



Robot-ethiek: de waardigheid van iedere mens moet centraal blijven staan

Tot de deelnemers aan de 25^{ste} Algemene vergadering van de Pauselijke Academie voor het Leven: Robotethiek - Mensen, machines en gezondheid

Paus Franciscus

25 februari 2019

Beste broeders en zusters,

Ik groet u van harte bij gelegenheid van uw algemene vergadering en ik dank Mgr. Paglia voor zijn vriendelijke woorden. Deze ontmoeting vindt plaats ter gelegenheid van het eerste jubileum van de Academie voor het leven: op 25 jaar na haar ontstaan. Bij deze belangrijke gebeurtenis heb ik de vorige maand een brief aan de president geschreven met de titel: *Humana communitas*. Het was vooral de wens om alle presidenten die elkaar hebben opgevolgd aan het hoofd van de Academie en alle leden te bedanken voor hun competente dienstwerk en hun edelmoedige inzet voor de bescherming en de bevordering van het menselijke leven gedurende deze werkzame 25 jaar.

We kennen de moeilijkheden waarmee onze wereld de strijd aanbindt. Het weefsel van de betrekkingen in gezin en maatschappij lijkt steeds meer te verslijten en er verspreidt zich een neiging om in zichzelf en zijn persoonlijke belangen opgesloten te raken, met ernstige gevolgen voor de "grote en beslissende kwestie van de eenheid in de mensenfamilie en haar toekomst" (brief *Humana communitas*, nr. 2). Er tekent zich een dramatische paradox af. Juist nu de mensheid de wetenschappelijke mogelijkheden bezit om te komen tot een eerlijk verdeeld welzijn volgens het gebod van God, zien we daarentegen een versterking van de conflicten en een toename van de ongelijkheid. De mythe van de vooruitgang volgens de filosofie van de Verlichting neemt in betekenis af en de opeenstapeling van de mogelijkheden die wetenschap en techniek ons hebben verschaft leidt niet altijd tot de resultaten waarop gehoopt werd. Inderdaad, aan de ene kant hebben de technische ontwikkelingen het ons mogelijk gemaakt problemen op te lossen die enkele jaren geleden nog onoverkomelijk waren en we zijn de onderzoekers die dergelijke resultaten mogelijk gemaakt hebben dankbaar; aan de andere kant zijn er moeilijkheden en dreigingen die soms sluipender zijn dan die van voorheen aan het licht gekomen. Het "kunnen doen" brengt het gevaar mee dat niet meer gezien wordt *wie* het doet en *voor wie* men het doet. Het technocratische systeem dat gebaseerd is op het criterium van de doeltreffendheid beantwoordt niet aan de diepste vragen die de mens zichzelf stelt. En al is het dan aan de ene kant onmogelijk om het zonder de hulpbronnen te stellen, aan de andere kant legt dit systeem zijn logica op aan hen die ze gebruiken. Toch is de techniek een kenmerk van het menselijke wezen. Men moet haar niet zien als een kracht die de mens vreemd en vijandig is, maar als een voortbrengsel van zijn vindingrijkheid, waardoor hij kan voorzien in de eisen die het leven stelt, voor zichzelf en voor de anderen. Het is dus een typisch menselijke manier om de wereld te bewonen. Maar de huidige ontwikkeling van de technische mogelijkheden brengt een gevaarlijke betovering met zich mee: in plaats van voor het menselijke leven de hulpmiddelen te leveren die de zorg ervoor verbeteren, lopen we het gevaar dat we het leven overlaten aan de logica van de voorzieningen die beslissen over de waarde ervan. Die omkering mondt uit in kwalijke resultaten: de machine blijft er niet bij om geheel alleen zichzelf te besturen, maar gaat uiteindelijk de mens besturen. Zo wordt de menselijke rede

teruggebracht tot een redelijkheid die is losgemaakt van haar effecten en die niet meer als menswaardig gezien mag worden.

We zien helaas de ernstige schade die aan de planeet, ons gemeenschappelijke huis, wordt toegebracht door het zonder onderscheid toepassen van technische middelen. Daarom is de wereld-bio-ethiek een belangrijke zaak om zich voor in te zetten. Zij laat zien dat men zich bewust is van de diepgaande invloed die de milieufactoren en maatschappelijke factoren op de gezondheid en het leven hebben. Dat is een aanpak die geheel in overeenstemming is met de integrale ecologie, zoals die beschreven en aanbevolen werd in de encycliek *Laudato si*. Het is bovendien op zijn plaats dat wij in de huidige wereld, die gekenmerkt wordt door een nauw contact tussen verschillende culturen, als gelovigen onze specifieke bijdrage leveren in het zoeken naar operationele criteria die met allen gedeeld kunnen worden en die gemeenschappelijke uitgangspunten zijn voor het maken van keuzes door het die de grote verantwoordelijkheid hebben om op nationaal en internationaal niveau beslissingen te nemen. Dat betekent ook dat men zich engageert in de dialoog over de mensenrechten, waarbij ook de plichten die ermee gepaard gaan duidelijk belicht worden. Zij vormen inderdaad het onderzoeksgebied voor een universele ethiek en we zien dat belangrijke vragen daarover door de traditie behandeld zijn door te putten uit het erfgoed van de natuurwet.

De brief *Humana communitas* verwijst expliciet naar het thema van “opkomende en met elkaar samenwerkende technologieën”. De mogelijkheid om op de levende materie in te grijpen op een niveau met een steeds kleinere schaal, om steeds grotere massa’s informatie te verwerken, om de hersenprocessen die te maken hebben met de activiteit van het kennen en het denken te volgen – en te beïnvloeden – heeft onmetelijke consequenties: dit raakt de eigenlijke drempel van de biologische specificiteit van de mens en wat hem in geestelijke zin verschillend maakt. Ik bevestig hierbij dat “wat het menselijke leven verschillend maakt een absoluut goed is” (nr. 4).

Het is belangrijk dit nog eens te herhalen: “kunstmatige intelligentie, robotica en andere technologische vernieuwingen moeten gebruikt worden om bij te dragen aan de dienst aan de mensheid en aan de bescherming van ons gemeenschappelijke huis en niet aan precies het tegenovergestelde daarvan, zoals helaas sommige schattingen voorzien.” (*Boodschap aan het World Economic Forum te Davos 12 januari 2019*). De waardigheid die elk menselijk wezen eigen is moet met vasthoudendheid geplaatst worden in het centrum van ons denken en handelen.

In dit kader is het op zijn plaats op te merken dat de benaming “kunstmatige intelligentie” het gevaar loopt bedrieglijk te zijn ook al kan ze zeker doeltreffend zijn. De termen verhullen het feit dat die functionele automatisen in hun kwaliteiten ver verwijderd blijven van die van het kennen en handelen die een menselijk privilege zijn – ook al zijn ze nuttig in het vervullen van slaafse taken (dat is de oorspronkelijke betekenis van de term “robot”). En juist daarom kunnen ze een maatschappelijk gevaar vormen. Er is trouwens al een risico dat de mens vertechnologiseerd wordt, in plaats van dat de techniek vermenselijkt wordt: men schrijft al te snel aan zogenaamde “intelligente machines” vermogens toe die echt menselijk zijn.

We dienen beter te begrijpen wat in deze context de betekenis is van intelligentie, bewustzijn, gevoelsleven, gevoelsmatige gerichtheid en de autonomie van het moreel handelen. De kunstmatige voorzieningen die menselijke vermogens nabootsen bezitten in werkelijkheid geen enkele menselijke hoedanigheid. Daarmee moet rekening gehouden worden om de regeling van hun gebruik en het onderzoek zelf te oriënteren op een constructieve en billijke interactie tussen de mensen en de meeste recente versies van die machines. Deze verspreiden zich immers in onze wereld en wijzigen het scenario van ons bestaan radicaal. Als we in staat zijn deze overwegingen te doen gelden in wat er gebeurt, dan zullen de buitengewone mogelijkheden van de nieuwe ontdekkingen hun zegeningen kunnen doen stralen over iedereen en over de hele mensheid.

De discussie tussen de specialisten zelf laat de ernstige problemen aangaande de bestuurbaarheid van de algoritmen die met enorme hoeveelheden data omgaan al zien. Evenzo stellen de technieken waarmee het genetisch erfgoed en de hersenfuncties gemanipuleerd kunnen worden ons voor ernstige ethische vraagstukken. De poging om het geheel van het gedachteleven, het gevoelsleven en de menselijke psyché te verklaren op grond van de functionele som van de samenstellende fysieke en organische onderdelen houdt in

ieder geval geen rekening met wat er aan het licht komt aan verschijnselen van de ervaring en het bewustzijn. Het verschijnsel mens is meer dan het resultaat van een optelling en samenstelling van elk van zijn elementen. Ook op dit terrein krijgt het axioma dat zegt dat het geheel meer is dan de som van de delen een nieuwe diepgang en een nieuwe betekenis.

Aan de andere kant leren we nu juist in deze lijn van de complexiteit in de synergie tussen *psyché* en *technè* dat, hetgeen we te weten komen over de hersenactiviteit nieuwe aanwijzingen levert voor de wijze waarop we het bewustzijn (van onszelf en van de wereld) en het menselijk lichaam zelf moeten zien: het is niet mogelijk om geen rekening te houden met het in elkaar grijpen van allerlei betrekkingen als we willen geraken tot een dieper begrip van de integrale menselijke dimensie.

Zeker, we kunnen geen metafysische gevolgtrekkingen maken uit de gegevens van de empirische wetenschappen. Maar we kunnen er wel aanwijzingen aan ontleen die richting geven aan ons antropologische denken, zelf op theologisch gebied, zoals dat trouwens in de geschiedenis altijd is gebeurd. Het zou immers tegen onze toch authentieke traditie ingaan als we nu zouden vasthouden aan een anachronistisch begrippenapparaat. We zouden dan niet in staat zijn op passende manier om te gaan met de veranderingen in de begrippen natuur en kunst, omstandigheden en vrijheid, middel en doel die is opgeroepen door de nieuwe cultuur van handelen die typisch is voor het tijdperk van de technologie. We worden opgeroepen de weg in te slaan die op beslissende wijze is gekozen door het Tweede Vaticaans Concilie. Dat roept op tot een vernieuwing van de theologische wetenschappen en een kritisch nadenken over de betrekking tussen Christelijk geloof en moreel handelen (vgl. *Optatam totius*, 16).

In onze deelname in de ethische samenwerking ten gunst van het leven zal onze inzet – zowel op intellectueel als op specialistisch gebied – een erezaak zijn. In een situatie waarin technologische voorzieningen die steeds vernuftiger zijn rechtstreeks ingrijpen op menselijke eigenschappen van lichaam en geest wordt het nu een dringende zaak om dit met alle mannen en vrouwen die zich inzetten voor wetenschappelijk onderzoek en de praktijk van de zorg. Dat is zeker een zware taak, gezien het snelle tempo waarin de vernieuwingen plaatsvinden. We worden hierin aangemoedigd en gesteund door het voorbeeld van leermeesters uit de gelovige intelligentsia, die met wijsheid en moed de processen van hun eigen tijd hebben benaderd met het oog op een begrijpen van het geloofsgoed op een verstandelijk niveau dat de mens waardig is.

Ik wens u toe dat u uw studie en onderzoek kunt voortzetten zodat het werk voor de bevordering en de bescherming van het leven steeds doeltreffender en vruchtbaarder zal worden. Moge de Maagd Maria u bijstaan en moge mijn zegen u begeleiden. En vergeet alstublieft niet voor mij te bidden. Dank u.

Vertaling: dr. J.A. Raymakers

To the participants of the 25th General Assembly of the Pontifical Academy for Life “Robo Ethics - Humans, Machines and Health”

Pope Francis

25 February 2019

Dear brothers and sisters,

I cordially greet you on the occasion of your General Assembly, and I thank Archbishop Paglia for his kind words. This meeting takes place in the first Jubilee of the Academy for Life: twenty-five years after its birth. On this important anniversary, last month I sent the president a letter entitled *Humana communitas*. I was moved to write this message first of all by the wish to thank all the presidents who have guided the Academy, and all the Members for their competent service and generous commitment to protecting and promoting human life during these twenty-five years of activity.

We know the difficulties with which our world struggles. The fabric of family and social relations seems increasingly to wear away, and there is a tendency to become wrapped up in oneself and one's own individual interests, with serious consequences for the “the decisive global issue of the unity of the human family and its future” (Letter *Humana communitas*, 2). A dramatic paradox is thus outlined: just when humanity possesses the scientific and technical capacities to achieve a justly distributed well-being, in accordance with how it was

delivered by God, we observe instead an exacerbation of conflicts and an increase in inequality. The enlightenment myth of progress is declining and the accumulation of the potentialities that science and technology have provided us does not always attain the desired results. Indeed, on the one hand, technological development has allowed us to solve problems that were insurmountable until a few years ago, and we are grateful to the researchers who have achieved these results; yet on the other hand, difficulties and threats, sometimes more insidious than the previous ones, have emerged. The possibility of doing something risks obscuring both the person who does, and the person doing it. The technocratic system based on the criterion of efficiency does not respond to the most profound questions that man poses; and if on the one hand it is not possible to do without its resources, on the other it imposes its logic on those who use them. Yet technology is characteristic of the human being. It should not be understood as a force that is alien to and hostile to it, but as a product of its ingenuity through which it provides for the needs of living for oneself and for others. It is therefore a specifically human mode of inhabiting the world. However, today's evolution of technical capacity casts a dangerous spell: instead of delivering the tools that improve their care to human life, there is the risk of giving life to the logic of the devices that decide its value. This reversal is destined to produce nefarious outcomes: the machine is not limited to driving alone, but ends up guiding man. Human reason is thus reduced to rationality alienated from effects, which cannot be considered worthy of mankind.

We see, unfortunately, the serious damage caused to the planet, our common home, from the indiscriminate use of technical means. This is why global bioethics is an important front on which to engage. It expresses awareness of the profound impact of environmental and social factors on health and life. This approach is very in tune with the integral ecology described and promoted in the Encyclical *Laudato si'*. Moreover, in today's world, in which there is close interaction between different cultures, we need to bring our specific contribution as believers to the search for universally shared operational criteria, so that they may be common points of reference for the choices of those who have the serious responsibility for taking decisions on national and international levels. This also means engaging in dialogue regarding human rights, clearly highlighting their corresponding duties. Indeed these constitute the ground for the common search for universal ethics, on which we find many questions that tradition has dealt with by drawing on the patrimony of natural law.

The Lettera *Humana communitas* explicitly recalls the theme of "emerging and converging technologies". The possibility of intervening on living material to orders of ever smaller size, to process ever greater volumes of information, to monitor - and manipulate - the cerebral processes of cognitive and deliberative activity, has enormous implications: it touches the very threshold of the biological specificity and spiritual difference of the human being. In this sense, I affirmed that "The distinctiveness of human life is an absolute good" (4).

It is important to reiterate: "Artificial intelligence, robotics and other technological innovations must be so employed that they contribute to the service of humanity and to the protection of our common home, rather than to the contrary, as some assessments unfortunately foresee" ([Message to the World Economic Forum in Davos, 12 January 2018](#)). The inherent dignity of every human being must be firmly placed at the centre of our reflection and action. In this regard, it should be noted that the designation of "artificial intelligence", although certainly effective, may risk being misleading. The terms conceal the fact that - in spite of the useful fulfilment of servile tasks (this is the original meaning of the term "robot"), functional automatisms remain qualitatively distant from the human prerogatives of knowledge and action. And therefore they can become socially dangerous. Moreover, the risk of man being "technologized", rather than technology humanized, is already real: so-called "intelligent machines" are hastily attributed capacities that are properly human.

We need to understand better what intelligence, conscience, emotionality, affective intentionality and autonomy of moral action mean in this context. Indeed, artificial devices that simulate human capabilities are devoid of human quality. This must be taken into account to guide the regulation of their use, and research itself, towards a constructive and equitable interaction between human beings and the latest versions of machines. Indeed these spread throughout our world and radically transform the scenario of our existence. If we can also make these references bear weight also in action, the extraordinary potential of the new discoveries may radiate their benefits on every person and on the whole of humanity.

The ongoing debate among specialists themselves already shows the serious problems of governability of algorithms that process huge amounts of data. Likewise, the technologies for the manipulation of genetic makeup and brain functions also pose serious ethical questions. In any case, the attempt to explain the whole of human thought, sensitivity, and psychism on the basis of the functional sum of its physical and organic parts, does not account for the emergence of the phenomena of experience and consciousness. The human phenomenon exceeds the result of the calculable assemblage of the individual elements. Also in this context, the axiom according to which the whole is superior to the parts takes on new depth and meaningfulness (see Apostolic Exhortation *Evangelii gaudium*, 234-237).

Precisely in this area of the complexity of the synergy of *psyche* and *techne*, on the other hand, what we are learning about cerebral activity provides new clues about the way of understanding the conscience (of self and of the world) and the human body itself: it is not possible to disregard the interweaving of multiple relationships for a deeper understanding of the integral human dimension.

Of course, we cannot make metaphysical deductions from the data provided by empirical sciences. We can, however, draw from them indications that instruct anthropological reflection, in theology too, as has always happened in its history. It would indeed be decidedly contrary to our more genuine tradition to become set on an anachronistic conceptual apparatus, incapable of adequately interacting with the transformations of the concept of nature and of artifice, conditioning and freedom, means and ends, induced by the new culture of acting, typical of the technological era. We are called to place ourselves on the path undertaken decisively by *Vatican Council II*, which calls for the renewal of theological disciplines and a critical reflection on the relationship between Christian faith and moral action (cf. *Optatam totius*, 16).

Our commitment - also intellectual and specialist - will be a point of honour for our participation in the ethical alliance in favour of human life. A project which, in a context in which increasingly sophisticated technological devices directly involve the human qualities of the body and the psyche, it becomes urgent to share with all men and women engaged in scientific research and care work. It is a difficult task, certainly, given the fast pace of innovation. The example of the teachers of the Christian intelligence, who entered with wisdom and audacity in the processes of their contemporary world, with a view to an understanding of the patrimony of the faith at the level of reason worthy of man, must encourage and sustain us.

I hope you will continue your study and research so that the work of the promotion and defence of life may be increasingly effective and fruitful. May the Virgin Mother assist you and my blessing accompany you. And please, do not forget to pray for me. Thank you.