

Human rights as benchmarks of artificial intelligence

Ernst Hirsch Ballin

WRR



The 'Working Papers' series includes studies that have been carried out as part of the WRR's work. The views and opinions expressed in these papers are those of the authors. A listing of all Working Papers can be found at www.wrr.nl

The Netherlands Scientific Council for Government Policy

Buitenhof 34

po Box 20004

2500 EA The Hague, The Netherlands

Phone +31 (0)70 356 46 00

Fax +31 (0)70 3564685

info@wrr.nl

www.wrr.nl

Design: Today, Utrecht

Working Paper number 46

e-ISBN: 978-90-83201-20-7

© The Netherlands Scientific Council for Government Policy,
The Hague 2021

The content of this publication may be (partially) used and reproduced for non-commercial purposes. The content may not change. Quotes must always be indicated.

Human rights as benchmarks of artificial intelligence

Ernst Hirsch Ballin

WRR

Content

	Foreword	7
	Introduction	8
1.	Human rights in the Anthropocene	9
2.	Artificial intelligence in relationships between persons and bodies	15
3.	Threatening and repressive artificial intelligence	21
4.	Supportive and liberating artificial intelligence	25
5.	Benchmarks	33
6.	Artificial intelligence in the Anthropocene	39
7.	Conclusion	40
	Bibliography	42

Foreword

This WRR Working Paper entitled ‘Human rights as benchmarks of artificial intelligence’ was written by Professor Ernst Hirsch Ballin, who is an advisory member of the Council of the Netherlands Scientific Council for Government Policy and distinguished university professor at Tilburg University.

This Working Paper considers the relationship between artificial intelligence and human rights, but in a way that differs from the approach often adopted. The author believes the impact of artificial intelligence will be understated if discussion is confined to rules intended to ‘frame’ its use: it is the working processes of the democratic constitutional state itself that need to be overhauled. Artificial intelligence must be seen as a facet of our social coexistence, with the constant moral and political challenge being to improve its dignity and sustainability. This Working Paper was completed on 12 July 2021.

The ‘Working Papers’ series comprises studies produced as part of the WRR’s work. Responsibility for the content rests with the authors.

Prof. J.E.J. (Corien) Prins
Chair of the WRR

Prof. F.W.A. (Frans) Brom
Secretary of the WRR

Introduction

Many people are concerned that the use of artificial intelligence (AI) will have a dehumanizing impact (in the sense of non-respect for human dignity). They appeal to human rights as a safeguard against such impact. This applies to both the establishment and use of AI systems. In the context of the OECD and – potentially with stronger binding force – the European Union, such appeals lead to proposals for regulation. Without in any way detracting from the importance of such regulation, we view the subject here from a different perspective. Aside from regulation, key points include training those required to work with artificial intelligence so that they can live with it, and at the same time gearing the administrative organization to the required working practices. In this essay we assess how human administrative intelligence needs to develop in order to handle artificial intelligence according to the standards of a democratic community governed by the rule of law.

1. Human rights in the Anthropocene

This essay considers the relationship between artificial intelligence and human rights, but in a way that differs from the approach that – albeit for good reason – is often adopted. The usual focus is on the ‘limits’ that human rights set on the ‘use’ of artificial intelligence. Here we consider first and foremost how artificial intelligence relates to the humanity (in the sense of respecting human dignity) of societal processes and organization. This may indeed lead to limits on the use of artificial intelligence, although there is no prospect of ever eliminating it from society (or parts of society). As Karl Jaspers wrote in 1946, technology and science spread around the world in the preceding four centuries and – with all the attendant consequences – became the ‘Weltschicksal’, or destiny of the world.¹ The emergence of artificial intelligence marks a new phase in this destiny.

This essay does not, therefore, ‘run checks’ of human rights as specifically formulated in treaties and constitutions to gauge the acceptability of practices based on artificial intelligence. Rather, our task is to specify how human administrative intelligence can relate to those practices.

To do so we must first explain how artificial intelligence can contribute to a political and social order that seeks to fulfil human rights while potentially also posing a threat to them. The Dutch government’s use of artificial intelligence, principally for administrative tasks, is rapidly outpacing the evolution of constitutional and administrative law, jeopardizing the balance between the duties of the state.² Human rights, as enshrined in treaties and constitutions, must constitute binding frameworks for all functions of the state. They may be very important for people or groups who rely on them in the event of violations. (The fundamental rights in the Dutch Constitution barely fulfil this function, since any assessment of laws against the Constitution in Article 120 has so far been ruled out.) At the same time, as the background and basis for these legal functions, human rights give expression to a distinct culture of coexistence in which human dignity is sacrosanct. This is characterized by relationships of mutual respect in which people do not subjugate others. That is a basis for what the European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR) refers to as a ‘democratic society’.³ The concern

1 Cited from a lecture to the *Recontres Internationales de Genève* in Assmann 2018.

2 Passchier 2020.

3 This formulation occurs many times in the restrictive clauses of fundamental rights. It is difficult to provide a comprehensive definition, but characteristics can be cited such as tolerance, social pluralism, relative openness and primacy of law, according to Logemann 2004.

in a social order underpinned by such a culture must be whether – and, if so, under what conditions – artificial intelligence enables practices that help to fulfil the fundamental rights of every human being, and where the opposite is the case. We first concern ourselves with these practices. As in the case of climate change, these are developments whose consequences may only manifest themselves over the longer term. Case law increasingly recognizes that these long-term effects must also be considered in the protection of fundamental rights.⁴

The risk of dehumanizing effects must not be taken lightly: it goes beyond incidental corrections by a supervisory body such as the Dutch Data Protection Authority or the courts. Even in a democratic society there is constant pressure on respect for each individual's personal dignity. This is a consequence not only of a deliberate pursuit of power at the expense of others, but also of the persistent, apparent attractiveness of models of political and social order aimed at controlling human lives, which have begun to use artificial intelligence for their propaganda and exercise of power.

The attractiveness of practices and movements that exercise power at the expense of others (for example certain population groups, or people who do not 'conform'), is aggravated by – for some more than others – frightening experiences. These include a lack of predictability, for example in the use of government schemes, the behaviour of outsiders, the operation of financial markets, the climate and the spread of diseases.

Hartmut Rosa has described such post-modern experiences as disillusion with the pursuit of 'Verfügbarkeit' (availability), to which even more vigorous efforts to exert control are not an appropriate response.⁵ Amid these tensions the rise of artificial intelligence plays an important role, because it holds out the promise – or allure – of 'taking back control' with the aid of technology. A painful example in the Netherlands has been the 'combating of fraud' in childcare benefits. On the other hand, the potential of artificial intelligence opens up entirely new ways of analysing inaccessible patterns to detect and tackle injustice and deprivation, for example to combat hunger.

Regardless of the extent to which a technology-fuelled pursuit of control constitutes a risky illusion or a real danger, or the extent to which this technology opens up new opportunities for humanity to come to terms with itself and its environment, the key point in both cases is the relationship of

4 A comparison on this point between the Urgenda case law and the need to protect citizens and businesses in the digital sphere can be found in Prins 2018.

5 Rosa 2018.

artificial intelligence to the culture of coexistence that is expressed through human rights. This essay will show that developments can go in either direction.

Artificial intelligence appeared on the stage of history at the point where the subject-object relationship of persons and things, including nature, reached a tipping point. The era that has become known in geological terms as ‘Anthropocene’⁶ presents a shocking paradox. While human activity has made such an impression on the planet that climate, biodiversity and atmospheric processes have changed irreversibly, humanity itself has lost the ability to shape its relationships on the primordial subject-object basis whereby man is seen as the controller – or at least the steward – of nature and other ‘things’⁷ that he can control for his own ends.

In another context⁸ I wrote on this:

The need to reconsider the relationship between human rights and entitlements to protection of living nature emerges from the recognition that, with the transition to the Anthropocene, humankind has arrived at the limits of its increasing control over nature.⁹ [...] Even though the human rights revolution of the 20th century produced an understanding of every human being as a “person”, the very distinction between “persons and things” [...] has become unsettled in the present “new axial age”.¹⁰ [...] This is specifically the case with the proprietary significance of the person/thing divide. Roberto Esposito explained in a genealogical investigation of this divide that it is designed to reduce every “thing”, including even enslaved human beings, to an object that can be owned by subjects performing the legal role of a “person”.¹¹ Under the present conditions of the Anthropocene, the “world” – in the sense of the ensemble of objects of possible domination – is resisting and escaping control by humans.

Human rights are an attempt to calibrate relationships between human beings and their relationships with all kinds of goods and resources against the principle of human dignity referred to earlier: human beings must not be reduced to objects of the exercise of power and exploitation, but on the contrary have an equal entitlement to respect and protection. Human rights have, so to speak, placed every human being on an equal pedestal: no one must be an object of his fellow man, as occurs in slavery, people trafficking and genocide. The downside, however, is that animals and, more generally, manifestations of the Earth’s natural environment, such as rainforests and rivers, are downgraded

6 Ellis 2018.

7 These are sometimes even deemed to include other people reduced to objects, as in the case of slavery.

8 Hirsch Ballin 2020.

9 Ellis 2018.

10 Toit 2016.

11 Esposito 2014 (2015).

to merely controllable objects that command no respect. This imbalance that is wont to creep into thinking on human dignity has been highlighted with growing emphasis since the end of the previous century.¹² Provided this criticism does not lead to a relativization of human rights, the awareness of a respectful association between human beings and their social and natural environment can help strengthen the foundations of human rights. In some legal systems this has been embodied in a recognition that, alongside man's commonly understood legal subjectivity, elements of nature also have rights.¹³ The extent to which and the way in which artificial intelligence should be included by analogy in the discourse on the protection of personal dignity is a question we shall return to at the end of this essay.

There is a further development under way that results from the paradox described earlier. If the human community is exposed to changes in an uncontrollable (or no longer controllable) complex environment, the protection afforded by human rights may be too weak. The desire for dominance that led to the Anthropocene will not simply disappear, as has been demonstrated once again by the coronavirus crisis of 2020/2021.¹⁴ This desire is a strong force in both economic and political terms, and its cultural driver is consumerism. It will seek to sustain itself, even at the expense of the protection of human beings in vulnerable positions.¹⁵ Examples include countries which, for economic reasons, take insufficient action to combat the spread of infection in poor districts and regimes that use the combating of the pandemic as a pretext for strengthening repressive controls, as a precursor to an authoritarian political order. In countries with a solid democratic structure there are nascent movements demanding 'freedom' from measures to combat the pandemic, whereas in fact they are sacrificing elderly people and other vulnerable groups to their own self-interest.

This raises the question of whether human rights can hold out as a means of protecting personal freedoms and life prospects if coercive measures appear to be justified by a sustained, kaleidoscopic state of emergency comprising drought, floods, food shortages, wars, civil wars, pandemics and (real) mass migration. Are not human rights just as vulnerable as the people who – both protected and protectors – attach importance to them? Is the protection afforded by human rights thus not vulnerable precisely when it is most needed?

¹² See Heeger 2014; Schaber 2014.

¹³ See Pietrzykowski 2017, 70: "It is necessary to distinguish full-fledged persons from non-personal subjects of law and make the latter clearly distinct from mere things." These are preliminary steps towards a post-anthropocentric legal theory; see Hirsch Ballin 2020, 75.

¹⁴ See Biebrichter 2021.

¹⁵ 'Citizens' have given way to 'consumers', according to Tjeenk Willink 2018. See Blühdorn and Deflorian 2019 on the negative effects of this on the possibilities for environmental and climate policy.

The rise of artificial intelligence coincides with the transformative crises of the incipient Anthropocene era. The emergence of artificial intelligence raises high expectations that people will be able to regain or even increase control over their lives, precisely at a time of looming major needs. On the other hand there are fears that artificial intelligence, alongside climate change, wars and pandemics, is itself one of the forces that humanity has summoned but may no longer be able to control. An example is when a large number of cases appear to require binding assessments (for example on legal positions or admission to and removal from posts or areas) to be dependent on an artificial intelligence-based estimation of the impacts.

This is the origin of the question this essay seeks to answer, namely the relationship between the rise of artificial intelligence and the possibility of dignified coexistence based on trust in the rule of law. *Is artificial intelligence one of the characteristic forces of the Anthropocene that will dethrone the freedom-seeking citizen of the democratic state governed by the rule of law? Or, in such circumstances, does it provide an opportunity to achieve a democratic society in which human rights are a key factor, subject to certain rules and conditions?* We will focus further on this question in section 2.

The answer to this question is open-ended. In this essay we will examine the possibilities for directing policy towards the second alternative. The fundamental rights of each human being¹⁶, their characteristic context of a ‘democratic society’ and their most recent articulation, from a development perspective, as Sustainable Development Goals (SDGs)¹⁷ will serve as the benchmarks. This examination must constantly contrast these possibilities with the other possibility: that of the first alternative and of the consequent threat of a community being overwhelmed by needs and emergency measures. This also impacts our understanding of human rights. The desirable linking of artificial intelligence to development and the fulfilment of human rights will only work if it is based on an ‘inclusive’ understanding of human rights themselves: i.e. not only as a ‘conventional’ constraint on government action, but also as principles of a political and legal order that enables everyone – including the historically disadvantaged and those made vulnerable by their life situation – to realize their life projects in freedom and help build a democratic society.

16 Fundamental rights (a synonym for basic rights) of every human being are ‘human rights’. There are also fundamental rights related to citizenship of a particular state, including voting rights; these are referred to as civil rights.

17 AIV 2019.

2. Artificial intelligence in relationships between persons and bodies

Artificial intelligence is defined with varying breadth in the public debate. At one extreme even algorithmic processing of big data is deemed to constitute artificial intelligence, while at the other extreme we find machine learning and, within that, deep learning. The draft regulation on artificial intelligence (Artificial Intelligence Act) published by the European Union on 21 April 2021¹ defines an ‘artificial intelligence system’ as “software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with”. This definition must therefore be read in conjunction with the list in Annex I to the draft regulation. This includes:

- a. Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
- b. Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;
- c. Statistical approaches, Bayesian estimation, search and optimization methods.

Legal definitions such as these must be specified precisely. In this essay we take artificial intelligence to mean algorithmic data processing that uses advanced technology and fulfils the succinct description offered by Margaret Boden: “Artificial intelligence (AI) seeks to make computers do the sorts of things that minds can do.”²

In a preliminary study for the Netherlands Scientific Council for Government Policy (WRR) Bakker and Korsten describe artificial intelligence as a system technology with all kinds of uses.³ This description immediately gives rise to *two observations*. The first is that this emphasis on open-endedness of potential uses places artificial intelligence in a position where it is close but not equal to human intelligence. Artificial intelligence is developing into an information technology with in principle unlimited but not yet all-encompassing potential. Human intelligence is also ‘open’, i.e. not limited to the processing of innate reflexes, but – as Thomas Aquinas said of the soul – *quodammodo omnia*,

1 com(2021) 206 final.

2 Boden 2018.

3 Bakker and Korsten 2020.

in a sense all things.⁴ *All things* in the sense of being entirely open to what is yet to be discovered, but of course not ‘absolutely’ all things, in advance and comprehensively, so only ‘in a sense’. The intangibility of ‘all things’ is linked to the limitations of human learning capacity and the availability of energy and information sources.

It is noteworthy that many people who write on artificial intelligence consider its definition to be a difficult issue, whereas the definition of human intelligence appears to be obvious and generally understood. Actually defining it is no less difficult, however.⁵ Creativity, inventiveness and serendipity result from an unparalleled ability of cognitive integration, for which, through cooperation between neurologists, system theoreticians and computer scientists, models are now being developed which bear little resemblance to a conventional computer. The competitive formation of interconnected serial and parallel processes in the brain forms the substrate of the openness of the human mind.

‘Deep learning’ with artificial intelligence can imitate the characteristics of human creativity and incorporate these in the operation of computers. In the case of human beings the situation is different: a person cannot be described as ‘being’ his intelligence; he does not ‘use his mind’ – like a computer – to achieve a prescribed purpose but to shape his own life.

The second observation is related to this. It concerns an implication of the (unlimited) potential uses of artificial intelligence for purposes outside the AI system. This means there are always human masters – defined in terms of the General Data Protection Regulation (GDPR) as controllers⁶ and in the system used in the new draft regulation on artificial intelligence as providers and users – who determine the purpose for which the artificial intelligence system is established and who can connect or disconnect that system to or from the energy and information sources without which artificial intelligence cannot operate. In informal terms, if an AI system operates in a way we dislike and even disregards human instructions, we can always pull the plug out. It is certainly possible that AI systems are being or already have been developed that control other systems, including the management of such ‘plugs’, but entirely

4 Thomas Aquinas, *Questiones Disputatae: De Anima*, *Questio unica de scientia Dei*, Art. II, Conclusio. Shanahan 2012.

5 See the definition in Article 4(7): “the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law.”

independent systems accompanying humans onto the battlefield can for now be attributed to science fiction that is more fiction than science.

Logically, the use of AI systems or complexes of such systems by human masters ('controllers') can be relationally embedded in two ways. The first possibility is that such use relates solely to the user, for his own assessments and decisions, as an *instrument*, i.e. as a technical support. The other possibility is that such use forms part of relationships with others (other individuals, partnerships or other groups, or entire communities), i.e. as a *medium* for evaluating or normative interaction, i.e. as a link that is not controlled by the human actor at the time of assessment. Automated marketing based on individual preference measurement and automated decision making, including 'autonomous weapon systems', are examples of this. The GDPR offers individual consumers a legal means to protect themselves from such unwanted marketing by withholding or revoking consent for the use of their personal data. With a few exceptions, Article 22 of the GDPR also prohibits automated decision-making without human intervention if it is disadvantageous to the data subject.⁷

This second category of potential uses of artificial intelligence implies that artificial intelligence itself is positioned in *legal relationships* as a (conditional) autonomous means of assessment and/or decision-making. This autonomy, unlike human autonomy, is not the result of 'self' – the way in which the subject relates to his own knowledge and ability – but an autonomy *conferred* on the system by the architect.

Reference can be made here to comments made in publications on the regulation of artificial intelligence with regard to the place it will occupy – and to some extent already has occupied – in relations between persons, bodies and organizations that are also legal relationships, including those under administrative and labour law, and other private law relationships.⁸ Artificial intelligence is a medium in the relationship between two or more legal subjects, possibly even between a public law body vested with authority and an entire community of persons subject to the same legal system. Unlike traditional media, artificial intelligence does not only reproduce epistemic and normative content, but also produces it. More advanced technology makes this production of original content possible and to some extent makes AI systems similar to natural human intelligence.

7 Article 23(1) states: "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her." The subsequent paragraphs contain conditional exceptions to the prohibition.

8 See for example Hermstrüwer 2020 on the development of legal relationships under administrative law.

The field of law and public administration has seen the development of systems that use artificial intelligence as an instrument as well as AI systems operating as a medium in legal relationships between two or more subjects. Artificial intelligence and, more generally, computers and robots are particularly suitable for carrying out routine-intensive tasks, but the nature of those routines differs greatly. Depending on the level of the available artificial intelligence, its use may characterize it more as an instrument or as a medium for decision-making.

Despite the increasing refinement of artificial intelligence, it still lacks interpersonal empathy; such empathy can at best be simulated. So as not to overstate the potential of artificial intelligence, we must be mindful that interpersonal relationships, including legal relationships, will ultimately always be part of what is processed in the physicality of human beings. The human mind is not only a kind of ‘controller’ in relation to a person’s own body, but is existentially bound up with it.⁹ Human dignity is violated if the human body is harmed, traded or abused; social intelligence can also respond to feelings that are conveyed without words.¹⁰

Our body is therefore necessary, but at the same time restricts the openness of the soul ‘to all things’, as mentioned above. The body also bears risk and may be attacked and used as a means to restrict the interhumanity of our spiritual life. Cultural, economic and political interaction can be promoted by opening up borders, or can be truncated by closing them and by excluding or enclosing individuals.

To conclude this section we will consider whether the difference between the creative openness of the human mind and artificial intelligence based on continued processing of acquired knowledge does indeed produce a definitive dividing line. The potential to approximate the functioning of the human mind – Margaret Boden’s definition of artificial intelligence referred to at the beginning of this section – is steadily growing. Studies of the current state of development of artificial intelligence show that it increasingly approximates human creativity. In his book on ‘The Digital Mind’ Arlindo Oliveira points to the far-reaching developments in the digitization of models of the human mind and in brain-machine interaction.¹¹ He believes the barrier that currently distinguishes the relative autonomy of existing artificial intelligence systems

9 A very brief theological digression: ‘immortality’, however simplified, cannot be construed as continued temporality; ‘eternity’ cannot be counted in centuries. What we can conceive of as human coexistence is tied to physicality.

10 See Hollak 1972 on man as a physical spirit.

11 Oliveira 2017.

from the full autonomy of what is labelled the ‘mind’ may well not be insuperable. If ‘digital persons’ were built, questions would be raised about their rights and obligations, the responsibility for their behaviour, and the implications of the fact that such artefacts with the characteristics of persons could be reproduced in unlimited numbers. (Even if we wished to identify them as bearers of rights and responsibilities, we would be better advised to exclude voting rights.) Whether the human mind can actually be copied to a digital system must be doubtful, however. Developments in research into human consciousness point to the phenomenon of quantum consciousness.¹² The implication would be that it is actually impossible to recreate the human mind.¹³ Substantial growth of artificial intelligence abilities must nevertheless be expected. It is already technologically possible to write text to the standard of a novel.¹⁴ Even complex reasoning in court judgements can be reconstructed and then applied to new situations.¹⁵

The same applies to artificial intelligence as a medium in the exercise of civil rights. People’s physical vulnerability is characteristic of the current crisis caused by a pandemic. In our interpersonal interaction we are limited by remoteness and a predominant reliance on telecommunication. Could the intervention of robots and physical structures equipped with artificial intelligence to some extent compensate for and make more bearable the limitations placed on personal and hence also physical meetings, or would that possibility confirm or even exacerbate our physical vulnerability? That will depend on whether artificial intelligence is available to everyone, particularly in democratic, verifiable structures, or is controlled by dominant and hegemonic powers.

12 Li et al. 2019.
13 Oliveira points out that “the quantum configuration of a system cannot be copied without interfering with the original” (Oliveira 2017, 259).
14 Price 2020.
15 Ashley 2017.

3. Threatening and repressive artificial intelligence

From the perspective of privacy and consumer protection, the use of artificial intelligence for marketing purposes is an important subject of debate and concern. This essay, however, focuses more on the use of artificial intelligence to fulfil public functions, because that is where the human rights situation is most urgent. Much has already been written about the use of artificial intelligence in administrative and legal assessments. Attention is focused particularly on possible impairment of fundamental rights, particularly the right to privacy, the right to a fair trial and the right to equal treatment. A frequently discussed danger is that prejudices will creep into the predictions that are fed into artificial intelligence systems,¹ for example because the calculations draw on a database that overrepresents suspects or convicts with characteristics that are irrelevant to the fair administration of criminal justice. Arithmetically intelligent but socially dumb calculations may then attribute a predictive value to these characteristics.

The potential victims are particularly people in disadvantaged positions, such as dissidents, the elderly, migrants and undocumented persons. These risks are difficult to eliminate; it is unrealistic to believe they can be corrected incrementally on the basis of ‘artificial intelligence for the good’. Jared Moore writes as follows about those who play down the problems²:

To paraphrase, they argue that the risks of AI technologies are important, but that the risks can only be solved by further development of AI technologies. The utopic notion of economic liberalism employs the same sort of rhetoric: because the free-market ideal has never been achieved, one can always argue that its failures are due to insufficiently free markets. Likewise, data scientists, instead of addressing critiques, focus on how to realize the ideal of datafication in society; they reinforce a technological determinism.

¹ WRR 2016. In the unsolicited opinion of the Advisory Division of the Council of State on the impact of digitization on relationships governed by the rule of law of 31 August 2018 (Parliamentary Papers II 2017/18, 26643, no. 557, §1) the Division also writes: “He also risks being confronted with decisions based on profiling and statistical relationships. This will not show that the citizen has acted culpably, but will merely arouse a suspicion based on general characteristics. It will give rise to a statistical reality that differs from the concrete facts.”

² Moore 2019.

The probability calculations in (selectively) self-learning algorithms project data, such as undesirable behaviour linked to demographic characteristics, onto all those who belong to the same statistical class. In order to provide a serious answer to the disadvantage that people may consequently suffer, we must delve more deeply into the operation of these systems.

It is therefore not – directly at any rate – a question of the qualities of the *intentions* with which artificial intelligence is used, but of the quality of its *use* and *operation*, either as an instrument or as a medium for assessment and decision-making. The key issues are integrity, care (for example the right definitions) and – as is always the case in subjects relevant to basic law – the proportionality of uses of artificial intelligence. In its report on ‘Big Data in a Free and Secure Society’ the WRR already issued recommendations in 2016 to strengthen the transparency, care and verifiability of the processing of large data collections with algorithms.³ In the draft regulation of April 2021 the European Union focuses with good reason on the use of data and sets related standards. It explains the reasons as follows⁴:

Artificial Intelligence (AI) is a fast-evolving family of technologies that can bring a wide array of economic and societal benefits across the entire spectrum of industries and social activities. By improving prediction, optimizing operations and resource allocation, and personalizing service delivery, the use of artificial intelligence can support socially and environmentally beneficial outcomes and provide key competitive advantages to companies and the European economy. Such action is especially needed in high-impact sectors, including climate change, environment and health, the public sector, finance, mobility, home affairs and agriculture. However, the same elements and techniques that power the socio-economic benefits of AI can also bring about new risks or negative consequences for individuals or the society. In light of the speed of technological change and possible challenges, the EU is committed to strive for a balanced approach. It is in the Union interest to preserve the EU’s technological leadership and to ensure that Europeans can benefit from new technologies developed and functioning according to Union values, fundamental rights and principles.

These risks, even if they arise primarily on an individual level, concern rule-of-law characteristics of the constitutional order and affect fundamental rights. The risks to the democratic character of the constitutional order are no less serious. A notorious case concerns the manipulation of voter preferences using algorithms developed by the company Cambridge Analytica. Sofia Ranchordás writes in this regard that this scandal “in 2017 exposed a collective

unawareness of how much disinformation individuals receive on a daily basis and how it influences their voting decisions, how their data is abused, and the lack of information as to how to protect themselves against both private and state surveillance”.⁵ The democratic constitutional order relies on the recognition and implementation of fundamental rights: the freedom of public debate described in Article 10(1) ECHR as “the freedom to hold opinions and to receive and impart information and ideas”, the right to take part in free elections (Article 3 of Protocol No. 1 to the ECHR), protection of honour and good name as a right that limits freedom of expression pursuant to Article 10(2) and the prohibition of discrimination with regard to these rights and freedoms (Article 14 ECHR).

Unlike the risk that that governments will impose unauthorized restrictions, i.e. violations, of these rights – a risk that appears somewhat dated in the Netherlands and neighbouring countries – the threats posed by artificial intelligence are difficult to identify and particularly difficult to trace back to particular individuals or organizations. The trolls and other intruders often come from far across the border using the World Wide Web. Although artificial intelligence can certainly benefit the expertise required in a properly functioning constitutional system, there are now emerging threats to *each* of these fundamental rights. These threats are reaching a new level of refinement and elusiveness with the production and distribution of ‘deep fakes’, i.e. counterfeits produced with advanced technology that appear to associate known persons, such as politicians, with remarks and conduct that undermine their position.⁶

Manipulation and deception strike at the heart of democracies governed by the rule of law. Onora O’Neill, author of authoritative works on *trust* and *trustworthiness*, said of this in an interview in 2018⁷:

Not deceiving is one of the fundamental duties. When I think about technology, I wonder whether we will have democracy in 20 years because if we cannot find ways to solve this problem, we will not. People are receiving messages and content which is distributed by robots, not by other human beings, let alone by other fellow citizens. It is frightening.

5 Ranchordás 2021.

6 Schindler 2020.

7 https://www.huffingtonpost.com/entry/onora-oneill-i-wonder-whether-we-will-have-democracy_us_5a4f8a12e4b0cd114bdb324f, cited by Tasioulas 2019.

These dangers to democracies governed by the rule of law are being exacerbated by measures taken to tackle the pandemic that broke out in 2020.⁸ As a form of state intrinsically bound up with human rights, the democracy governed by the rule of law is fragmenting worldwide as climate change and its effects on food and other essential resources, ethnically fuelled wars and civil wars and pandemics play into the hands of authoritarian movements. This confirms the Schmittian model of politics as enmity (parties that see themselves as *friends* against *foes*), occupation of territories (*Landnahme*) and sovereign control of the state of emergency (*Notstand*) and hence increases its attractiveness among people who have grown insecure.⁹ Artificial intelligence also plays a role, because it makes it possible to deceive people by means of conspiracy theories¹⁰ and other falsifications of politically relevant facts and to use individual targeting to fuel support for leaders who will entice or coerce them into continued acquiescence.

Günter Frankenberg points in particular to the risks associated with making the state of emergency permanent, setting aside the normal operation of a democratic constitutional state. The trigger may be a natural or environmental disaster, an infectious disease or a revolution.¹¹ Risk prevention is good, but ‘hyperprevention’ of all kinds of life risks negates the operation of the rule of law as a basis for trust.¹² In this context the use of artificial intelligence starts to look grim.

8 “The protection of human rights, including freedom of expression, internet freedom and independent journalism, is the foundation of well-functioning democracies” (Sweijts et al. 2020).
9 Frankenberg 2014. Cf. Müller 2001 on the after-effects of these compromising ideas.
10 Butter 2018.
11 Frankenberg 2014.
12 Ibid., 94, 110, 146.

4. Supportive and liberating artificial intelligence

This section describes the prospect of controlled and well-considered use of artificial intelligence as an aid to or a means of decision-making, taking into account qualitative conditions. This prospect is the complete opposite of the subject of the previous section. It must be noted that it can only be achieved if at the same time action is taken to prevent the occurrence of the serious risks referred to above.

In a report compiled on behalf of the Ministry of Internal Affairs and Kingdom Relations, Vetzo, Gerards and Nehmelman discuss the broader issue of the relationship between the use of algorithms and fundamental rights. Artificial intelligence is a facet of this subject, in which the distinctive characteristic is ‘a high degree of autonomy’: “AI applications can perform complex tasks, without human monitoring or supervision.”¹ These build further on the “presence of large volumes of data (Big Data)” and the Internet of Things that gathers data and feeds it back to systems where it is analysed using algorithms.²

Aside from the possible invasion of privacy, the authors consider that excessive individualization of decision-making is problematic. It can undermine “human dignity and personal autonomy”.³ We have already indicated elsewhere that in the not too distant future artificial intelligence will give courts and administrative bodies previously unimaginable abilities to base decisions on very refined, algorithmic rules.⁴ Busch and De Franceschi describe such rules as “fine-grained” and “personalized”, but for the parties concerned such microscopic standard-setting will not necessarily be any more convincing.⁵ They suggest that coarse-grained rules deliberately disregard personal characteristics and thus create a “protective space” of individual freedom into which the law does not intrude.⁶ The advantages and disadvantages of ‘customization’ have also been discussed in relation to the preliminary advice given by Sandra van Heukelom-Verhage for the Administrative Law Association (*Vereniging voor bestuursrecht – VAR*).⁷ Extreme refinement, when enshrined in law, is therefore disproportionately intrusive. The relationship between the different phases in the administrative law standardization chain

1 Vetzo et al. 2018.

2 Ibid.

3 Ibid.

4 Hirsch Ballin 2020.

5 Busch and De Franceschi 2018, 413.

6 Ibid., 424.

7 Wolswinkel 2020a.

(law, delegated legislation, policy rules and concrete decisions) is consequently changed, eroding the constitutional guarantee of the principle of legitimacy and requiring a different implementation of democratic control in relation to rule-of-law guarantees.⁸

Examples of excessive data recording include citizen scoring in China and some African countries⁹: data on desirable and undesirable conduct is recorded to compile information that can be passed on and fed back through a ‘social credit system’. This extremely fine-meshed form of conduct regulation intrudes far more deeply into people’s lives than the ‘ordinary’ criminal law and is totally at odds with a democratic society.

In order to counter the risk of biases in algorithms and the use of – proper or improper – profiling, Vetzo et al. recommend strengthening the procedural rights of the parties concerned.¹⁰ The “neutrality and transparency of algorithms” must be assured.¹¹ Wolswinkel provides more detailed recommendations on the establishment and assessment of algorithmic decision-making in public administration. With reference to the *Aerius* case law of the Administrative Jurisdiction Division of the Council of State¹² and other sources he finds three criteria: 1. The use of objective criteria in decision-making rules, 2. The transparency of those criteria, and 3. The consistency of the resulting decisions, in order to ensure that the consequences for one interested party are proportionate to those for another.¹³ The question is whether the steps accepted in case law, however important they are, will ultimately prove sufficient. For example, is a prohibition of the use of a decision-making rule between the parties in an administrative lawsuit sufficient if that decision-making rule is embedded in a generally applied working process? “If the application of an algorithmic decision-making rule leads, for example, to a violation of fundamental rights, for example in the case of discriminatory bias, there appears to be more need to apply an absolute prohibition of use, which nevertheless appears to require legislation,” according to Goossens et al.¹⁴

8 Goossens et al. 2021.

9 Straaten 2019.

10 Vetzo et al. 2018.

11 Ibid.

12 ECLI:NL:RVS:2017:1259; ECLI:NL:RVS:2018:2454.

13 Wolswinkel 2020. The principles of materiality, care, proportionality and rationale are central in Adriaansz 2020.

14 Goossens et al. 2021, 14.

Such requirements are concrete safeguards against the use of algorithms with hidden biases, a danger which WRR had already highlighted in its report on Big Data.¹⁵ The importance of the criteria identified by Wolswinkel lies in the fact that they not only counter such undesirable use of artificial intelligence, but also indicate clearly how it can be properly used. His criteria therefore have a different perspective on algorithmic decision-making than a mere response of rejection and he explains this on the basis of the ambiguity of ‘arbitrariness’. In the traditional approach, which he refers to as ‘analogue administrative law’, arbitrariness is a scourge, as it causes unwarranted unequal treatment. In cases in which the same good cannot accrue to everyone, the algorithms of ‘digital administrative law’ offer the possibility of lawful ‘arbitrariness’ due to randomization. The importance of preventing personally or politically motivated arbitrariness in administrative law means that interested parties – reflecting the previously mentioned requirement of Article 22 of the GDPR – may have an interest in the use or support of algorithmic decision-making.¹⁶ Artificial intelligence can be designed in such a way as to exclude unlawful and careless elements of the decision-making process. This is therefore the opposite of the aforementioned microscopic, ‘fine-grained’ regulation.

Good use of artificial intelligence is also possible from the perspective of judicial practice. Micha-Manuel Bues refers to the possibility of mutual learning by man and machine. Instead of ‘artificial intelligence’ it is then more accurate to speak of augmented intelligence.¹⁷ Appropriate software can be used to access and make optimum use of experts’ practical knowledge, for example.¹⁸ Such considerations point to the importance of professional handling of artificial intelligence in law and administration, with professional knowledge that is not only of a technical nature.

The conditions under which quality improvement can be achieved with artificial intelligence in the legal and administrative field must therefore be among the benchmarks we seek in this essay, but we must remain mindful of the associated risks noted in numerous observations, such as the undermining of privacy, biased assessments and invisible unequal treatment. Section 2:4 of the General Administrative Law Act describes this in terms appropriate for human decision-makers: “The administration shall perform its task without bias.” Artificial intelligence must also be unbiased, but that is only possible if the designated benchmarks are complied with. In the administrative sphere it

15

WRR 2016.

16

Wolswinkel 2020.

17

Bues 2018.

18

Ibid., 284.

is necessary particularly to guard against the use of artificial intelligence that is not or not sufficiently objective and transparent, and which consequently undermines the impartiality required of administration in a free society. The law, and not the functionality of an algorithm for administrative optimization, must determine which decision-making factors can be deemed relevant.

As usual, such guarantees concern the rights and freedoms that citizens of a democratic constitutional state can be expected to demand. Artificial intelligence will then not only threaten to constrain these abilities of free citizens in a democratic society, but will also support them.

It is important to consider this point again with regard to the anthropological basis of this perspective on the phenomenon of artificial intelligence. In the second section of this essay it was already emphasized that the human mind, unlike artificial intelligence that can be switched on and off, is more than just an operating system of a locomotor system. Due to their own physicality, people live in a world which they can endlessly explore, and which in the Anthropocene they have irrevocably changed, but which – despite long-cherished illusions of control and the resulting degradation of living nature – they will not be able to subjugate.

These are the characteristics of human freedom that we seek to protect and realize with the legal and political institutions of human rights. Human existence is by and large *precarious*, as can be seen among other things in the methods of repression that people use in their physicality, and in the concern that artificial intelligence will threaten their spiritual freedom. People's equality and autonomy cannot be taken for granted if legal distinctions and restrictions are eliminated. Therefore, as Martha Albertson Fineman argues in a fundamentally critical article, thinking about human rights must take account of inherent and unavoidable inequalities in smaller and larger societal relationships.¹⁹ She posits a theory based on the vulnerability of human beings. That is not only vulnerability to physical or emotional harm, nor one limited to so-called vulnerable groups, but “the continuous susceptibility to change in both our bodily and social well-being that all human beings experience”. Here we must consider among other things the different ways in which “power and privilege are conferred through the operation of social institutions, relationships and the creation of social identities, sometimes inequitably”.²⁰ These key principles allow a sharper focus on human rights and public responsibility for their fulfilment. The idealized identity of people as political and legal

19

Fineman 2017.

20

Ibid.

subjects must be confronted with the reality that we are “embodied beings”, as “social beings embedded in social institutions and relationships”.²¹

A positive impact of artificial intelligence on human rights must be linked to these anthropological and social realities; taking into account requirements of care, transparency and proportionality is a precondition, but is not sufficient. A key issue remains how artificial intelligence will be implemented in the fulfilment of public functions. In contrast to the neoliberal fixation on personal responsibility and a small state, public functions must be aimed at giving people access, as rights holders, to “resources that will enable them to endure or prosper from change, even harm, throughout institutions and relationships across the life-course.”²² Such an approach concerns individuals in their social context and focuses on the promotion of resilience, which gives them “the means and ability to recover from harm, setbacks and the misfortunes that affect our lives. [...] Resilience is the critical, yet incomplete, solution to our vulnerability” and involves access to physical, personal, social, ecological and existential resources.²³

This viewpoint goes to the essence of our concept of human rights, and hence also their importance for a recalibration of the (possible) regulation of uses of artificial intelligence that enter society. The promotion of resilience in people’s lives, in response to their inherent and inevitable vulnerability, should be the benchmark for artificial intelligence that benefits the fulfilment of human rights.

Such a dynamic concept of human rights – in contrast to the view of human rights as individual protection in a static society, which was widely held until recently – was also developed in the research by William Paul Simmons, in the footsteps of Emanuel Levinas, into the significance of human rights for “the marginalized Other”. In particular he considers the obstacles that refugees and other marginalized people have to overcome in realizing their life projects, or – in the Latin American context, their *proyectos de la vida*.²⁴ The approach to human rights based on insight into their vulnerability and the need for greater resilience is far more comprehensive, however, and could be described as universal, like the idea of human rights itself. The implementation of the rights laid down in a universal declaration may differ, however, depending on the nature and cause of people’s vulnerability in a specific sociogeographic context.

21

Ibid.

22

Ibid. With regard to neoliberal scepticism towards democracy and the part it plays in causing the political vacuum in the operation of the EU, see also WRR 2018, 54, and Biebricher 2021, 160 and 325.

23

Fineman 2017.

24

Simmons 2011.

Although Fineman relates vulnerability and resilience universally to the human condition, people in the global South, including Africa, the Indian subcontinent, South-East Asia and Central and South America are particularly exposed to need, threats and other dangers. Warnings are currently being voiced about the potential of artificial intelligence to further disadvantage the global South and its populations and expose them to poverty and experimentation. “The degree to which the AI industry is willing to experiment on human populations, in the name of innovation, should make us uncomfortable,” says Chinmayi Arun.²⁵ In these parts of the world people are dependent on the technology industry located in the United States, China or the European Union. “Vulnerable Southern populations in particular are at risk from the surveillance and other forms of discrimination, bias and poorly tailored outcomes that will result from AI that is designed with no regard to their local contexts.”²⁶

Artificial intelligence is also important, however, for the entitlement to be free of need and have good collective facilities, which equally belongs to the sphere of fundamental rights. Artificial intelligence is seen as a technology offering unprecedented predictive power.²⁷ This raises the question of whether artificial intelligence can be further developed in such a way that it supports the combating of those risks. In order to pursue the objective of ‘zero hunger’ in India, for example, the resources of ‘pioneering technologies’ are used, including blockchain and artificial intelligence.²⁸ Aside from the serious risks of ‘killer robots’, for example, the work of Human Rights Watch also includes opportunities to highlight crimes against humanity by using “new technological methods — including AI — for its investigations”²⁹:

These new tools include remote sensing via satellite and drone data, analytics from public datasets, and investigations using videos and photos posted to social media. Remote sensing is essential in situations where researchers can’t access a conflict zone or closed country — a major issue for the human rights and humanitarian community.

Artificial intelligence can therefore also help in detecting and eliminating inequalities, including in the Western world. The means available to reduce economic inequality, for example in relation to technologically refined marketing, are still very underdeveloped. Government policy, for example, has to make do with clearly deficient ‘purchasing power snapshots’, as

25 Arun 2019.

26 Ibid.

27 Agrawal et al. 2018.

28 Chavan and Breyer 2020.

29 Salian 2019.

decision-making has been heavily dependent on forecasts derived from policy models built by humans. The reality of economics and the labour market is so complex, however, that more reliable forecasts have only been made possible by data processing based on self-learning algorithms. Maarten Goos and Anna Salomons point to the importance of predictive algorithms, for example with regard to the labour market and unemployment data.³⁰ Similar developments are possible with regard to other determinants of people's socioeconomic position. Both in the Western world and the global South, artificial intelligence can help to improve the quality and flexibility of government policy.

Artificial intelligence has highly significant potential with regard to SDGs. Goralski and Tan cite the following examples³¹:

- “Smart water management systems powered by ai replicate the way that humans learn in an ever-changing environment that maximizes decisions and investments in the water management infrastructure”;
- “PlantVillage, a research and development project based at Penn State University, brings the power of ai to small farmers of 2 ha or less that produce most of the developing world's food”; and
- “Innovative use of ai to evaluate the purity of water systems will allow municipalities and eventually individuals to inexpensively identify waterborne disease infested waters and map bacteria and eventually viruses in those waters.”

These examples again relate to the global South, but SDGs apply to all countries, including the Netherlands. Policy for the development of AI applications with a view to fulfilling social rights is therefore important for both the Dutch contributions to development issues elsewhere and to the policy in Dutch and European society. This applies, for example, to the use of artificial intelligence in ‘smart’ city management. In a sensible combination of governance and openness, urban planning – including economic, social and ecological development – can be better geared to the real needs of the population.³²

30 Goos and Salomons 2021.

31 Goralski and Tan 2020.

32 Davis 2020.

5. Benchmarks

The main question raised by the emergence of artificial intelligence is actually *how we see human intelligence*, i.e. man's intelligence as a social being. After all, it determines how we should shape the relationship with artificial intelligence and, as far as necessary, regulate it. Neoliberalist theory sees *homo oeconomicus* as subordinate to the pursuit of macroeconomic growth¹; artificial intelligence is then a possible contribution to revenue models and only needs to be regulated for competition purposes. The situation is different when human intelligence is construed as a capability that allows respectful relationships – and mutual dialogue – with fellow man. A person's intelligence transcends the individual's life and extends to the lives of others and the community as a whole, leading to questions which, in the political theory of the democratic constitutional state, are associated with human rights. For various reasons the traditional sources of legitimacy are insufficient in the case of algorithmic administration.²

This essay therefore explores the place that artificial intelligence can acquire in relation to human intelligence with regard to fulfilling fundamental human and citizens' rights, particularly in the performance of public functions. This question cannot be answered in the form of a to-do list, because all the terms in this research problem are dynamic and developing, both the phenomenon of artificial intelligence itself and the nature and scale of public functions, and – not least – the notion of human rights themselves. Hence the most appropriate conclusion to this essay appears to be to propose a number of benchmarks for artificial intelligence policy derived from the current discourse on human rights.

In cyberspace, the usual starting point is the obligation under international law to “respect, protect and fulfil human rights obligations”.³ Nonetheless, specifically with regard to artificial intelligence, more detailed or even new standards have been proposed in treaties (or draft treaties) and international statements of principle.

The national level is clearly insufficient for these cross-border developments. The OECD has now issued a *Recommendation of the Council on Artificial Intelligence*,⁴ with a scope comparable to that of the more robust guidelines

1 According to Wendy Brown, who reviews the beliefs of Foucault here. See Brown 2015, 83.

2 Meijer et al. 2021.

3 See Rona and Aarons 2016, 4-5.

4 OECD/LEGAL/0449, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>, of 22 May 2019. On the subject see Vöneky 2020.

of 8 April 2019 issued by the High Level Expert Group on AI established by the European Commission; these guidelines have now been adopted in the aforementioned draft regulation. Key points of the Ethics Guidelines for Trustworthy AI include human supervision, privacy, respect for diversity, combating discrimination, responsibility towards future generations and fairness and accountability; these points must be assessed on the basis of specific test criteria.⁵ Andrea Renda notes in this regard that monitoring ‘ethics’ in the development of artificial intelligence actually concerns the trustworthiness of the technology.⁶ An initiative group centred on *Die Zeit* has gone a step further by setting out fundamental rights in the form of a *Charter of Fundamental Digital Rights of the European Union*,⁷ which could be made legally binding. This remains a long way off, however, and may follow the enactment of the currently proposed regulation known as the Artificial Intelligence Act.

In addition to the question of whether and, if so, how artificial intelligence could be standardized, the question arises as to whether the Dutch Data Protection Authority’s current supervision of the processing of personal data remains sufficient or whether a specialist authority should be established, as called for in the United States.⁸

These European developments are of great importance. The aim is always to prevent people’s individuality being lost in probability calculations. A lack of safeguards

raises the spectre of a government that intrudes invisibly but efficiently into people’s private lives, seeing individuals as controllable objects that can be lumped into general categories. This was addressed in the third section of this essay. The fourth section concerned the possibility of an opposite approach: is reciprocity possible in the use of the medium of artificial intelligence in legal and other relationships? This will depend partly on whether the medium is controlled solely by one party or is incorporated in the assessment processes on a mutual basis. A somewhat deeper question was whether artificial intelligence could be used to reposition ourselves more strongly in the complexity of our social environments. This brings us to the following three benchmarks.

5 Ethical Guidelines for Trustworthy AI of the EU High Level Expert Group on Artificial Intelligence. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

6 Renda 2020.

7 <https://digitalcharta.eu/wp-content/uploads/DigitalCharter-English-2019-Final.pdf>

8 Wheeler et al. 2020.

a. **Connecting artificial intelligence with people's life projects, taking into account their vulnerability.** The formal conception of equality abstracted from real-life situations, which has been criticized by Martha Fineman, concerns not only the right to equal treatment but also the full set of 'classic' human rights that should apply equally to everyone. Their form and content are based on the setting of standards and limits for government action, but that means they overlook the differences in people's life situations that feature in a more comprehensive vision of human rights.

Her call for the key criterion to be people's vulnerability also means that in our thinking about the fulfilment of human rights people must no longer be lumped together in undifferentiated populations detached from their life circumstances. In this regard the effects of using artificial intelligence are ambiguous. Policy based on predictions generated with artificial intelligence will mask or even preclude personal diversity. Artificial intelligence can also be used, however, to correct the complex processes that prevent people sustaining or educating themselves and otherwise developing their life projects.

Examples of this are already occurring, as in the combating of food shortages. An international group of researchers from academia and industry led by Koen Peters, Hein Fleuren and Dick den Hertog (ups) has developed mathematical models jointly with the United Nations World Food Programme to analyse the operation of the food supply chain.⁹ Building on that basis, in the Zero Hunger Lab Koen Peters and Hein Fleuren found algorithmic resources to improve food distribution in areas at risk of hunger. "Thanks to the optimus model, in some cases up to 15-20% more people can be fed with the same aid budget. Currently the research of wfp with Zero Hunger Lab is focused on exploring extensions of the optimus model to make it more robust against uncertain parameters such as prices and capacities."¹⁰ The right to water, the right to health and the right to a healthy living environment are among the New Human Rights investigated by an international group of researchers. Many of those rights – including those relating to genetics, IT and indigenous peoples whose life circumstances are threatened – are associated with complex, dynamic processes.¹¹ This also makes it harder to determine what the protective effect of human rights should be and how it should operate. In order to develop the expertise required to determine in an appropriate and hence timely fashion where the right to water or health is threatened, intuitive empirical knowledge will need to be combined with artificial intelligence.

9 Peters et al. 2021.

10 <https://www.tilburguniversity.edu/current/news/more-news/zero-hunger-lab-and-wfp-finalists-franz-edelman-award>

11 Arnauld et al. 2020.

b. Requirements for use as an *instrument* in contexts sensitive to human rights. Artificial intelligence can make predictions and estimates with varying degrees of plausibility. When it comes to imposing sanctions or restrictions on people, however, it cannot take the place of conclusive, verifiable evidence. Artificial intelligence can nevertheless be a valuable and increasingly reliable aid to decision-making, in which objectivity, transparency and testability must be guaranteed. A substantial and, according to reasonable expectations, effective step to increase transparency is the establishment of public databases and their mandatory use for the fulfilment of public functions.¹²

Various aspects of the regulation of artificial intelligence are widely supported. The ‘regulatory guidelines’ in which these can be summarized concern areas including quality, non-discrimination, proportionality, transparency, liability, democratic scrutiny and harmlessness to future generations.¹³ The European Union’s draft regulation on artificial intelligence focuses systematically on the use of these techniques. The regulation is geared systemically to the use of the techniques. Certain types of use are entirely prohibited, while others are designated as ‘high-risk AI systems’ and subject to strict requirements.

c. Linking artificial intelligence in all its phases to humanity, respect for diversity and support for freely accepted life projects. These benchmarks relate both to the purpose for which artificial intelligence is used and to its organization. Where artificial intelligence systems are self-learning and produce decisions independently, there is a danger of overstating the objective quality. That requires a critical approach on the part of those who establish and use AI systems, with an awareness that the facts known hitherto about discrimination by or with algorithms is due to inattention to such risks among the responsible officials.¹⁴

Although human intelligence may show many deficiencies and often fails to match artificial intelligence in terms of objectivity, it will not be possible to dispense with the safeguard in Article 22 of the GDPR until a high level of reliability of artificial intelligence can be guaranteed. When it comes to both the processing and moral assessment of information, the human mind can do more than just reproduce the means by which similar situations were previously assessed.

¹² Wittner 2019.

¹³ Hoffmann-Riem 2020, 11-12.

¹⁴ See Prins 2021.

There are now sometimes calls for such artificial intelligence systems that can emulate the operation of the human mind to be equated for certain legal purposes with human subjects and to be treated as persons or legal entities.¹⁵ There are nevertheless reservations about ‘promoting’ AI systems to holders of fundamental rights. Tasioulas warns against overestimating the capabilities of artificial intelligence and consequently allowing “our ethical thought to be hi-jacked by utopian (or dystopian) speculation based on possibilities that lie, at best, in the remote future”.¹⁶

Rather, we must take a leap forward in *connecting* human and artificial intelligence. A substantial part of the risks and disadvantages ascribed to artificial intelligence actually reflect a lack of critical expertise among people who rely on AI assessments. In order to counter risks, it is therefore necessary not only to set limits on artificial intelligence but also to develop natural human intelligence. Experience shows that natural human intelligence can be developed more effectively, and often much more effectively, if it is able to work with artificial intelligence, for example to determine what should or should not be included in databases used by the AI system.

6. Artificial intelligence in the Anthropocene

People are right when they say – as they so often do – that artificial intelligence will ‘inevitably’ penetrate our lives. But this assertion gives the impression that artificial intelligence is a force originating externally, whereas in fact it is merely a new stage of technological development brought about by man. In this regard artificial intelligence is no different from all the other human-originated developments from which we can no longer escape and which have been grouped under the name of Anthropocene. This does not mean, however, that history will henceforth only be something that happens to man. The practical and hence moral and political question remains how people in the present and future of the Anthropocene will relate to each other.

That is the relevance of the benchmarks embedded in human rights as discussed in this essay. Human rights are a characteristic facet of the type of political order that we call a ‘democratic state governed by the rule of law’. They will only work if the political order also works, i.e. if it is indeed democratic and governed by the rule of law. This means rights must be protected, citizens must have effective control of political decision-making and public functions must be professionally fulfilled.¹

There are processes under way whereby humanity is increasingly exposed to what it has brought upon itself, such as climate changes that are affecting the life situations of large population groups. After all, advanced expertise is required to assess risks and anticipate threats such as famine and disease, and by democratic means, i.e. not on the basis of a permanent state of emergency. Hence the constitutional question arises as to whether artificial intelligence will better equip governments, social actors and active citizens to cope with the complexity of the Anthropocene. That question must be asked even if we already acknowledge that it is impossible to control those complex processes and that any attempts to do so will be undesirable. More specifically, can artificial intelligence help to identify elements in these processes that could still be standardized and adjusted?

The public functions referred to here are among those required by human rights and SDGs. Furthermore, artificial intelligence can have a democratizing effect on the accessibility of such expertise and prevent the rulers monopolizing or ignoring expertise.

1

This is a reformulation of the constitutional values identified by Daniel Halberstam, namely voice, rights and expertise. See Halberstam 2017.

7. Conclusion

How does artificial intelligence relate to the humanity (in the sense of respecting human dignity) of societal processes and orders, particularly the policies of national and transnational institutions? That has been the subject of this essay. It is important to assess these developing relationships as a moral question – provided it does not take on the semblance of ethical laundering of practices that are essentially continued without correction.¹ New working processes have been developed in the professions in which artificial intelligence is used, such as marketing, medicine and counterterrorism. The situation will be no different in public administration and government policy. A process of serious reflection is required on this matter to indicate how the use of artificial intelligence can be embedded in interpersonal relationships governed by the rule of law. This is necessary both to prevent a calamity and to optimize the constitutional embedding.²

Such an arrangement must ensure that an advance in expertise (and access to it) in government or stakeholder groups is not achieved at the expense of people and groups in a weaker, more vulnerable position, leading to further alienation of democracy³ and greater socioeconomic inequality. The potential of comprehensive, coherent expertise can be used to facilitate a democratic society and support the development of life projects. The public system must consolidate the beneficial use of AI by providing countervailing powers, such as supervisory authorities, and sufficiently implement existing countervailing powers, such as courts and complaint-handling bodies (because monitoring of databases and the operation of algorithms is highly demanding specialist work). Trusted intermediaries have played a role in ensuring the integrity of information gathering and assessments since time immemorial; here too changes are under way involving the use of advanced information technology, such as distributed trust using secure multipart computation. Cryptocurrencies are an example of their use – perhaps not the most socially desirable one – but the technology also lends itself to protection against data manipulation, for example.⁴

1 Bietti 2020; Yeung et al. 2020.

2 An initiative for such constitutional recalibration was developed in the previously cited preliminary advice for the Constitutional Law Conference 2020 of Goossens et al. 2021.

3 Schäfer and Zürn 2021, 120-127.

4 Goossens et al. 2020, 141-143.

It is by these means that trust in justice and the fulfilment of human rights must be achieved in the context of the 21st century – but the power that can be derived from the use of artificial intelligence may also produce the opposite outcome.

This essay thus leads to one burning question: *Is the arrival of artificial intelligence in the Anthropocene a portent of humanity being sucked into the black hole of a world made inhospitable by our own actions, or the means by which, amid the complexity of that world, people will give free rein to their inspiration, reposition themselves and achieve fragments of constitutionality?* To achieve the latter outcome we must not ignore artificial intelligence but must engage with it to create genuine ‘augmented intelligence’, in accordance with the principles in section 5.

We must not allow artificial intelligence to be abused but must use it to achieve good ends. That requires the government, civil society organizations and citizens to remain connected with each other in institutions and working processes in which artificial intelligence is used. Like every other facet of the policy, such use can be directed towards a dignified future, or not, or – as is alas so often the case – only to a limited extent, because we no longer care. We would be underestimating the impact of artificial intelligence if discussion about it were confined to rules intended to ‘frame’ its use: it is the working processes of the democratic constitutional state itself that need to be overhauled. Artificial intelligence must be seen as a facet of our social coexistence, with the constant moral and political challenge being to improve its dignity and sustainability.

12 July 2021

Bibliography

- Adriaansz, C. (2020) 'De rechtmatigheid van algoritmische besluitvorming in het licht van het zorgvuldigheidsbeginsel en het motiveringsbeginsel', *Nederlands Tijdschrift voor Bestuursrecht* 2020(100): 240-249.
- Agrawal, A., J. Gnas and A. Goldfarb (2018) *Prediction Machines: The Simple Economics of Artificial Intelligence*, Boston: Harvard Business Review Press.
- AIV (2019) *Duurzame ontwikkelingsdoelinden en mensenrechten. Een noodzakelijk verbond*, The Hague: Advisory Council on International Affairs.
- Arnauld A. von, K. von der Decken and M. Susi (2020) *The Cambridge Handbook of New Human Rights: Recognition, Novelty, Rhetoric*, Cambridge/New York: Cambridge University Press.
- Arun, C. (2019) 'AI and the Global South: Designing for Other Worlds', in M. Dubber et al. (eds.).
- Ashley, K.D. (2017) *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, Cambridge/New York: Cambridge University Press.
- Assmann, J. (2018) *Achsenzeit – Eine Archäologie der Moderne*, Munich: C.H. Beck.
- Bakker, S. and P. Korsten (2020) *Artificiële intelligentie als een general purpose technology: Strategische belangen en verantwoorde inzet in historisch perspectief*, WRR Working Paper no. 41, The Hague Netherlands Scientific Council for Government Policy.
- Biebricher, T. (2021) *Die politische Theorie des Neoliberalismus*, Berlin: Suhrkamp.
- Bietti, E. (2020) 'From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy', pp. 210-219 in *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. DOI: <https://doi.org/10.1145/3351095.3372860>.
- Blühdorn, I. and M. Deflorian (2019) 'The Collaborative Management of Sustained Unsustainability: On the Performance of Participatory Forms of Environmental Governance', *Sustainability* 11(4/1189): 17.
- Boden, M. (2018) *Artificial Intelligence: A Very Short Introduction*, Oxford: Oxford University Press.
- Brown, W. (2015) *Undoing the Demos: Neoliberalism's Stealth Revolution*, New York: Zone Books.
- Bues, M.-M. (2018) 'Artificial intelligence im Recht', in Hartung et al. (eds.).
- Busch, C. and A. de Franceschi (2018) 'Granular Legal Norms: Big Data and the Personalization of Private Law', pp. 408-424 in V. Mak, E. Tjon Tin Tai

- and A. Berlee (eds.) *Research Handbook in Data Science and Law*, Cheltenham/Northampton, MA: Edward Elgar Publishing.
- Butter, M. (2018) “*Nichts ist, wie es scheint*” – *Über Verschwörungstheorien*, Berlin: Suhrkamp.
- Chavan, M. and Y.A. Breyer (2020) ‘Supply Chain Management and Social Enterprise Towards Zero Hunger: The Akshaya Patra Foundation in India’, in Wood et al. (eds.).
- Davis, M.F. (2020) ‘Get Smart: Human Rights and Urban Intelligence’, *Fordham Urban Law Journal* 47: 971-991.
- Dubber, M.D., F. Pasquale and S. Das (eds.) (2020) *Oxford Handbook of Ethics of AI*, Oxford: Oxford University Press.
- Düwell, M., J. Braarvig, R. Brownsword, et al. (eds.) (2014) *The Cambridge Handbook of Human Dignity: Interdisciplinary Perspectives*, Cambridge: Cambridge University Press.
- Ellis, E.C. (2018) *Anthropocene. A Very Short Introduction*, Oxford: Oxford University Press.
- Esposito, R. (2014) *Le persone e le cose*, Turin: Einaudi.
- Esposito, R. (2015) *Persons and Things: From the Body’s Point of View*, Cambridge/Oxford/Boston/New York: Polity Books.
- Fineman, M.A. (2017) ‘Vulnerability and Inevitable Inequality’, *Oslo Law Review* 4: 133-149.
- Frankenberg, G. (2014) *Political Technology and the Erosion of the Rule of Law*, Cheltenham/Northampton, MA: Edward Elgar.
- Goos, M. and A. Salomons (2021) ‘Data op de Arbeidsmarkt: toepassingen en uitdagingen’, pp. 113-132 in: Hirsch Ballin et al. (eds.).
- Goossens, J., E.M.H. Hirsch Ballin and E.Y. van Vugt (2021) ‘Algoritmische beslisregels vanuit constitutioneel oogpunt – Tweedeling tussen algemene regels en concrete toepassing onder druk’, *Tijdschrift voor Constitutioneel Recht* 12: 4-19.
- Goossens, J., K. Verslype and E. Tjong Tjin Tai (2020) *Blockchain en smart contracts: Herijking van de rol van de vertrouwde tussenpersoon in de algoritmische samenleving*, The Hague: Sdu.
- Goralski, M.A. and T.K. Tan (2020) ‘Artificial intelligence and sustainable development’, *The International Journal of Management Education* 18(1): 100330.
- Halberstam, D. (2017) ‘The Promise of Comparative Administrative Law: A Constitutional Perspective on Independent Agency’, pp. 139-158 in S. Rose-Ackerman, P.L. Lindseth and B. Emerson (eds.) *Comparative Administrative Law*, 2nd ed., Cheltenham/Northampton, MA: Edward Elgar Publishing.
- Hartung, M., M.-M. Bues and G. Halbleib (2018) *Legal Tech: Die Digitalisierung des Rechtsmarkts*, Munich: C.H. Beck/Vahlen.

- Heeger, R. (2014) 'Dignity only for humans? A controversy', pp. 541-545 in Düwell et al. (eds.).
- Hermstrüwer, Y. (2020) 'Artificial Intelligence and Administrative Decisions Under Uncertainty', pp. 199-223 in Wischmeyer and Rademacher (eds.).
- Hirsch Ballin, E. (2020) *Advanced Introduction to Legal Research Methods*, Cheltenham/Northampton MA: Edward Elgar.
- Hirsch Ballin, E.M.H., A.P.C.M. Jaspers, J.A. Knottnerus and H. Vinke (eds.) (2021) *De toekomst van de sociale zekerheid: de menselijke maat in een solidaire samenleving*, The Hague: Boom juridisch.
- Hoffmann-Riem, W. (2020) 'Artificial Intelligence as a Challenge for Law and Regulation', pp. 1-29 in Wischmeyer and Rademacher (eds.).
- Hollak, J.H.A. (1972) 'Vrijheid en noodzakelijkheid', in J.H.A. Hollak (2010) *Denken als bestaan. Het werk van Jan Hollak*, Budel: Damon.
- Li, T., H. Tang, J. Zhu and J.H. Zhang (2019) 'The finer scale of consciousness: quantum theory', *Annals of translational medicine*, 7(20): 585. DOI: 10.21037/atm.2019.09.09.
- Logemann, A. (2004) *Grenzen der Menschenrechte in demokratischen Gesellschaften: Die "demokratische Gesellschaft" als Determinante der Grundrechtsschranken in der Europäischen Menschenrechtskonvention*, Baden-Baden: Nomos.
- Meijer, A., S. Grimmelikhuijsen and M. Bovens (2021) 'De legitimiteit van algoritmisch bestuur: een systematisch overzicht van bedreigingen en oplossingsrichtingen', *Nederlands Juristenblad* 96(18): 1471-1478.
- Moore, J. (2019) 'AI for Not Bad', *Frontiers in Big Data*, 11 September 2019. DOI: <https://doi.org/10.3389/fdata.2019.00032>.
- Müller, J.-W. (2001) *Ein gefährlicher Geist: Carl Schmitts Wirkung in Europa*, Darmstadt: Wissenschaftliche Buchgesellschaft.
- Oliveira, A. (2017) *The Digital Mind: How Science is Redefining Humanity*, Cambridge, MA/London: The MIT Press.
- Passchier, R. (2020) 'Digitalisering en de (dis)balans binnen de trias politica', *Ars Aequi*, October: 916-927.
- Peters, K., S. Silva, R. Gonçalves, M. Kavelj, H. Fleuren, D. den Hertog, O. Ergun, M. Freeman (2021) 'The Nutritious Supply Chain', *INFORMS Journal on Optimization* 3(2): 200-226.
- Pietrzykowski T, (2017) 'Towards modest naturalization of personhood in law', *Revus (Journal for Constitutional Theory and Philosophy of Law / Revija za ustavno teorijo in filozofijo prava)* 32: 59-71.
- Price, D.A. (2020) 'An AI Breaks the Writing Barrier – A new system is shocking experts with its ability to use and understand language as well as human beings do', *Wall Street Journal* 22 August 2020.
- Prins, C. (2018) 'Urgenda en digitalisering', *Nederlands Juristenblad* 93(22): 1563.

- Prins, C. (2021) 'Discriminerende algoritmes', *Nederlands Juristenblad* 96(20): 1637.
- Ranchordás, S. (2021) 'Connected but still excluded? Digital exclusion beyond Internet Access', in: M. Ienca, O. Pollicino, L. Liguori, et al. (eds.) *The Cambridge Handbook of Life Sciences, Informative Technology and Human Rights*, Cambridge: Cambridge University Press.
- Rende, A. (2020) 'Europe: Toward a Policy Framework for Trustworthy AI', in Dubber et al. (eds.).
- Rona, G. and L. Aarons (2016) 'State Responsibility to Respect, Protect and Fulfill Human Rights Obligations in Cyberspace', *J. Nat'l Security L. & Pol'y*. Accessed at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2859615.
- Rosa, H. (2018) *Unverfügbarkeit*, Vienna/Salzburg: Residenz Verlag.
- Salian, I. (2019) 'AI in the Sky Aids Feet on the Ground Spotting Human Rights Violations', *NVIDIA* 4 April 2019. Accessed at <https://blogs.nvidia.com/blog/2019/04/04/human-rights-watch-ai-gtc/>.
- Schaber, P. (2014) 'Dignity only for humans? On the dignity and inherent value of non-human beings', pp. 546-550 in Düwell et al. (eds.).
- Schäfer, A. and M. Zürn (2021) *Die demokratische Regression: Die politischen Ursachen des autoritären Populismus*, Berlin: Suhrkamp.
- Schindler, H.-J. (2020) 'Die Demokratisierung von Deepfakes: Wie technologische Entwicklung unseren gesellschaftlichen Konsens beeinflussen kann', *Auslandsinformationen* (1): 64-73.
- Shanahan, M. (2012) 'The brain's connective core and its role in animal cognition', *Philos Trans R Soc Lond B Biol Sci* 367(1603): 2704-2714.
- Simmons, W.P. (2011) *Human Rights Law and the Marginalized Other*, Cambridge/New York: Cambridge University Press.
- Straaten, J. van der (2019) *AlgorithmWatch. Identity management and citizen scoring in Ghana, Rwanda, Tunisia, Uganda, Zimbabwe and China*.
- Sweijts, T., F. Remmits, H. van Manen and F. Bekkers (2020) *De Veiligheidsimplicaties van de Pandemie: De Impact van Covid-19 op Europese Veiligheid*, The Hague: Hague Centre for Strategic Studies.
- Tasioulas, J. (2019) 'First Steps Towards an Ethics of Robots and Artificial Intelligence', *Journal of Practical Ethics* 7: 61-95.
- Tjeenk Willink, H. (2018) *Groter denken, kleiner doen: Een oproep*, Amsterdam: Prometheus.
- Toit, C.W. de (2016) 'Human uniqueness on the brink of a new axial age: From separation to reintegration of humans and nature', *HTS Theologische Studies/Theological Studies* 72: a3487.
- Vetzo, M.J., J.H. Gerards and R. Nehmelman (2018) *Algoritmes en grondrechten*, The Hague Boom juridisch.

- Vöneky, S. (2020) 'How Should We Regulate AI? Current Rules and Principles as Basis for "Responsible Artificial Intelligence"', pp. 279-310 in A. Bruns, Y. Bu, H. Merkt, S. Meier, M. Pawlik, E. Takahashi, S. Vöneky and J. von Hein (eds.) *Legal Theory and Interpretation in a Dynamic Society*, Baden Baden: Nomos Verlagsgesellschaft mbH & Co. KG.
- Wheeler, T., P. Verveer and G. Kimmelman (2020) *New Digital Realities; New Oversight Solutions*, Shorenstein Center Discussion Papers, Cambridge, MA: Harvard Kennedy School.
- Wischmeyer, T. and T. Rademacher (eds.) (2020) *Regulating Artificial Intelligence*, Cham: Springer.
- Wittner, F. (2019) 'A public database as a way towards more effective algorithm regulation and transparency?', pp. 175-192 in L. Reins (ed.) *Regulating New Technologies in Uncertain Times*, Information Technology and Law Series vol. 32, The Hague: T.M.C. Asser Press/Springer.
- Wolswinkel J. (2020) *Willekeur of algoritme? Laveren tussen analoog en digitaal bestuursrecht*, Tilburg: Tilburg University.
- Wolswinkel, J. (2020a) 'Maatwerk, digitalisering en bestuursrecht: onlosmakelijk verbonden? Bespreking van VAR-preadvies "Maatwerk bieden in een gedigitaliseerde en datagedreven samenleving; #HoeDan?"', *Nederlands Tijdschrift voor Bestuursrecht* 2020(137): 321-325.
- Wood, L., L.P. Tan, Y.A. Breyer and S. Hawse (eds.) (2020) *Industry and Higher Education: Case Studies for Sustainable Futures*, Singapore: Springer.
- WRR (2016) *Big data in een vrije en veilige samenleving*, Amsterdam: Amsterdam University Press.
- WRR (2018) *Europese variaties*, The Hague: Netherlands Scientific Council for Government Policy.
- Yeung, K., A. Howes, and G. Pogrebna (2020) 'AI Governance by Human Rights-Centered Design, Deliberation, and Oversight: An End to Ethics Washing', in Dubber et al. (eds.).

Human rights as benchmarks of artificial intelligence

WRR

e-ISBN 978-90-83201-20-7



9 789083 201207 >