



Brussel, 8.4.2019
COM(2019) 168 final

**MEDEDELING VAN DE COMMISSIE AAN HET EUROPEES PARLEMENT, DE
RAAD, HET EUROPEES ECONOMISCH EN SOCIAAL COMITÉ EN HET COMITÉ
VAN DE REGIO'S**

Vertrouwen kweken in mensgerichte kunstmatige intelligentie

1. INLEIDING – DE EUROPESE STRATEGIE VOOR KI

Kunstmatige intelligentie (KI) kan van de wereld een betere plaats maken, met een betere gezondheidszorg, een lager energieverbruik, veiliger auto's, en een efficiënter gebruik van water en natuurlijke hulpbronnen door landbouwers. KI kan worden ingezet om milieu- en klimaatveranderingen te voorspellen en om financiële risico's beter te beheren, en maakt een op onze behoeften afgestemde productie met minder afval mogelijk. Voorts kan KI fraude en virtuele dreigingen helpen opsporen en rechtshandavingsinstanties in staat stellen om criminaliteit efficiënter te bestrijden.

KI kan de hele samenleving en de economie ten goede komen. Het is een strategische technologie die wereldwijd in snel tempo wordt ontwikkeld en gebruikt. Maar KI brengt ook nieuwe uitdagingen met zich mee voor de toekomst van werk, en doet juridische en ethische vragen rijzen.

Om deze uitdagingen het hoofd te bieden en de mogelijkheden van KI optimaal te benutten, heeft de Commissie in april 2018 een Europese strategie¹ gepubliceerd. In de strategie wordt benadrukt dat de mens centraal moet staan bij de ontwikkeling van KI (**mensgerichte KI**). Er wordt een driedelige aanpak voorgesteld, die erin bestaat de technologische en industriële capaciteit van de EU te vergroten en het gebruik van KI op te voeren in de hele economie, voorbereidingen te treffen voor sociaal-economische veranderingen en te zorgen voor een passend ethisch en juridisch kader.

Ter verwezenlijking van de KI-strategie **heeft de Commissie** in december 2018 een **gecoördineerd plan inzake KI²** voorgesteld, dat **samen met de lidstaten** is opgesteld om synergieën te ontwikkelen, data – de grondstof voor veel KI-toepassingen – te bundelen en gezamenlijke investeringen aan te moedigen. Het doel is om grensoverschrijdende samenwerking te bevorderen en alle actoren te mobiliseren om de publieke en private investeringen in de komende tien jaar te verhogen tot **ten minste 20 miljard EUR per jaar³**. De Commissie heeft haar investeringen in KI in het kader van Horizon 2020 verdubbeld en is van plan om jaarlijks 1 miljard EUR uit Horizon Europa en het programma Digitaal Europa te investeren ter ondersteuning van met name gemeenschappelijke dataruimten voor gezondheid, vervoer en fabricage, en grote experimenteerfaciliteiten, zoals slimme ziekenhuizen en infrastructuur voor geautomatiseerde voertuigen en een strategische onderzoeksagenda.

Met het oog op de uitvoering van een dergelijke gemeenschappelijke strategische agenda voor onderzoek, innovatie en uitrol heeft de Commissie de **dialogoog met alle relevante belanghebbenden** uit onderzoeksinstituten, overheidsinstanties en het bedrijfsleven geïntensiveerd. Daarnaast wordt het nieuwe programma Digitaal Europa van cruciaal belang om KI in alle lidstaten beschikbaar te maken voor kleine en middelgrote ondernemingen via digitale innovatiehubs, versterkte test- en experimenteerfaciliteiten, dataruimten en opleidingsprogramma's.

Europa kan, voortbouwend op zijn reputatie inzake veilige en hoogwaardige producten, via een ethische benadering van KI het vertrouwen van de burger in de digitale ontwikkeling versterken en Europese KI-bedrijven een concurrentievoordeel bezorgen. Deze mededeling heeft als doel de aanzet te geven tot een uitgebreide proeffase met een zo breed mogelijke

¹ COM(2018) 237.

² COM(2018) 795.

³ Met dit doel voor ogen heeft de Commissie voorgesteld dat de Unie in de volgende programmeringsperiode (2021-2027) jaarlijks ten minste 1 miljard EUR aan middelen van Horizon Europa en Digitaal Europa toewijst voor investeringen in KI.

betrokkenheid van belanghebbenden om de praktische uitvoering van ethische richtsnoeren voor de ontwikkeling en het gebruik van KI te testen.

2. VERTROUWEN KWEKEN IN MENSGERICHTE KI

De Europese KI-strategie en het gecoördineerde plan maken duidelijk dat **vertrouwen een noodzakelijke voorwaarde is voor een mensgerichte benadering van KI**: KI is geen doel op zich, maar een instrument dat ten dienste staat van mensen, met als uiteindelijke doel het menselijke welzijn te vergroten. Om dit te bereiken **moet de betrouwbaarheid van KI gewaarborgd zijn**. De waarden waarop onze samenlevingen zijn gegrondvest, moeten ten volle worden geïntegreerd in de manier waarop KI zich ontwikkelt.

De **waarden** waarop de Unie berust, zijn **eerbied voor de menselijke waardigheid, vrijheid, democratie, gelijkheid, de rechtsstaat en eerbiediging van de mensenrechten**, waaronder de rechten van personen die tot minderheden behoren⁴. Alle lidstaten hebben deze waarden gemeen in samenlevingen die worden gekenmerkt door pluralisme, non-discriminatie, verdraagzaamheid, rechtvaardigheid, solidariteit en gelijkheid. Daarnaast verenigt het **EU-Handvest van de grondrechten** alle persoonlijke, burgerlijke, politieke, economische en sociale rechten die mensen in de EU hebben, in één tekst.

De EU beschikt over een **sterk regelgevingskader** dat de wereldwijde norm voor mensgerichte KI zal bepalen. De algemene verordening gegevensbescherming garandeert een hoog niveau van bescherming van persoonsgegevens en vereist dat maatregelen worden genomen ter bescherming van gegevens door ontwerp en door standaardinstellingen⁵. De verordening betreffende het vrije verkeer van niet-persoonsgebonden gegevens neemt belemmeringen voor het vrije verkeer van dergelijke gegevens weg en zorgt ervoor dat alle categorieën gegevens overal in Europa kunnen worden verwerkt. De onlangs aangenomen cyberbeveiligingsverordening zal het vertrouwen in de onlinewereld helpen versterken, wat ook het doel is van de voorgestelde e-privacyverordening⁶.

Toch brengt KI nieuwe uitdagingen mee omdat zij machines in staat stelt om te "leren" en beslissingen te nemen en uit te voeren zonder menselijk ingrijpen. Binnen afzienbare tijd zal dit soort functionaliteit standaard worden toegepast in diverse goederen en diensten, van smartphones tot geautomatiseerde voertuigen, robots en onlinetoepassingen. De kans bestaat echter dat door algoritmes genomen beslissingen gebaseerd zijn op onvolledige en dus onbetrouwbare gegevens, dat ze gemanipuleerd worden door cyberaanvallers, of dat ze vertekend of gewoonweg verkeerd zijn. Het ondoordacht toepassen van de zich ontwikkelende technologie kan problematische gevolgen hebben en burgers terughoudend maken om de technologie te aanvaarden of te gebruiken.

⁴ Tevens is de EU partij bij het VN-Verdrag van de Verenigde Naties inzake de rechten van personen met een handicap.

⁵ Verordening (EU) 2016/679. De algemene verordening gegevensbescherming (AVG) waarborgt het vrije verkeer van persoonsgegevens binnen de Unie. Zij bevat bepalingen inzake uitsluitend op geautomatiseerde verwerking, met inbegrip van profilering, gebaseerde besluitvorming. De betrokken personen hebben het recht in kennis te worden gesteld van het bestaan van geautomatiseerde besluitvorming en nuttige informatie te ontvangen over de onderliggende logica en over het belang en de verwachte gevolgen van de verwerking voor hen. Zij hebben in dergelijke gevallen ook het recht op menselijke tussenkomst, het recht om hun standpunt kenbaar te maken en het recht om het besluit aan te vechten.

⁶ COM(2017) 10.

In plaats daarvan moet KI-technologie worden ontwikkeld op een manier die de mens centraal stelt en daardoor vertrouwen inboezemt. Dit betekent dat KI-toepassingen niet alleen in overeenstemming moeten zijn met de wet maar ook met ethische beginselen, en dat onbedoelde schade bij het gebruik ervan moet worden voorkomen. In elk stadium van de ontwikkeling van KI moet diversiteit worden gewaarborgd op het gebied van geslacht, ras of etnische afkomst, godsdienst of overtuiging, handicap en leeftijd. KI-toepassingen moeten burgers empoweren en hun grondrechten eerbiedigen. Zij moeten de capaciteiten van mensen vergroten in plaats van ze te vervangen, en zij moeten ook toegankelijk zijn voor mensen met een handicap.

Er is dan ook behoefte aan **ethische richtsnoeren** die voortbouwen op het bestaande regelgevingskader en die moeten worden toegepast door ontwikkelaars, leveranciers en gebruikers van KI op de interne markt, zodat in alle lidstaten een gelijk speelveld op ethisch vlak tot stand komt. Daarom heeft de Commissie de door haar opgerichte **deskundigengroep op hoog niveau inzake KI**⁷, die een breed scala aan belanghebbenden vertegenwoordigt, belast met de opstelling van ethische richtsnoeren voor KI en van een reeks aanbevelingen voor een breder KI-beleid. Tegelijkertijd werd een **Europese KI-alliantie**⁸ in het leven geroepen om, als een open multistakeholderplatform met meer dan 2 700 leden, bredere input te leveren voor de werkzaamheden van de deskundigengroep.

In december 2018 heeft de deskundigengroep op hoog niveau inzake KI een eerste ontwerp van de ethische richtsnoeren gepubliceerd. Na een **raadpleging van belanghebbenden**⁹ en **vergaderingen met vertegenwoordigers van de lidstaten**¹⁰ heeft de deskundigengroep KI de Commissie in maart 2019 een herzien document bezorgd. In de tot dusver ontvangen feedback hebben de belanghebbenden zich over het algemeen verheugd getoond over de praktische aard van de richtsnoeren en over de concrete richtlijnen die zij ontwikkelaars, leveranciers en gebruikers van KI geven om betrouwbaarheid te waarborgen.

2.1. Door de deskundigengroep op hoog niveau inzake KI opgestelde richtsnoeren voor betrouwbare KI

De richtsnoeren die door de deskundigengroep op hoog niveau inzake KI zijn opgesteld en waarop deze mededeling betrekking heeft¹¹, bouwen met name voort op de werkzaamheden van de Europese Groep ethiek van de exacte wetenschappen en de nieuwe technologieën en het Bureau voor de grondrechten.

Volgens de richtsnoeren vereist "betrouwbare KI" dat wordt voldaan aan de volgende drie voorwaarden: 1) zij moet in overeenstemming zijn met de wet; 2) zij moet ethische beginselen eerbiedigen en 3) zij moet robuust zijn.

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 In het kader van de raadpleging werden opmerkingen ontvangen van 511 organisaties, verenigingen, ondernemingen, onderzoeksinstituten, individuele personen en andere respondenten. Een overzicht van de ontvangen feedback is te vinden op: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ De lidstaten hebben de werkzaamheden van de deskundigengroep positief ontvangen. In de conclusies van de Raad van 18 februari 2019 is onder meer kennis genomen van de aanstaande publicatie van de ethische richtsnoeren en is steun uitgedrukt voor het streven van de Commissie om wereldwijd een Europese ethische norm te introduceren. <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/nl/pdf>

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

Uitgaande van deze drie voorwaarden en van de in punt 2 beschreven Europese waarden worden in de richtsnoeren zeven essentiële vereisten genoemd waaraan KI-toepassingen moeten voldoen om als betrouwbaar te worden beschouwd. De richtsnoeren bevatten tevens een beoordelingslijst aan de hand waarvan kan worden nagegaan of aan deze eisen is voldaan.

De zeven essentiële vereisten zijn:

- Invloed en toezicht door mensen
- Technische robuustheid en veiligheid
- Privacy en datagovernance
- Transparantie
- Diversiteit, non-discriminatie en billijkheid
- Maatschappelijk en ecologisch welzijn
- Verantwoordingsplicht

Die vereisten zijn bedoeld voor alle KI-systemen in verschillende omgevingen en bedrijfstakken. Met het oog op de concrete en evenredige uitvoering van de systemen moet echter ook rekening worden gehouden met de specifieke context waarin ze worden gebruikt, uitgaande van de mogelijke impact die ze kunnen hebben. Zo is een KI-toepassing die een lezer een ongeschikt boek voorstelt, veel minder gevaarlijk dan een toepassing die een kankerdiagnose mist, en kan het toezicht in het eerste geval dus minder streng zijn.

De door de deskundigengroep op hoog niveau opgestelde richtsnoeren zijn niet bindend en creëren dus geen nieuwe wettelijke verplichtingen. Dat neemt natuurlijk niet weg dat veel bestaande (vaak gebruiks- of domeinspecifieke) bepalingen van het recht van de Unie al betrekking hebben op een of meer van deze essentiële vereisten, bijvoorbeeld veiligheid, gegevensbescherming, privacy of milieubescherming.

De Commissie is ingenomen met de werkzaamheden van de deskundigengroep op hoog niveau inzake KI en stelt prijs op de input van de groep voor de beleidsvorming.

2.2. Essentiële vereisten voor betrouwbare KI

De Commissie steunt de volgende essentiële vereisten voor betrouwbare KI, die gebaseerd zijn op Europese waarden. Zij moedigt de belanghebbenden aan de vereisten toe te passen en de bijbehorende beoordelingslijst te testen, teneinde een klimaat van vertrouwen te scheppen dat bevorderlijk is voor de ontwikkeling en het gebruik van KI. De Commissie ontvangt graag feedback van belanghebbenden om na te gaan of de beoordelingslijst verder moet worden bijgewerkt.

I. Invloed en toezicht door mensen

KI-systemen moeten mensen helpen om betere en beter geïnformeerde keuzes te maken die aansluiten bij hun doelstellingen. Zij moeten een bloeiende en rechtvaardige samenleving mogelijk maken door menselijke handelingen en de **grondrechten** te ondersteunen, zonder de menselijke autonomie te verminderen, in te perken of te misleiden. Het systeem moet zo functioneren dat het algemene **welzijn van de gebruiker** centraal staat.

Menselijk toezicht kan ervoor zorgen dat een KI-systeem de menselijke autonomie niet ondermijnt of andere schadelijke effecten veroorzaakt. Afhankelijk van het specifieke op KI gebaseerde systeem en het toepassingsgebied ervan, moet worden voorzien in passende vormen van **controlemaatregelen**, die onder meer betrekking hebben op de aanpasbaarheid, nauwkeurigheid en verklaarbaarheid van op KI gebaseerde systemen¹². **Toezicht** is mogelijk via governancemechanismen, zoals het volgen van een "human in the loop"-, "human on the loop"- of "human in command"-benadering¹³. Er dient te worden gewaarborgd dat overheidsinstanties in staat zijn om hun toezichtsbevoegdheden uit te oefenen overeenkomstig hun mandaat. Naarmate er minder menselijk toezicht op een KI-systeem mogelijk is, zijn onder voor het overige gelijke omstandigheden uitgebreidere tests en striktere governance nodig.

II. Technische robuustheid en veiligheid

Voor betrouwbare KI zijn algoritmen nodig die veilig, betrouwbaar en robuust genoeg zijn om fouten of inconsistenties tijdens alle fasen van de levenscyclus van het KI-systeem aan te pakken en om adequaat om te gaan met foutieve resultaten. KI-systemen moeten **betrouwbaar** zijn, moeten veilig genoeg zijn om **weerstand** te kunnen bieden tegen zowel openlijke aanvallen als meer subtiele pogingen om data of de algoritmen zelf te manipuleren, en moeten in geval van problemen kunnen terugvallen op een **noodplan**. De beslissingen van de systemen moeten **nauwkeurig** zijn, of ten minste hun niveau van nauwkeurigheid correct weergeven, en de resultaten moeten **reproduceerbaar** zijn.

Voorts moeten in KI-systemen mechanismen voor veiligheid en beveiliging door ontwerp worden geïntegreerd om ervoor te zorgen dat ze bij elke stap **aantoonbaar veilig** zijn, met oog voor de fysieke en mentale veiligheid van alle betrokkenen. Dit omvat de mogelijkheid om onbedoelde gevolgen of fouten in de werking van het systeem tot een minimum te beperken en waar mogelijk ongedaan te maken. Er moet worden voorzien in processen ter verduidelijking en beoordeling van de mogelijke risico's door het gebruik van KI-systemen, in diverse toepassingsgebieden.

III. Privacy en datagovernance

Privacy en **gegevensbescherming** moeten worden gewaarborgd in **alle stadia** van de levenscyclus van een KI-systeem. Uit digitale bestanden van menselijk gedrag kunnen KI-systemen mogelijk niet alleen iemands persoonlijke voorkeur, leeftijd en geslacht afleiden, maar ook zijn of haar seksuele geaardheid en religieuze of politieke opvattingen. Mensen zullen pas vertrouwen hebben in gegevensverwerking als zij erop kunnen rekenen dat zij de volledige controle hebben over hun eigen gegevens en dat hun gegevens niet

¹² Op grond van de algemene verordening gegevensbescherming hebben personen het recht niet te worden onderworpen aan een uitsluitend op geautomatiseerde verwerking gebaseerd besluit waaraan voor hen rechtsgevolgen zijn verbonden of dat hen anderszins in aanmerkelijke mate treft (artikel 22 AVG).

¹³ "Human in the loop" (HITL) verwijst naar menselijke tussenkomst in elke beslissingscyclus van het systeem, wat in veel gevallen niet mogelijk en ook niet wenselijk is. "Human on the loop" (HOTL) verwijst naar de mogelijkheid van menselijke tussenkomst tijdens de ontwerpcyclus van het systeem en toezicht op de werking van het systeem. "Human in command" (HIC) verwijst naar het vermogen om toezicht te houden op het geheel van de activiteiten van het KI-systeem (met inbegrip van de bredere economische, maatschappelijke, juridische en ethische impact ervan) en om te beslissen wanneer en hoe het systeem in een bepaalde situatie mag worden gebruikt. Zo kan worden beslist dat in een bepaalde situatie geen KI-systeem wordt gebruikt, dat een zekere mate van menselijk ingrijpen mogelijk is tijdens het gebruik van het systeem of dat een door het systeem genomen beslissing van de hand kan worden gewezen.

worden gebruikt om hen te schaden of te discrimineren.

Behalve aan vereisten inzake de bescherming van privacy en persoonsgegevens moet ook worden voldaan aan vereisten die KI-systemen van hoge kwaliteit waarborgen. De kwaliteit van de gebruikte datasets is van essentieel belang voor de prestaties van KI-systemen. Wanneer gegevens worden verzameld, kunnen die sociaal geconstrueerde vooroordelen weerspiegelen, of onjuistheden en fouten bevatten. Pas nadat die zijn aangepakt, mag een dataset worden gebruikt om een KI-systeem te trainen. Bovendien moet de **integriteit** van de gegevens worden gewaarborgd. De gebruikte processen en datasets moeten worden getest en gedocumenteerd bij elke stap (planning, training, testen, toepassing enz.). Dit geldt ook voor KI-systemen die niet intern zijn ontwikkeld maar elders zijn aangeschaft. Ten slotte moet de **toegang** tot de gegevens adequaat worden beheerd en gecontroleerd.

IV. Transparantie

De **traceerbaarheid** van KI-systemen moet worden gewaarborgd. Het is belangrijk dat zowel de door de systemen genomen beslissingen als het hele besluitvormingsproces worden geregistreerd en gedocumenteerd (met inbegrip van een beschrijving van de manier waarop de data zijn verzameld en gelabeld, en een beschrijving van het gebruikte algoritme). In dit verband moet, voor zover mogelijk, worden gezorgd voor **verklaarbaarheid** van het algoritmische besluitvormingsproces, in een aan de betrokken personen aangepaste vorm. Het lopende onderzoek ter ontwikkeling van verklaarbaarheidsmechanismen moet worden voortgezet. Daarnaast moet uitleg beschikbaar zijn over de mate waarin een KI-systeem invloed heeft op en vorm geeft aan het organisatorische besluitvormingsproces, de ontwerpkeuzes voor het systeem en de reden voor de toepassing ervan (zodat niet alleen transparantie wordt geboden op het niveau van de data en het systeem maar ook op het niveau van het bedrijfsmodel).

Ten slotte is het belangrijk om de mogelijkheden en beperkingen van het KI-systeem op adequate wijze **kenbaar te maken** aan de verschillende betrokken partijen, op een manier die past bij de gebruikssituatie. Bovendien moeten KI-systemen als zodanig herkenbaar zijn, zodat de gebruikers zich ervan bewust zijn dat ze te maken hebben met een KI-systeem en weten wie ervoor verantwoordelijk is.

V. Diversiteit, non-discriminatie en billijkheid

Datasets die door KI-systemen worden gebruikt (voor de training of de werking van het systeem), kunnen aangetast zijn door onbedoelde historische vooroordelen, lacunes en slechte governance modellen. Als dergelijke vooroordelen worden overgenomen, kan dat leiden tot (in)directe discriminatie. Schade kan ook het gevolg zijn van de opzettelijke uitbuiting van (consumenten)vooroordelen of van oneerlijke concurrentie. Ook de manier waarop een KI-systeem wordt ontwikkeld, kan gekleurd zijn door vooroordelen (bijvoorbeeld de manier waarop de programmeercode van een algoritme wordt geschreven). Hiermee moet van bij het begin van de ontwikkeling van het systeem rekening worden gehouden.

Deze zorgpunten kunnen ook worden aangepakt door **diverse ontwerpteam**s op te zetten en door te zorgen voor mechanismen ter bevordering van **participatie**, met name van burgers, in de ontwikkeling van KI. Het is raadzaam om de belanghebbenden die gedurende de levenscyclus van het systeem direct of indirect met het systeem te maken

kunnen krijgen, te raadplegen. KI-systemen moeten het hele scala van menselijke vaardigheden, capaciteiten en noden in aanmerking nemen en via een universeel-ontwerpaanpak zorgen voor toegankelijkheid, waarbij wordt gestreefd naar een gelijke toegang voor personen met een handicap.

VI. Maatschappelijk en ecologisch welzijn

Wil KI betrouwbaar zijn, dan moet rekening worden gehouden met de impact ervan op het **milieu en op andere wezens met gevoel**. Idealiter moeten alle mensen, met inbegrip van de toekomstige generaties, over biodiversiteit en een leefbaar milieu kunnen beschikken. Daarom moeten duurzame en **ecologisch verantwoorde** KI-systemen worden aangemoedigd. Dat geldt ook voor KI-oplossingen die gericht zijn op bepaalde problemen van mondiale omvang, bijvoorbeeld in het kader van de VN-doelstellingen inzake duurzame ontwikkeling.

Bovendien moet de impact van KI-systemen niet alleen worden bekeken vanuit een individueel perspectief, maar ook vanuit het perspectief van de **samenleving als geheel**. Het gebruik van KI-systemen moet zorgvuldig worden overwogen, met name in situaties die verband houden met het democratische bestel, waaronder opinievorming, politieke besluitvorming of verkiezingen. Voorts moet ook de **sociale impact** van KI in overweging worden genomen. KI-systemen kunnen sociale vaardigheden versterken, maar kunnen deze ook in gevaar brengen.

VII. Verantwoordingsplicht

Er moeten mechanismen worden ingevoerd om de verantwoordelijkheid en de verantwoordingsplicht voor KI-systemen en de resultaten ervan te waarborgen, zowel vóór als na de uitvoering ervan. De **controleerbaarheid** van KI-systemen is in dit verband van cruciaal belang. De beoordeling van de systemen door interne en externe auditors en de beschikbaarheid van evaluatieverslagen doen de betrouwbaarheid van de technologie sterk toenemen. Externe controleerbaarheid moet met name worden gewaarborgd in toepassingen die van invloed zijn op de grondrechten, waaronder veiligheidskritische toepassingen.

Potentiële negatieve gevolgen van KI-systemen moeten worden geïdentificeerd, beoordeeld, gedocumenteerd en tot een minimum worden beperkt. Effectbeoordelingen zijn daartoe een geschikt middel. Dergelijke beoordelingen moeten in verhouding staan tot de omvang van de risico's waarmee de KI-systemen gepaard gaan. **Trade-offs** tussen de vereisten zijn vaak onvermijdelijk, maar moeten op een rationele en methodologische wijze worden aangepakt en moeten worden verantwoord. Ten slotte moet worden voorzien in toegankelijke mechanismen voor **adequate verhaalmogelijkheden** in geval van ongerechtvaardigde negatieve effecten.

2.3. Volgende stappen: proeffase met een zo breed mogelijke betrokkenheid van belanghebbenden

Het bereiken van een consensus over de essentiële vereisten voor KI-systemen is een eerste belangrijke stap in de richting van richtsnoeren voor ethische KI. Als volgende stap zal de Commissie ervoor zorgen dat deze richtsnoeren kunnen worden getest en in de praktijk kunnen worden toegepast.

Daarom zal de Commissie de aanzet geven tot een gerichte proeffase om gestructureerde feedback van belanghebbenden te verkrijgen. Daarbij zal met name worden gefocust op de beoordelingslijst die de deskundigengroep op hoog niveau voor elk van de essentiële vereisten heeft opgesteld.

De werkzaamheden zullen uit twee onderdelen bestaan: i) een proeffase voor de richtsnoeren, waarbij belanghebbenden worden betrokken die KI ontwikkelen of gebruiken, waaronder overheidsdiensten; en ii) een voortdurende raadpleging en bewustmaking van verschillende groepen belanghebbenden, waaronder de industrie en de dienstensector, in alle lidstaten.

- (i) Vanaf juni 2019 zullen alle belanghebbenden en burgers de beoordelingslijst kunnen testen en feedback kunnen geven over wat beter kan. Daarnaast zal de deskundigengroep op hoog niveau inzake KI samen met belanghebbenden uit de private en de publieke sector een diepgaande evaluatie verrichten om meer gedetailleerde feedback te verkrijgen over hoe de richtsnoeren op zoveel mogelijk gebieden kunnen worden toegepast. Alle feedback over de werkbaarheid en de haalbaarheid van de richtsnoeren wordt uiterlijk eind 2019 geëvalueerd.
- (ii) Parallel hiermee zal de Commissie verdere bewustmakingsactiviteiten organiseren, waarbij vertegenwoordigers van de deskundigengroep op hoog niveau inzake KI de richtsnoeren kunnen voorstellen aan relevante belanghebbenden in de lidstaten, waaronder de industrie en de dienstensector, die daarmee een extra kans krijgen om opmerkingen te maken en bij te dragen aan de KI-richtsnoeren.

De Commissie zal rekening houden met de werkzaamheden van de deskundigengroep inzake ethiek voor geconnecteerd en geautomatiseerd rijden¹⁴. Zij zal ook samenwerken met door de EU gefinancierde onderzoeksprojecten inzake KI en met relevante publiek-private partnerschappen met het oog op de uitvoering van de essentiële vereisten¹⁵. Zo zal de Commissie, in overleg met de lidstaten, de ontwikkeling ondersteunen van een gemeenschappelijke databank met medische beelden die in een eerste fase op de meest voorkomende vormen van kanker zijn toegespitst, zodat algoritmen kunnen worden getraind om met een zeer hoge nauwkeurigheid een diagnose te stellen. De samenwerking tussen de Commissie en de lidstaten maakt het ook mogelijk om meer grensoverschrijdende corridors op te zetten voor het testen van geconnecteerde en geautomatiseerde voertuigen. In het kader van die projecten zouden de richtsnoeren moeten worden toegepast en getest, zodat de resultaten kunnen worden meegenomen in het evaluatieproces.

De proeffase en de raadpleging van belanghebbenden zullen gebaat zijn bij de bijdrage van de Europese KI-alliantie en het AI-on-demandplatform AI4EU. Het AI4EU-project¹⁶, dat in januari 2019 van start is gegaan, brengt algoritmen, instrumenten, datasets en diensten bijeen om organisaties, met name kleine en middelgrote ondernemingen, te helpen om KI-oplossingen te implementeren. De Europese KI-alliantie en AI4EU zullen zich samen blijven inspannen voor de mobilisatie van het KI-ecosysteem in heel Europa, ook met het oog op het beproeven van de ethische richtsnoeren voor KI en de bevordering van mensgerichte KI.

Begin 2020 zal de deskundigengroep op hoog niveau inzake AI de richtsnoeren evalueren en bijwerken. Zij zal daarbij voortbouwen op de evaluatie van de feedback die

¹⁴ Zie de mededeling van de Commissie met betrekking tot geconnecteerde en geautomatiseerde mobiliteit (COM(2018) 283).

¹⁵ In het kader van het Europees Defensiefonds zal de Commissie ook specifieke ethische richtsnoeren opstellen voor de evaluatie van projectvoorstellen op het gebied van KI voor defensie.

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

tijdens de proeffase wordt ontvangen. **De Commissie** zal op basis van de evaluatie en de opgedane ervaring **de resultaten evalueren en eventuele volgende stappen voorstellen**.

Ethische KI levert alleen maar voordelen op. Door ervoor te zorgen dat de fundamentele waarden en rechten worden geëerbiedigd, op zichzelf al van essentieel belang, wordt de acceptatie door het publiek bevorderd en zien de Europese KI-bedrijven hun concurrentievoordeel toenemen dankzij de totstandbrenging van een merk van mensgerichte, betrouwbare KI, dat bekendstaat om zijn ethische en veilige producten. Daarmee wordt meer in het algemeen voortgebouwd op de sterke reputatie van Europese bedrijven wat de productie van veilige en betrouwbare producten van hoge kwaliteit betreft. De proeffase zal ertoe bijdragen dat KI-producten aan deze belofte voldoen.

2.4. Op weg naar internationale ethische richtsnoeren voor KI

Nadat het Japanse voorzitterschap van de G7 het onderwerp in 2016 hoog op de agenda had gezet, zijn de internationale discussies over KI-ethiek sterk toegenomen. De ontwikkeling van KI wordt op internationaal niveau gekenmerkt door een nauwe verwevenheid op het vlak van gegevenscirculatie, de ontwikkeling van algoritmen en investeringen in onderzoek. Daarom **zal de Commissie haar inspanningen voortzetten om de aanpak van de Unie wereldwijd ingang te doen vinden en een consensus tot stand te brengen over een mensgerichte KI**¹⁷.

Het werk van de deskundigengroep op hoog niveau inzake KI, en meer in het bijzonder de lijst van vereisten en de inbreng van de belanghebbenden, biedt de Commissie aanvullende waardevolle input voor haar deelname aan internationale discussies. De Europese Unie kan een leidende rol spelen bij de ontwikkeling van internationale KI-richtsnoeren en, indien mogelijk, een bijbehorend beoordelingsmechanisme.

Daarom zal de Commissie het volgende doen:

Zorgen voor meer samenwerking met gelijkgestemde partners:

- nagaan in hoeverre convergentie mogelijk is met de ontwerpen van ethische richtsnoeren van derde landen (bijvoorbeeld Japan, Canada, Singapore) en, vertrekkend vanuit die groep van gelijkgestemde landen, zich voorbereiden op een bredere discussie, ondersteund door acties ter uitvoering van het partnerschapsinstrument voor samenwerking met derde landen¹⁸; en

¹⁷ De hoge vertegenwoordiger van de Unie voor buitenlandse zaken en veiligheidsbeleid zal, met de steun van de Commissie, voortbouwen op overleg in de Verenigde Naties, het Global Tech Panel en andere multilaterale fora, en zal met name instaan voor de coördinatie van voorstellen voor de aanpak van de complexe veiligheidsproblemen die zich kunnen voordoen.

¹⁸ Verordening (EU) nr. 234/2014 van het Europees Parlement en de Raad van 11 maart 2014 tot vaststelling van een partnerschapsinstrument voor samenwerking met derde landen (PB L 77 van 15.3.2014, blz. 77). Zo zal het geplande project inzake "een internationale alliantie voor een mensgerichte aanpak van kunstmatige intelligentie" gezamenlijke initiatieven met gelijkgestemde partners faciliteren, met als doel de bevordering van ethische richtsnoeren en de vaststelling van gemeenschappelijke beginselen en operationele conclusies. Het zal de EU en gelijkgestemde landen in staat stellen om operationele conclusies te bespreken die voortvloeien uit de door de deskundigengroep voorgestelde ethische richtsnoeren inzake KI, om zo tot een gemeenschappelijke aanpak te komen. Het zal er ook voor zorgen dat de opname van KI-technologie wereldwijd kan worden gemonitord. Ten slotte voorziet het project in de organisatie van activiteiten op het gebied van publieksdiplomatie in de marge van internationale evenementen, bijvoorbeeld van de G7, de G20 en de Organisatie voor Economische Samenwerking en Ontwikkeling.

- onderzoeken hoe bedrijven uit niet-EU-landen en internationale organisaties kunnen bijdragen aan de "proeffase" van de richtsnoeren door tests en valideringsactiviteiten te verrichten.

Een actieve rol blijven spelen in internationale discussies en initiatieven:

- actief zijn in multilaterale fora zoals de G7 en de G20;
- een dialoog aangaan met niet-EU-landen en bilaterale en multilaterale bijeenkomsten organiseren om een consensus te bereiken over mensgerichte KI;
- deze visie promoten door bij te dragen tot desbetreffende normalisatieactiviteiten in internationale normalisatieorganisaties; en
- in samenwerking met relevante internationale organisaties zorgen voor een betere verzameling en verspreiding van informatie over het overheidsbeleid.

3. CONCLUSIES

De EU is gegrondvest op een reeks fundamentele waarden waarop een sterk en evenwichtig regelgevingskader is gebouwd. Voortbouwend op dit bestaande regelgevingskader moeten ethische richtsnoeren worden opgesteld voor de ontwikkeling en het gebruik van KI, die als nieuwe technologie specifieke problemen met zich meebrengt. Slechts indien KI wordt ontwikkeld en gebruikt in overeenstemming met breed gedeelde ethische waarden, kan zij als betrouwbaar worden beschouwd.

In dat licht bezien verwelkomt de Commissie de input van de deskundigengroep op hoog niveau inzake KI. Op basis van de essentiële vereisten waaraan KI moet voldoen om als betrouwbaar te worden beschouwd, zal de Commissie nu de aanzet geven tot een gerichte proeffase, zodat de daaruit voortvloeiende ethische richtsnoeren voor de ontwikkeling en het gebruik van KI in de praktijk kunnen worden toegepast. De Commissie zal zich ook inspannen om tot een brede maatschappelijke consensus over mensgerichte KI te komen, waaronder met alle belanghebbenden en onze internationale partners.

De ethische dimensie van KI is geen overbodige luxe of optie, maar moet volledig worden geïntegreerd in de ontwikkeling van KI. Door te streven naar een mensgerichte, op vertrouwen gebaseerde KI kunnen we ervoor zorgen dat onze maatschappelijke kernwaarden worden geëerbiedigd en dat Europa en zijn industrie zich kunnen positioneren als voortrekkers van geavanceerde KI die overal ter wereld het vertrouwen geniet.

Om de ethische ontwikkeling van KI in Europa in een bredere context te garanderen, streeft de Commissie naar een alomvattende aanpak via met name de volgende activiteiten, die uiterlijk in het derde kwartaal van 2019 moeten zijn uitgevoerd:

- In het kader van Horizon 2020 zal een reeks **netwerken van excellentiecentra op het gebied van KI-onderzoek** worden opgericht. De Commissie zal tot vier netwerken selecteren die zich richten op belangrijke wetenschappelijke of technologische uitdagingen, zoals verklaarbaarheid en geavanceerde mens-machine-interactie, die van belang zijn met het oog op betrouwbare KI.
- Er wordt van start gegaan met het opzetten van **netwerken van digitale-innovatiehubs**¹⁹ die focussen op KI in de industrie en op big data.

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

- Samen met de lidstaten en de belanghebbenden zal de Commissie voorbereidende besprekingen opstarten met het oog op de ontwikkeling en de implementatie van **een model voor het delen van gegevens en een optimaal gebruik van gemeenschappelijke dataruimten**, met bijzondere aandacht voor vervoer, gezondheidszorg en industriële productie.²⁰

Daarnaast werkt de Commissie aan een verslag over de uitdagingen die KI meebrengt voor de kaders voor veiligheid en aansprakelijkheid, en aan een leidraad voor de uitvoering van de productaansprakelijkheidsrichtlijn²¹. Tegelijkertijd is de Gemeenschappelijke Onderneming Europese high-performance computing (EuroHPC)²² bezig met de ontwikkeling van de volgende generatie supercomputers. Computercapaciteit is immers essentieel voor de verwerking van data en de training van KI, en Europa moet de volledige digitale waardeketen beheersen. Het lopende partnerschap met de lidstaten en de industrie op het gebied van micro-elektronische componenten en systemen (Ecsel)²³ en ook het European Processor Initiative²⁴ zullen bijdragen tot de ontwikkeling van energiezuinige processortechnologie voor krachtige en geavanceerde computers die betrouwbaar en veilig zijn.

Net als de werkzaamheden met betrekking tot ethische richtsnoeren voor KI, bouwen al deze initiatieven voort op een **nauwe samenwerking tussen alle belanghebbenden**, de lidstaten, de industrie, maatschappelijke actoren en burgers. Alles welbeschouwd laat de Europese benadering van kunstmatige intelligentie zien hoe economisch concurrentievermogen en maatschappelijk vertrouwen moeten uitgaan van dezelfde fundamentele waarden en elkaar wederzijds versterken.

²⁰ De nodige middelen zullen worden vrijgemaakt uit Horizon 2020 (in het kader waarvan voor de periode 2018-2020 bijna 1,5 miljard EUR wordt uitgetrokken voor KI) en Horizon Europa (de opvolger van Horizon 2020), het digitale onderdeel van de Connecting Europe Facility en vooral het toekomstige programma Digitaal Europa. Er zullen ook middelen voor projecten beschikbaar zijn van de particuliere sector en programma's in de lidstaten.

²¹ Zie de mededeling van de Commissie over "Kunstmatige intelligentie voor Europa, (COM(2018) 237).

²² <https://eurohpc-ju.europa.eu>

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu