhed om, hvem der har *ansvaret* og hvornår. Både for databehandlingen og sagsbehandlingen. Samtidig kan ledelsen mangle indsigt i, hvor og hvordan AI faktisk anvendes i praksis, hvilket gør det vanskeligt at føre tilsyn og sikre kvalitet. Utilstrækkelig ansvarsplacering kan svække tilliden til socialforvaltningen.

Forklarligheden udfordres i særlig grad ved uofficiel anvendelse af AI, fordi det ofte er uklart, hvordan et givent output er fremkommet. Medarbejderen vil ikke have indsigt i, hvilke kilder, data eller antagelser systemet baserer sine svar på, og kan derfor have vanskeligt ved at redegøre for indholdet. Samtidig kan manglende dokumentation af, hvornår og hvordan AI er anvendt, gøre det svært at forklare, hvordan en tekst, vurdering eller formulering er blevet til. Dette svækker muligheden for at skabe gennemsigtighed i sagsbehandlingen og kan gøre det vanskeligt at efterprøve grundlaget for en vurdering. Borgere vil ofte ikke vide, at AI har været anvendt i forbindelse med en opgave, vurdering eller formulering, fordi anvendelsen sker uofficielt og uden faste procedurer for oplysning om brugen. Den manglende institutionelle indsigt i anvendelsen kan ligeledes gøre det vanskeligt at give borgere meningsfulde forklaringer på, hvordan deres data er blevet behandlet, og hvordan det har påvirket deres sagsbehandling generelt i socialforvaltningen.

For borgerne kan det opleves som *uværdigt*, at deres informationer deles med en AI-teknologi og det kan forværres, når AI-teknologien ikke har en kontrolleret anvendelse. Det kan opleves som en instrumentalisering af personen, frem for en behandling med behørigt

hensyn. Samtidig, hvis medarbejdere anvender AI til at formulere vurderinger eller beskrivelser af borgere uden tilstrækkelig faglig kontrol, kan der opstå upræcise eller stereotype beskrivelser, som kan virke stigmatiserende. Den uofficielle karakter af anvendelsen kan også betyde, at borgere behandles som genstand for eksperimenterende arbejdspraksisser. Manglende organisatoriske rammer for respektfuld anvendelse af AI kan bidrage til en forvaltningskultur, hvor hensynet til effektivitet får forrang frem for hensynet til menneskers *værdighed*.

Uofficiel anvendelse af AI-teknologier, kan altså gøre stor *skade*. Samtidig er der få argumenter, som taler for den gavn teknologien kan gøre. Ud over de dataetiske principper, risikerer borgerne at miste deres tillid både til deres sagsbehandler og til socialforvaltningen.

Fem dataetiske råd omkring uofficiel anvendelse af AI:

1. Generelt bør al uofficiel anvendelse af AI undgås.
2. Myndigheder på socialområdet skal være tydelige omkring deres forventning til medarbejderne vedrørende brugen af AI (uanset form).
3. Der skal udvikles klare politikker for anvendelse af AI, herunder hvad der er tilladt og ikke tilladt.
4. Medarbejdere i socialforvaltningen skal uddannes i at forstå, hvordan AI kan bruges og de potentielle konsekvenser af forskellige former for anvendelse.
5. Myndigheden og ledelsen skal tage ansvar for ikke at skabe rammer, hvor uofficiel anvendelse af AI kan opleves som nødvendig for at kunne løse opgaverne til tiden.

Kapitel 3
Anbefalinger

Afsluttende præsenterer Dataetisk Råd her vores anbefalinger til ansvarlig anvendelse af AI-teknologier på socialområdet. Anbefalingerne er baseret på rådsprojektets fund og dataetiske overvejelser, både i de dataetiske analyser og i de møder, vi har haft med aktører på området. Vores erfaringer fra projektet viser, at AI-teknologier ikke er en nem og hurtig løsning, men hvis de anvendes ansvarligt, kan de i mange tilfælde bidrage til at styrke de faglige indsatser på socialområdet. Dataetisk Råd er fortaler for, at de seks anbefalinger bør følges i både eksisterende og kommende AI-projekter, og gerne så tidligt i forløbet som muligt. Anbefalingerne er især rettet mod kommunale beslutningstagere, men de er bestemt også relevante for medarbejdere og borgere på socialområdet.

Dataetisk Råds seks anbefalinger til anvendelsen af AI på socialområdet

1 Formålet med anvendelsen af AI skal være klart

Inden en AI-teknologi tages i brug, skal det være klart, hvorfor man vil anvende den. Samtidig skal det være tydeligt, hvilken gavn teknologien skal gøre og hvordan man vil undgå, at den gør skade.

Derfor bør kommunalbestyrelsen og den administrative ledelse klart kommunikere forventninger og succeskriterier for AI-teknologier, samt løbende evaluere teknologien.

2 Ansvaret for AI skal altid kunne placeres hos mennesker

Ansvaret for socialfaglige vurderinger og beslutninger skal altid være tydeligt placeret hos ledelsen og blandt sagsbehandlere. AI-teknologier kan ikke holdes ansvarlige.

Kommunalbestyrelsen og den administrative ledelse bør derfor etablere faste procedurer for godkendelse, kvalitetssikring og ansvarlighed, så det altid er klart, hvem der bærer det endelige faglige ansvar for AI-teknologien og dens konsekvenser.

3 Anvendelsen af AI skal være gennemsigtig og til at forklare

AI-teknologier må aldrig skabe tvivl om, hvordan socialfaglig dokumentation er fremkommet. Det skal være gennemsigtigt for både ledelse, sagsbehandlere og borgere, hvornår en AI-teknologi anvendes, hvilke data der indgår og hvordan teknologien fungerer.

Derfor bør kommunalbestyrelsen og den administrative ledelse sikre, at anvendelsen af AI er tydelig, dokumenteret og forståelig for både medarbejdere og borgere.

4 Datakvaliteten på socialområdet skal styrkes

Datakvaliteten på socialområdet varierer i dag, samtidig er datakvalitet afgørende for, hvordan en AI-teknologi fungerer. Derfor skal prioriteringen af AI-indsatser afstemmes med kvaliteten af de data, der er til rådighed og der skal arbejdes systematisk med at styrke datagrundlaget på socialområdet, hvis potentialet skal indfries.

Derfor bør kommunalbestyrelsen og den administrative ledelse fortsat arbejde for at skabe fælles standarder og kvalitetssikring af data på socialområdet.

5 Anvendelsen af AI skal fremme borgernes værdighed

Anvendelsen af AI skal understøtte en individuel indsats, der tager højde for borgernes forskellige livssituationer, samt fremmer en værdig adgang til socialområdet. Samtidig skal AI anvendes inden for klare tekniske og organisatoriske rammer, der beskytter borgernes privatliv og understøtter borgernes mulighed for indsigt og medinddragelse.

Derfor bør kommunalbestyrelsen og den administrative ledelse etablere faste procedurer for at sikre en individuel sagsbehandling, aktiv borgerinddragelse og beskyttelse af borgernes privatliv.

6 AI skal styrke socialfagligheden

AI bør anvendes som et støtteværktøj, der både gavner og styrker arbejdet på socialområdet. Den socialfaglige vurdering og relation til borgeren skal fortsat være central. Sagsbehandlerne skal derfor inddrages samt have de nødvendige kompetencer og ressourcer til at forstå, udfordre og kvalitetssikre teknologiens bidrag.

Derfor bør kommunalbestyrelsen og den administrative ledelse altid inddrage ansatte på socialområdet i anvendelsen af AI og investere i løbende uddannelse, faglig sparring og kvalitetssikring.

Om Dataetisk Råd

Dataetisk Råd blev etableret i 2019 og består af 16 rådsmedlemmer inklusiv en formand.

Formålet med Dataetisk Råd er at skabe et forum, hvor etiske spørgsmål om forholdet mellem på den ene side fordelene ved anvendelse af ny teknologi og på den anden side hensynet til borgernes grundlæggende rettigheder og samfundsmæssige værdier, kan drøftes, og hvor en varig indsats for understøttelse af en ansvarlig anvendelse af data kan forankres. Samfundets behov for at anvende data og at udnytte de teknologiske muligheder skal tilgodeses og respekteres. Dataetisk Råd kan bidrage til, at brugen af data sker dataetisk forsvarligt og dermed bidrage til den positive udvikling i anvendelse af data og teknologier.

Dataetisk Råd beskæftiger sig med at **skabe dataetisk debat og offentlig opmærksomhed.**

Rådet blev etableret for at opretholde et højt niveau af tillid til digitale offentlige tjenester og brugen af persondata i det offentlige. Gennem anbefalinger, redegørelser, udtalelser og debatskabende aktiviteter imødekommer Dataetisk Råd behovet for løbende debat og offentlig opmærksomhed om dataetiske dilemmaer ved nye teknologier og teknologiske muligheder.

Dataetisk Råd understøtter løbende en **ansvarlig og bæredygtig dataanvendelse i erhvervslivet og i den offentlige sektor**. Rådet arbejder for en ansvarlig tilgang til dataanvendelse både nationalt og internationalt, og yder vejledning til folketinget, regeringen og myndigheder i forbindelse med gennemførelse af relevante initiativer. Ligeledes rådgiver rådet også andre interessenter i samfundet omkring dataetik og dataetiske overvejelser.

Dataetisk Råd understøttes af et sekretariat i Nationalt Center for Etik. For mere information, kontakt gerne Dataetisk Råd på: dataetiskraad@dketik.dk