



Bruxelles, den 8.4.2019
COM(2019) 168 final

**MEDDELELSE FRA KOMMISSIONEN TIL EUROPA-PARLAMENTET, RÅDET,
DET EUROPÆISKE ØKONOMISKE OG SOCIALE UDVALG OG
REGIONSUDVALGET**

Opbygning af tillid til menneskecentreret kunstig intelligens

DA

DA

1. INDLEDNING – DEN EUROPÆISKE STRATEGI FOR KUNSTIG INTELLIGENS

Kunstig intelligens har potentialet til at forandre vores verden til det bedre: den kan forbedre sundhedspleje, mindske energiforbrug, gøre biler sikrere og gøre det muligt for landbrugere at anvende vand og naturressourcer mere effektivt. Kunstig intelligens kan bruges til at forudsige miljø- og klimaændringer og forbedre den finansielle risikostyring, og den giver redskaber til at fremstille produkter - med mindre affald, der er skræddersyede til vores behov. Den kan også bidrage til at afsløre svig og trusler mod cybersikkerheden og gør det muligt for de retshåndhævende myndigheder at bekæmpe kriminalitet mere effektivt.

Hele samfundet og økonomien kan drage fordel af den kunstige intelligens. Det er en strategisk teknologi, som udvikles og anvendes i hastigt stigende grad over hele verden. Ikke desto mindre medfører kunstig intelligens også nye udfordringer for fremtidens arbejde, og den rejser en række juridiske og etiske spørgsmål.

For at imødegå disse udfordringer og få det bedste ud af de muligheder, som kunstig intelligens giver, offentliggjorde Kommissionen en europæisk strategi¹ i april 2018. Strategien sætter mennesket i centrum for udviklingen af kunstig intelligens – **menneskecentreret kunstig intelligens**. Det er en trestrengt tilgang til at fremme EU's teknologiske og industrielle kapacitet og indførelsen af kunstig intelligens på tværs af de økonomiske sektorer, forberede samfundet på socioøkonomiske forandringer og sikre en hensigtsmæssig etisk og juridisk ramme.

For at strategien for kunstig intelligens kan give resultater, **udviklede Kommissionen i samarbejde med medlemsstaterne en koordineret plan for kunstig intelligens²**, som blev præsenteret i december 2018, om at skabe synergier, samle data – råmateriale for mange applikationer baseret på kunstig intelligens – og øge fælles investeringer. Målet er at fremme samarbejde over landegrænser og mobilisere alle aktører til at øge de offentlige og private investeringer til **mindst 20 milliarder EUR** om året i løbet af det næste årti³. Kommissionen fordoblede sine investeringer i kunstig intelligens i Horisont 2020 og planlægger at investere 1 milliard EUR årligt fra Horisont Europa og programmet for et digitalt Europa. Målet er at støtte især fælles dataområder inden for sundhed, transport og produktion samt store forsøgsfaciliteter såsom intelligente hospitaler og infrastrukturer for automatiserede køretøjer samt en strategisk forskningsdagsorden.

For at gennemføre en sådan fælles strategisk forsknings-, innovations- og etableringsdagsorden har Kommissionen intensiveret sin **dialog med alle relevante interessenter** fra erhvervslivet, forskningsinstitutter og offentlige myndigheder. Det nye program for et digitalt Europa vil også være afgørende i indsatsen for at gøre kunstig intelligens tilgængelig for små og mellemstore virksomheder i alle medlemsstater gennem digitale innovationsknudepunkter, øget afprøvning og forsøgsfaciliteter, dataområder og uddannelsesprogrammer.

Europas etiske tilgang til kunstig intelligens tager udgangspunkt i dets ry for at levere sikre produkter af høj kvalitet, og tilgangen styrker dermed borgernes tillid til den digitale udvikling og sigter mod at skabe en konkurrencefordel for europæiske virksomheder inden for kunstig intelligens. Formålet med denne meddelelse er at iværksætte en omfattende pilotfase

¹ COM(2018) 237.

² COM(2018) 795.

³ For at opnå dette mål foreslog Kommissionen, at Unionen i løbet af den næste programmeringsperiode 2021-2027 afsætter mindst 1 milliard EUR om året til finansiering via Horisont Europa og programmet for et digitalt Europa af investeringer i kunstig intelligens.

med bredest mulig deltagelse af interessenter for at afprøve den praktiske gennemførelse af de etiske retningslinjer for udvikling og anvendelse af kunstig intelligens.

2. OPBYGNING AF TILLID TIL MENNESKECENTRERET KUNSTIG INTELLIGENS

Den europæiske strategi for kunstig intelligens og den koordinerede plan gør det klart, at **tillid er en forudsætning for at sikre en menneskecentreret tilgang til kunstig intelligens**: Kunstig intelligens er ikke et mål i sig selv, men derimod et værktøj, der skal hjælpe mennesker med at nå det ultimative mål, nemlig at øge vores velfærd. For at opnå dette **bør pålideligheden af kunstig intelligens sikres**. De værdier, vores samfund bygger på, skal integreres fuldt ud i den måde, hvorpå kunstig intelligens udvikles.

Unionen bygger på **værdierne respekt for den menneskelige værdighed, frihed, demokrati, ligestilling, retsstaten og respekt for menneskerettighederne**, herunder rettigheder for personer, der tilhører mindretal⁴. Disse værdier er fælles for alle medlemsstaternes samfund, der er præget af pluralisme, ikkediskrimination, tolerance, retfærdighed, solidaritet og lighed. Derudover samler **EU's charter om grundlæggende rettigheder** – i en enkelt tekst – de personlige, borgerlige, politiske, økonomiske og sociale rettigheder, som borgerne i EU er omfattet af.

EU har en **stærk lovgivningsramme**, der vil sætte den globale standard for menneskecentreret kunstig intelligens. Databeskyttelsesforordningen (GDPR) garanterer en høj standard for beskyttelse af personoplysninger og kræver gennemførelse af foranstaltninger til sikring af indbygget databeskyttelse og standarddatabeskyttelse⁵. Forordningen om fri udveksling af andre data end personoplysninger fjerner hindringer for fri udveksling af andre data end personoplysninger og sikrer behandling af alle datakategorier overalt i Europa. Den nyligt vedtagne forordning om cybersikkerhed vil bidrage til at styrke tilliden til den digitale verden, og den foreslåede e-databeskyttelsesforordning⁶ har det samme mål.

Kunstig intelligens rejser imidlertid nye udfordringer, da den gør det muligt for maskiner at "lære" og træffe og gennemføre beslutninger uden menneskelig indgriben. Om ikke så længe vil denne form for funktionalitet blive standard i mange typer varer og tjenester, fra smartphones til automatiserede biler, robotter og onlineapplikationer. Beslutninger taget ved hjælp af algoritmer kan dog være et resultat af data, som er ufuldstændige og derfor ikke er pålidelige, de kan være manipulerede via cyberangreb, være påvirket af bias eller simpelthen forkerte. Uigennemtænkt anvendelse af teknologien, mens den stadig udvikler sig, kan derfor føre til problematiske resultater og gøre borgerne tilbageholdende med at acceptere eller anvende den.

Teknologien for kunstig intelligens bør i stedet udvikles på en måde, der sætter mennesket i centrum og dermed fortjener borgernes tillid. Det betyder, at applikationer med kunstig intelligens ikke kun skal være i overensstemmelse med lovgivningen, men også skal

⁴ Derudover er EU part i FN's konvention om rettigheder for personer med handicap.

⁵ Forordning (EU) 2016/679. Databeskyttelsesforordningen (GDPR) garanterer fri udveksling af personoplysninger inden for Unionen. Den indeholder bestemmelser om beslutningsprocessen, der alene er baseret på automatisk behandling, herunder profilering. De berørte personer har ret til at blive informeret om, at automatisk beslutningstagning findes, samt modtage relevante oplysninger om den logik, der ligger bag automatisk beslutningstagning, om betydningen af behandlingen, og hvilke konsekvenser den kan have for dem. I sådanne tilfælde har de også ret til menneskelig indgriben, at give udtryk for deres synspunkter og at anfægte afgørelsen.

⁶ COM(2017) 10.

overholde de etiske principper og sikre, at deres gennemførelse undgår utilsigtet skade. Mangfoldighed i køn, race eller etnisk oprindelse, religion eller tro, handicap og alder bør sikres på alle niveauer af udviklingen af kunstig intelligens. Kunstig intelligens-applikationer bør give borgerne indflydelse og respektere deres grundlæggende rettigheder. De bør sigte mod at styrke folks evner, ikke erstatte dem, og også gøre det muligt for personer med handicap at få adgang til dem.

Derfor er der et behov for **etiske retningslinjer**, der bygger på den eksisterende reguleringsramme, og som skal anvendes af udviklere, leverandører og brugere af kunstig intelligens på det indre marked, så der skabes lige etiske vilkår for alle medlemsstater. Kommissionen har derfor nedsat en **ekspertgruppe på højt niveau vedrørende kunstig intelligens**⁷, som repræsenterer en bred vifte af interessenter og har fået til opgave at udarbejde en række etiske retningslinjer for kunstig intelligens samt en række henstillinger om bredere politik for kunstig intelligens. Samtidig blev **den europæiske alliance vedrørende kunstig intelligens**⁸ oprettet. Alliancen er en åben multiinteressentplatform med over 2 700 medlemmer, som er oprettet for at give et bredere input til arbejdet i ekspertgruppen på højt niveau vedrørende kunstig intelligens.

Ekspertgruppen på højt niveau vedrørende kunstig intelligens offentliggjorde et første udkast til de etiske retningslinjer i december 2018. Efter en **høring af interessenter**⁹ og **møder med repræsentanter fra medlemsstaterne**¹⁰ forelagde ekspertgruppen vedrørende kunstig intelligens et revideret dokument for Kommissionen i marts 2019. Generelt har interessenter i deres feedback indtil nu set med tilfredshed på den praktiske karakter af retningslinjerne og den konkrete vejledning, som de giver udviklere, leverandører og brugere af kunstig intelligens om, hvordan man kan sikre pålidelighed.

2.1. Retningslinjer for pålidelig kunstig intelligens udarbejdet af ekspertgruppen på højt niveau vedrørende kunstig intelligens

Retningslinjerne udarbejdet af ekspertgruppen på højt niveau vedrørende kunstig intelligens, som denne meddelelse henviser til¹¹, bygger især på arbejdet udført af Den Europæiske Gruppe vedrørende Etik inden for Naturvidenskab og Ny Teknologi og Den Europæiske Unions Agentur for Grundlæggende Rettigheder.

Retningslinjerne fastslår, at for at opnå "pålidelig kunstig intelligens" er tre komponenter nødvendige: 1) den skal være i overensstemmelse med loven, 2) den skal opfylde etiske principper, og 3) den skal være robust.

Baseret på disse tre komponenter og de europæiske værdier beskrevet i afsnit 2 identificerer retningslinjerne syv centrale krav, som kunstig intelligens-applikationer skal opfylde for at

⁷ <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

⁸ <https://ec.europa.eu/digital-single-market/en/european-ai-alliance>

⁹ Høringen resulterede i bemærkninger fra 511 organisationer, foreninger, virksomheder, forskningsinstitutter, enkeltpersoner og andre. Et resumé af denne feedback kan findes på: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ Ekspertgruppens arbejde blev positivt modtaget af medlemsstaterne sammen med Rådets konklusioner, der blev vedtaget den 18. februar 2019, blandt andet med hensyn til den kommende offentliggørelse af de etiske retningslinjer og støtte til Kommissionens bestræbelser på bringe en europæisk etisk tilgang frem på den globale scene: <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/da/pdf>

¹¹ <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

blive betragtet som pålidelige. Retningslinjerne indeholder også en evalueringsliste, der gør det lettere at kontrollere, om kravene er opfyldt.

De syv centrale krav er følgende:

- Menneskelige aktiviteter og tilsyn udført af mennesker
- Teknologisk robusthed og sikkerhed
- Privatlivets fred og datastyring
- Gennemsigtighed
- Mangfoldighed, ikkediskrimination og retfærdighed
- Social og miljømæssig velfærd
- Ansvarlighed

Selv om disse krav skal gælde for alle kunstig intelligens-systemer i forskellige sammenhænge og erhverv, bør den specifikke kontekst, hvori de anvendes, tages i betragtning ved deres konkrete og forholdsmæssige gennemførelse, idet der tages udgangspunkt i en virkningsbaseret tilgang. Det er f.eks. langt mindre farligt, hvis en applikation med kunstig intelligens foreslår en upassende bog, end hvis den fejldiagnosticerer kræft, hvorfor den kan underkastes et mindre stringent tilsyn, når den indgår i førstnævnte kontekst.

Retningslinjerne, som ekspertgruppen på højt niveau om kunstig intelligens har udarbejdet, er ikke bindende og skaber dermed ikke nye juridiske forpligtelser. Der er dog mange eksisterende (og ofte anvendelses- eller områdespecifikke) EU-retlige bestemmelser, som allerede afspejler et eller flere af disse centrale krav, f.eks. sikkerhed, beskyttelse af personoplysninger, miljøbeskyttelsesregler samt regler om beskyttelse af privatlivets fred.

Kommissionen glæder sig over arbejdet i ekspertgruppen på højt niveau vedrørende kunstig intelligens og betragter det som et værdifuldt input til den politiske beslutningstagning.

2.2. Centrale krav til pålidelig kunstig intelligens

Kommissionen støtter følgende centrale krav til pålidelig kunstig intelligens, som er baseret på europæiske værdier. Den tilskynder interessenterne til at anvende kravene og afprøve den evalueringsliste, der gør kravene operationelle, for at skabe det rette tillidsmiljø for en vellykket udvikling og anvendelse af kunstig intelligens. Kommissionen glæder sig over feedback fra interessenter med henblik på at kunne vurdere, om denne evalueringsliste i henhold til retningslinjerne kræver yderligere justeringer.

I. Menneskelige aktiviteter og tilsyn udført af mennesker

Kunstig intelligens-systemer bør støtte enkeltpersoner i at træffe bedre og mere informerede valg i overensstemmelse med deres mål. Disse systemer bør fungere som katalysatorer for et blomstrende og retfærdigt samfund ved at støtte menneskelige aktiviteter og **grundlæggende rettigheder** og ikke mindske, begrænse eller vildlede menneskers autonomi. **Brugerens** overordnede **velbefindende** bør være af central betydning for systemets funktionalitet.

Tilsyn udført af mennesker hjælper med at sikre, at et kunstig intelligens-system ikke

underminerer den menneskelige autonomi eller skaber andre negative virkninger. Afhængigt af det specifikke kunstige intelligens-system og dets anvendelsesområde bør passende grader af **kontrollforanstaltninger** sikres¹², herunder systemernes tilpasningsevne, nøjagtighed og forklarlighed. **Tilsyn** kan opnås ved hjælp af forvaltningsmekanismer som f.eks. at sikre en tilgang med enten menneskelig indgriben (human-in-the-loop), menneskelig overvågning (human-on-the-loop) eller menneskeligt tilsyn (human-in-command).¹³ Det skal sikres, at offentlige myndigheder har mulighed for at udøve deres tilsynsbeføjelser i overensstemmelse med deres mandat. Jo mindre tilsyn, der kan udføres af et menneske over et kunstigt intelligens-system, jo mere omfattende afprøvning og strengere styring vil alt andet lige være påkrævet.

II. Teknologisk robusthed og sikkerhed

Pålidelig kunstig intelligens kræver algoritmer, der er sikre, stabile og robuste nok til at håndtere fejl eller uoverensstemmelser i alle livscyklusfaser af kunstig intelligens-systemet samt håndtere fejlagtige resultater på tilfredsstillende vis. Kunstige intelligens-systemer skal være **pålidelige** og sikre nok til at være **modstandsdygtige** over for både åbenlyse angreb og mere subtile forsøg på selv at manipulere data eller algoritmer, og de skal sikre en **nødplan** i tilfælde af problemer. Deres beslutninger skal være **nøjagtige** eller i det mindste korrekt afspejle deres grad af nøjagtighed, og deres resultater bør være **reproducerbare**.

Derudover bør kunstige intelligens-systemer integrere mekanismer for beskyttelse og indbygget sikkerhed for at garantere, at de er **beviseligt sikre** på alle trin, idet der tages hensyn til alle berørte parter fysiske og mentale sikkerhed. Dette omfatter minimering og om muligt reparation af utilsigtede konsekvenser eller fejl i systemets drift. Der bør indføres procedurer til afklaring og vurdering af potentielle risici i forbindelse med anvendelsen af kunstige intelligens-systemer på tværs af forskellige anvendelsesområder.

III. Privatlivets fred og datastyring

Privatlivets fred og **databeskyttelse** skal garanteres i **alle livscyklusfaser** for et kunstigt intelligens-system. Digitale registre over menneskelig adfærd kan give kunstige intelligens-systemer mulighed for at udlede ikke blot enkeltpersoners præferencer, alder og køn, men også deres seksuelle orientering samt religiøse eller politiske overbevisning. For at gøre det muligt for enkeltpersoner at stole på databehandlingen skal det sikres, at de har fuld kontrol over deres egne data, og at disse data ikke vil blive brugt til at skade eller diskriminere dem.

Ud over at beskytte privatlivets fred og personoplysninger skal kravene være opfyldt for at sikre kunstige intelligens-systemer af høj kvalitet. Kvaliteten af de anvendte datasæt er af

¹² Databeskyttelsesforordningen giver enkeltpersoner ret til ikke at blive genstand for en beslutning, der alene er baseret på automatisk databehandling, når dette har retsvirkning for brugerne eller på tilsvarende vis i væsentlig grad berører dem (artikel 22 i databeskyttelsesforordningen).

¹³ Menneskelig indgriben refererer til menneskelig indgriben i enhver beslutningsproces i systemet, hvilket i mange tilfælde hverken er muligt eller ønskeligt. Menneskelig overvågning refererer til muligheden for menneskelig indgriben i systemets designcyklus og overvågningen af systemets drift. Menneskeligt tilsyn refererer til muligheden for at føre tilsyn med systemets aktiviteter (herunder dets bredere økonomiske, sociale, juridiske og etiske virkninger) samt evnen til at afgøre, hvornår og hvordan systemet skal anvendes i en given situation. Dette kan omfatte en beslutning om ikke at bruge et kunstigt intelligens-system i en given situation, at fastsætte niveauer for den menneskelige skønsmargin ved brug af systemet eller at sikre muligheden for at tilsidesætte en afgørelse truffet af systemet.

afgørende betydning for systemernes ydeevne. Når der indsamles data, kan disse afspejle socialt konstrueret forudindtagethed eller være behæftet med unøjagtigheder, fejl og mangler. Dette skal håndteres, inden et kunstig intelligens-system bliver oplært med et givet datasæt. Derudover skal **integriteten** af dataene sikres. Processer og datasæt, der anvendes, skal testes og dokumenteres på hvert trin, som f.eks. planlægning, oplæring, afprøvning og etablering. Dette bør også gælde for systemer med kunstig intelligens, som ikke er udviklet internt, men derimod er erhvervet andetsteds. Endelig skal **adgangen** til data styres og kontrolleres på passende vis.

IV. Gennemsigtighed

Sporbarheden af alle systemer med kunstig intelligens skal sikres; det er vigtigt at logge på og dokumentere både de beslutninger, der er truffet af systemerne, og hele processen (herunder en beskrivelse af dataindsamling og -mærkning samt en beskrivelse af den anvendte algoritme), som førte til beslutningen. I forbindelse hermed bør det så vidt muligt sikres, at den algoritmiske beslutningsproces kan **forklares** på en måde, som er tilpasset de involverede personer. Igangværende forskning bør følges med henblik på at udvikle forklaringsmekanismerne. Desuden bør der gives forklaringer på, i hvilket omfang et kunstig intelligens-system påvirker og former den organisatoriske beslutningsproces, systemets designvalg samt begrundelsen for at anvende det (hvilket dermed sikrer ikke blot data- og systemgennemsigtighed, men også gennemsigtighed i forretningsmodellen).

Endelig er det vigtigt at **informere** de berørte interessenter om systemets kapacitet og begrænsninger på en hensigtsmæssig måde i forhold til det aktuelle tilfælde. Derudover bør kunstig intelligens-systemer kunne identificeres som sådanne og dermed sikre, at brugerne ved, at de interagerer med et system med kunstig intelligens, samt hvilke personer der er ansvarlige for det.

V. Mangfoldighed, ikkediskrimination og retfærdighed

Datasæt, der anvendes af kunstig intelligens-systemer (både til oplæring og drift) kan lide under inddragelsen af utilsigtet historisk forudindtagethed, ufuldstændigheder og dårlige forvaltningsmodeller. Fortsættelsen af en sådan forudindtagethed vil kunne føre til (in)direkte diskrimination. Skader kan også opstå ved forsætlig udnyttelse af (forbrugerens) forudindtagethed eller ved illoyal konkurrence. Derudover kan måden, hvorpå kunstig intelligens-systemer er udviklet (f.eks. måden, hvorpå programmeringskoden for en algoritme er skrevet) også være negativt påvirket af bias. Sådanne problemer bør tages op fra begyndelsen af systemets udvikling.

En anden måde at håndtere disse problemer på er at oprette **forskelligartede designteams** samt mekanismer, der sikrer, at især borgere **deltager** i udviklingen af kunstig intelligens. Det anbefales at høre interessenter, som direkte eller indirekte kan blive berørt af systemet gennem dets livscyklus. Kunstig intelligens-systemer bør tage hele spektret af menneskelige evner, færdigheder og krav i betragtning og sikre tilgængelighed ved at følge en universel designtilgang for at opnå lige adgang for alle personer med handicap.

VI. Social og miljømæssig velfærd

Hvis kunstig intelligens skal være pålidelig, bør indvirkningen på **miljøet og andre følende væsener** tages i betragtning. Ideelt set bør alle mennesker, herunder kommende

generationer, drage fordel af biodiversitet og et beboeligt miljø. Der bør derfor tilskyndes til kunstig intelligens-systemer med bæredygtighed og **miljømæssigt ansvar**. Det samme gælder for kunstig intelligens-løsninger, der vedrører områder af global betydning, som f.eks. FN's mål for bæredygtig udvikling.

Derudover bør indvirkningen af kunstig intelligens-systemer ikke kun tages i betragtning fra et subjektivt perspektiv, men også fra et perspektiv, der ser **samfundet som en helhed**. Anvendelsen af kunstig intelligens-systemer bør nøje overvejes, særlig i situationer, der vedrører den demokratiske proces, herunder meningsdannelse, politisk beslutningstagning eller i forbindelse med valg. Derudover bør også den **sociale indvirkning** af kunstig intelligens tages i betragtning. Kunstig intelligens-systemer kan anvendes til at forbedre sociale færdigheder, men de kan også bidrage til deres forværring.

VII. Ansvarlighed

Der bør indføres mekanismer til at sikre ansvar og ansvarlighed for kunstig intelligens-systemer og deres resultater både før og efter gennemførelse. Muligheden for **audit** af kunstig intelligens-systemerne er afgørende i denne henseende, da evaluering af disse systemer af interne og eksterne auditorer, og tilgængeligheden af sådanne evalueringsrapporter, i høj grad bidrager til teknologiens pålidelighed. Ekstern audit bør især sikres i forbindelse med applikationer, der påvirker grundlæggende rettigheder, herunder sikkerhedskritiske applikationer.

Kunstig intelligens-systemers **potentielle negative virkninger** bør identificeres, evalueres, dokumenteres og minimeres. Anvendelsen af konsekvensevaluering gør denne proces lettere. Disse evalueringer bør være i overensstemmelse med omfanget af de risici, som kunstig intelligens-systemerne udgør. **Kompromiser** mellem kravene er ofte uundgåelige og bør behandles på en rationel og metodisk måde samt være behørigt begrundede. Endelig bør tilgængelige mekanismer sikre **passende erstatning** i tilfælde af urimelige, negative konsekvenser.

2.3. Næste skridt: pilotfase med bredest mulig deltagelse af interessenter

Den første vigtige milepæl i udarbejdelse af retningslinjer for etisk kunstig intelligens er at opnå enighed om disse centrale krav for kunstig intelligens-systemer. Kommissionen sikrer som det næste skridt, at retningslinjerne kan afprøves og gennemføres i praksis.

Til dette formål vil Kommissionen nu iværksætte en målrettet pilotfase, der er udformet med henblik på feedback fra interessenter. Denne øvelse vil især fokusere på evalueringslisten, som ekspertgruppen på højt niveau har udarbejdet for hver af de centrale krav.

Denne indsats vil bestå af to dele: i) en pilotfase for retningslinjer med inddragelse af interessenter, som udvikler eller anvender kunstig intelligens, herunder offentlige myndigheder, og ii) en fortsat interessenthøring og bevidstgørelsesproces på tværs af medlemsstaterne og forskellige interessentgrupper, herunder industri- og servicesektorer:

- i) Fra juni 2019 vil alle interessenter og enkeltpersoner blive inviteret til at teste evalueringslisten og give feedback om, hvordan den kan forbedres. Ekspertgruppen på højt niveau vedrørende kunstig intelligens vil derudover gennemføre en detaljeret undersøgelse med interessenter fra den private og offentlige sektor for at indsamle mere udførlig feedback om, hvordan retningslinjerne kan gennemføres inden for en

lang række anvendelsesområder. Feedback vedrørende retningslinjernes gennemførlighed og anvendelighed vil blive evalueret ved udgangen af 2019.

- ii) Samtidig vil Kommissionen tilrettelægge yderligere opsøgende aktiviteter, hvilket giver repræsentanter fra ekspertgruppen på højt niveau vedrørende kunstig intelligens mulighed for at præsentere retningslinjerne for relevante interessenter i medlemsstaterne, herunder industri- og servicesektorer, samt give disse interessenter en ekstra mulighed for at kommentere på og bidrage til retningslinjerne for kunstig intelligens.

Kommissionen vil tage hensyn til ekspertgruppens arbejde om etik i forbindelse med opkoblet og automatiseret kørsel¹⁴ samt arbejde med EU-finansierede forskningsprojekter om kunstig intelligens og med relevante offentlig-private partnerskaber om gennemførelse af de centrale krav¹⁵. Eksempelvis vil Kommissionen, i samarbejde med medlemsstaterne, støtte udviklingen af en fælles database over sundhedsbilleder, som i første omgang blev anvendt til de mest udbredte kræftformer, for at algoritmer kan trænes til at diagnosticere symptomer med meget høj nøjagtighed. Samarbejdet mellem Kommissionen og medlemsstaterne muliggør ligeledes et stigende antal testkorridorer på tværs af grænserne til afprøvning af opkoblede og automatiserede køretøjer. Retningslinjerne bør anvendes i forbindelse med disse projekter og testes, og resultaterne vil indgå i evalueringsprocessen.

Pilotfasen og interessenthøringen vil få gavn af bidraget fra den europæiske alliance vedrørende kunstig intelligens og AI4EU, som er en on-demand platform for kunstig intelligens. AI4EU-projektet¹⁶, som blev lanceret i januar 2019, samler algoritmer, værktøjer, datasæt og tjenester for at hjælpe organisationer, særlig små og mellemstore virksomheder, med at gennemføre kunstig intelligens-løsninger. Den europæiske alliance vedrørende kunstig intelligens vil sammen med AI4EU fortsat mobilisere kunstig intelligens-økosystemet over hele Europa, også med henblik på at afprøve retningslinjerne for kunstig intelligens og fremme respekten for menneskecentreret kunstig intelligens.

I begyndelsen af 2020 vil gruppen på højt niveau for kunstig intelligens tage udgangspunkt i evalueringen af den feedback, der er modtaget i pilotfasen og gennemgå samt ajourføre retningslinjerne. Baseret på denne gennemgang og den erhvervede erfaring vil Kommissionen evaluere resultatet og foreslå eventuelle næste skridt.

Etisk kunstig intelligens er en win-win-løsning. Det er ikke kun vigtigt i sig selv at sikre respekten for grundlæggende værdier og rettigheder; det gør det også lettere for offentligheden at acceptere kunstig intelligens og øger den konkurrencemæssige fordel for europæiske virksomheder med kunstig intelligens ved at etablere et menneskecentreret og pålideligt kunstig intelligens-brand, der er kendt for etiske og sikre produkter. Dette bygger mere generelt på de europæiske virksomheders gode ry for at levere sikre produkter af høj kvalitet. Pilotfasen vil bidrage til at sikre, at kunstig intelligens-produkter lever op til dette løfte.

2.4. På vej mod internationale etiske retningslinjer for kunstig intelligens

¹⁴ Se Kommissionens meddelelse om opkoblet og automatiseret mobilitet, COM(2018) 283.

¹⁵ Inden for rammerne af Den Europæiske Forsvarsfond vil Kommissionen også udarbejde specifikke etiske retningslinjer for evaluering af projektforslag inden for kunstig intelligens med henblik på forsvar.

¹⁶ <https://ec.europa.eu/digital-single-market/en/news/artificial-intelligence-ai4eu-project-launches-1-january-2019>

Internationale drøftelser for etik vedrørende kunstig intelligens er intensiveret, efter at Japans G7-formandskab i 2016 satte emnet højt på dagsordenen. I betragtning af de internationale forbindelser mellem udvikling af kunstig intelligens med hensyn til data-cirkulation, algoritmisk udvikling og forskning vil **Kommissionen fortsætte sine bestræbelser på at bringe Unionens tilgang frem på den globale scene og opnå enighed om en menneskecentreret kunstig intelligens**¹⁷.

Arbejdet, som ekspertgruppen på højt niveau vedrørende kunstig intelligens har udført, og mere specifikt listen over krav samt processen med inddragelse af interessenter, giver Kommissionen yderligere værdifuldt input, som bidrager til de internationale drøftelser. Den Europæiske Union kan have en ledelsesrolle i udviklingen af internationale retningslinjer og om muligt en tilhørende evalueringsmekanisme.

Derfor vil Kommissionen:

styrke samarbejdet med ligesindede partnere ved at:

- udforske, i hvor høj grad konvergens med tredjelandes udkast til etiske retningslinjer kan opnås (f.eks. Japan, Canada, Singapore) og, ud fra denne gruppe af ligesindede lande, forberede en bredere debat, der støttes af foranstaltninger til gennemførelse af partnerskabsinstrumentet for samarbejde med tredjelands¹⁸, samt
- udforske, hvordan virksomheder fra tredjelands og internationale virksomheder kan bidrage til retningslinjernes "pilotfase" gennem afprøvning og validering.

Fortsætte med at spille en aktiv rolle i internationale drøftelser og initiativer ved at:

- bidrage til multilaterale fora som G7 og G20
- indgå i dialog med tredjelands og tilrettelægge bi- og multilaterale møder for at skabe enighed om menneskecentreret kunstig intelligens
- bidrage til relevante standardiseringsaktiviteter i internationale standardudviklende organisationer for at fremme denne vision samt
- styrke indsamling og formidling af indsigt i offentlige politikker gennem samarbejde med relevante internationale organisationer.

3. KONKLUSIONER

EU bygger på en række grundlæggende værdier og har på baggrund af disse anlagt en stærk og balanceret lovgivningsramme. På grundlag af denne eksisterende lovgivningsramme er

¹⁷ Unionens højtstående repræsentant for udenrigsanliggender og sikkerhedspolitik vil med støtte fra Kommissionen tage udgangspunkt i høringer i FN, Global Tech Panel og andre multilaterale fora og især koordinere forslag til løsning af de komplekse sikkerhedsudfordringer, der er forbundet hermed.

¹⁸ Europa-Parlamentets og Rådets forordning (EU) nr. 234/2014 af 11. marts 2014 om oprettelse af et partnerskabsinstrument for samarbejde med tredjelands (EUT L 77 af 15.3.2014, s. 77). Eksempelvis vil det planlagte projekt om "En international alliance for en menneskecentreret tilgang til kunstig intelligens" fremme fælles initiativer med ligesindede partnere med henblik på at fremme etiske retningslinjer og vedtage fælles principper og operationelle konklusioner. Det vil gøre det muligt for EU og ligesindede lande at drøfte operationelle konklusioner, der følger af de etiske retningslinjer om kunstig intelligens, som er foreslået af ekspertgruppen med henblik på at opnå en fælles tilgang. Derudover vil projektet gøre det muligt at overvåge anvendelsen af teknologien for kunstig intelligens på den globale scene. Endelig planlægges det som led i projektet at organisere offentlige diplomatiske aktiviteter, der ledsager internationale begivenheder, f.eks. af G7, G20 og Organisationen for Økonomisk Samarbejde og Udvikling.

der, fordi det er så nyt og på grund af de særlige udfordringer, som denne teknologi medfører, behov for etiske retningslinjer for udvikling og anvendelse af kunstig intelligens. Kun hvis kunstig intelligens udvikles og anvendes på en måde, der respekterer udbredte etiske værdier, kan den betragtes som pålidelig.

Med henblik herpå glæder Kommissionen sig over de synspunkter, som ekspertgruppen på højt niveau vedrørende kunstig intelligens har fremlagt. På grundlag af de krav, som er centrale for, at kunstig intelligens kan betragtes som pålidelig, vil Kommissionen nu iværksætte en målrettet pilotfase for at sikre, at de resulterende etiske retningslinjer for udvikling og anvendelse af kunstig intelligens kan gennemføres i praksis. Kommissionen vil også arbejde for at skabe en bred enighed i samfundet om menneskecentreret kunstig intelligens, herunder med alle involverede interessenter og vores internationale partnere.

Den etiske dimension af kunstig intelligens er ikke en luksusfunktion eller add-on: den skal være en integreret del af udviklingen af kunstig intelligens. Ved at stræbe mod menneskecentreret kunstig intelligens, der er baseret på tillid, sikrer vi respekten for vores grundlæggende samfundsmæssige værdier og skaber et særligt varemærke for Europa og dets industri som leder inden for banebrydende kunstig intelligens, der anses som pålidelig i hele verden.

For at sikre den etiske udvikling af kunstig intelligens i Europa i en bredere sammenhæng arbejder Kommissionen hen imod en samlet tilgang, herunder især de følgende tiltag, som skal gennemføres senest i tredje kvartal af 2019:

- Kommissionen vil begynde at lancere en række **netværk af førende kunstig intelligens-forskningscentre** gennem Horisont 2020. Den vil udvælge op til fire netværk med fokus på videnskabelige eller teknologiske udfordringer såsom forklarlighed og avanceret interaktion menneske-maskine, som er de vigtigste forudsætninger for pålidelig kunstig intelligens.
- Den vil begynde at oprette netværk af **digitale innovationsknudepunkter**¹⁹ med fokus på kunstig intelligens i produktion og big data.
- I samarbejde med medlemsstater og interessenter vil Kommissionen starte de indledende drøftelser for at udvikle og gennemføre **en model for dataudveksling og optimal udnyttelse af fælles dataområder** med særligt fokus på transport, sundhedspleje og industriel produktion.²⁰

Derudover arbejder Kommissionen på en rapport om de udfordringer, som kunstig intelligens udgør for sikkerheds- og ansvarsrammerne samt et vejledende dokument om gennemførelsen af direktivet om produktansvar²¹. Samtidig vil det europæiske fællesforetagende for højtydende databehandling (EuroHPC)²² udvikle den næste generation af supercomputere, da databehandlingskapacitet er af afgørende betydning for databehandling og oplæring af kunstig intelligens, og Europa er nødt til at beherske hele den digitale værdikæde. Det aktive partnerskab med medlemsstaterne og industrien om mikroelektroniske komponenter og

¹⁹ <http://s3platform.jrc.ec.europa.eu/digital-innovation-hubs>

²⁰ De nødvendige ressourcer vil blive mobiliseret fra Horisont 2020 (i henhold til hvilke der afsættes næsten 1,5 milliarder EUR til kunstig intelligens for perioden 2018-2020) og dennes planlagte efterfølger Horisont Europa, som er den digitale del af Connecting Europe-faciliteten og især det fremtidige program for et digitalt Europa. Projektet vil også trække på ressourcer fra den private sektor og medlemsstaternes programmer.

²¹ Se Kommissionens meddelelse om kunstig intelligens for Europa, COM(2018) 237.

²² <https://eurohpc-ju.europa.eu>

systemer (ECSEL)²³ samt det europæiske processor-initiativ²⁴ vil bidrage til udviklingen af processorteknologi med lavt effektforbrug til pålidelig og sikker højtydende databehandling og edge computing.

Ligesom arbejdet med etiske retningslinjer for kunstig intelligens bygger alle disse initiativer på **tæt samarbejde mellem alle berørte interessenter**, medlemsstater, industrien, samfundsaktører og borgere. Generelt set viser Europas tilgang til kunstig intelligens, hvordan økonomiske konkurrenceevne og samfundets tillid skal tage udgangspunkt i de samme grundlæggende værdier og gensidigt styrke hinanden.

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu