ASIAN
HORIZONS
Vol. 18, No 3, September 2024
Pages: 375-389

ARTIFICIAL INTELLIGENCE AND THE COMPLEXITY OF ETHICS

Peter G. Kirchschlaeger♦
University of Lucerne, Switzerland

Abstract

While reflecting upon artificial intelligence, one of its characteristics is often highlighted: its complexity. Sometimes the complexity of artificial intelligence is even used as an argument against holding humans responsible for it. At the same time, surprisingly the complexity of ethics is usually perceived with a reductionist understanding of ethics. In this article, the concept "artificial intelligence" itself is critically reviewed resulting in the introduction of a more adequate term: "databased systems." Beyond that, I argue against the possibility of "ethical" databased systems and in favour of databased systems with ethics. Finally, the complexity of ethics and its consequences for the ethical dimension of technology-based innovation will be in the focus.

♦**Peter G. Kirchschlaeger** is Full Professor of Theological Ethics and Director of the Institute of Social Ethics (ISE) at the Faculty of Theology of the University of Lucerne. He is Research Fellow at the University of the Free State, Bloemfontein. He had his studies in Theology and Judaism in Lucerne, in Rome (Gregoriana), and Jerusalem (2001: MA University of Lucerne), and studies in Philosophy, Religious Studies, and Political Science at the University of Zurich (2003: MA University of Zurich). 2005-2006: Research Stay at the University of Chicago Divinity School with a scholarship of the Swiss National Science Foundation. 2008: PhD University of Zurich. 2012: Habilitation University of Fribourg. 2011-2015 member of the Board of the Swiss Centre of Expertise in Human Rights; 2013 Visiting Scholar at the University of Technology Sydney, 2013-2014; Guest-Professor at the Faculty of Theology and Religious Studies at the Katholieke Universiteit Leuven; 2013-2017 Fellow at the Raoul Wallenberg Institute of Human Rights and Humanitarian Law; 2015-2019 Guest-Lecturer at the Leuphana University Lueneburg; and 2015-2017 Visiting Fellow at Yale University. One of his research focuses is Ethics of Artificial Intelligence. Email: Peter.Kirchschlaeger@unilu.ch

Keywords: Artificial Intelligence; Complexity of Ethics; Conscience; Databased Systems; *Epikeia*; Ethical Principles and Norms; Freedom; Moral Technologies

1. Introduction

In the discussion about artificial intelligence—including the ethical discourse on the opportunities and risks of this technology-based innovation[1]—, the complexity of artificial intelligence is referenced. In this exchange, surprisingly the complexity of ethics is usually not mentioned. Instead, one can meet a reductionist understanding of ethics. In this article, I want to address this. After discussing the concept "artificial intelligence" and neglecting the possibility of an "ethical" artificial intelligence and proposing artificial intelligence with ethics, the complexity of ethics and its consequences for the ethical dimension of technology-based innovation will be analysed. This is of particular relevance for South-East Asia as it is a major player in the rise of artificial intelligence.

2. Artificial Intelligence? Databased Systems!

Of course, artificial intelligence possesses a technological complexity. At the same time, confronted with the question for the definition of "artificial intelligence," one becomes aware of its conceptual blurriness,[2] something which should be overcome from an ethical perspective.

> I don't quite know whether it is especially computer science or its sub-discipline Artificial Intelligence that has such an enormous affection for euphemism. We speak so spectacularly and so readily of computer systems that understand, see, decide, make judgments (…) without ourselves recognizing our own superficiality and immeasurable naiveté with respect to these concepts. And, in the process of so speaking, we anesthetize our ability to (…) become conscious of its end use (…) One can't escape this state without asking, again and again: 'What do I actually do?' 'What is the final application and use of the products of my work?' and ultimately: 'Am I content or ashamed to have contributed to this use?'[3]

Artificial intelligence can be defined as "machines that are able to 'think' in a human like manner and possess higher intellectual abilities and professional skills, including the capability of correcting

---

[1]I'm grateful to Aaron Butler for the valuable comments on this article.

[2]Ohly Lukas, *Ethik der Robotik und der Künstlichen Intelligenz*, Berlin: Lang, 2019, 20.

[3]Joseph Weizenbaum, "Not Without Us," *SIGCAS Computers and Society* 16, 2-3 (1986) 2-7.

themselves from their own mistakes"[4] or as "the science and engineering of machines with capabilities that are considered intelligent by the standard of human intelligence."[5] The term 'artificial' in "artificial intelligence" highlights that the "intelligence (is) displayed or simulated by technological means."[6]

From an ethical perspective, the above-mentioned starting-point is criticized: "Intelligence is not limited to solving a particular cognitive problem, it depends on *how* that happens."[7] In view of the nature of artificial intelligence, doubts arise from an ethical perspective as to whether the term is even adequate, because artificial intelligence strives to imitate human intelligence, but this is limited to a certain area of intelligence (e.g., certain cognitive capacities).[8] Furthermore, it is to be assumed that artificial intelligence can at best become similar to human intelligence in certain areas of intelligence, but can never become identical or equivalent.

Finally, criticizing "artificial intelligence" on a conceptual level is ethically pertinent, which can, for example, be demonstrated by the use of the term "trustworthy Artificial Intelligence":

> First of all, the (…) central idea of a 'trustworthy AI' is conceptual nonsense. Machines are not trustworthy, only humans can be trustworthy—or not. (…) The story of the trustworthy AI is a marketing narrative thought of by the industry, a good-night-story for customers of tomorrow. In reality, it is about developing future markets and using debates on ethics as elegant, public decoration for a large-scale investment strategy.[9]

The term "databased systems" would be more appropriate than "artificial intelligence," because this term describes what actually

---

[4]Spyros G. Tzafestas, Roboethics, *A Navigating Overview*, Cham: Springer, 2016, 25. See also Lukas Ohly, *Ethik der Robotik und der Künstlichen Intelligenz*, Berlin: Lang, 2019, 22-25.

[5]Philip Jansen, Stearns Broadhead, Rowena Rodrigues, David Wright, Philp Brey, Alice Fox and Ning Wang, *State-of-the-Art Review*, Draft of the D4.1 deliverable submitted to the European Commission on April 13, 2018. A report for the SIENNA Project, an EU H2020 research and innovation program under grant agreement no. 741716, 5.

[6]Mark Coeckelberg, *AI Ethics*, London: MIT Press, 2020, 203.

[7]Catrin Misselhorn, *Grundfragen der Maschinenethik*, Ditzingen: Reclam, 2018, 17.

[8]See Hubert L. Dreyfus, *What Computers Can't Do. The Limits of Artificial Intelligence*, New York: MIT Press, 1972, 29; Hubert L. Dreyfus and Stuart E. Dreyfus, *Mind Over Machine*, New York: The Free Press, 1986.

[9]Thomas Metzinger, "Nehmt der Industrie die Ethik weg! EU-Ethikrichtlinien für Künstliche Intelligenz," in *der Tagesspiegel*, April 8, 2019.

constitutes "artificial intelligence": generation, collection, and evaluation of data; databased perception (sensory, linguistic); databased predictions; databased decisions. In addition, the term "databased systems" also underscores the main strength and the main weakness of the present technological achievement in this field. The mastery of an enormous quantity of data represents the key asset of databased systems.

Pointing to its core characteristic, namely of being based on data and relying exclusively on data in all its processes, its own development, and its actions—more precisely its reactions to data—lifts the veil of the inappropriate attribution of the myth of "intelligence" covering substantial problems and challenges of databased systems. This allows more accurateness, adequacy, and precision in the critical reflection on databased systems. The untraceability, unpredictability, and unexplainability of the algorithmic processes resulting in databased evaluation, databased predictions and databased decisions ("black-box-problem"[10]), its wide vulnerability to systemic errors, its deep susceptibility for confusing causality with correlation (e.g., high consumption of ice-creams by children in a summer-month and high number of children car-accidents due to more mobility during vacation in the same summer-month correlate but there is not any causal relationship between the two statistics, meaning ice-cream-consumption does not cause car-accidents),[11] and its high probability of biased and discriminatory data leading to biased and discriminatory databased evaluations, predictions, and decisions comprise its major disadvantages.[12] "Algorithms are opinions embedded in codes. They are not objective."[13] They are not neutral. They serve specific goals and purposes.

---

[10]See Will Knight, "The Dark Secret at the Heart of AI," *MIT Technology Review*, April 11, 2017; Yavar Bathaee, "The Artificial Intelligence Black Box and the Failure of Intent and Causation," *Harvard Journal of Law & Technology* 31, 2 (Spring 2018) 889-938; Will Knight, "The Financial World Wants to Open AI's Black Boxes," *MIT Technology Review*, April 13, 2017; Davide Castelvecchi, "Can We Open the Black Box of AI?" *Nature* 538, 7623 (2016) 21-23.

[11]See Gudmund R. Iversen and Mary Gergen, *Statistics: The Conceptual Approach*, New York: Springer, 1997, 317-318.

[12]See UNESCO, *Steering AI and Advanced ICTs for Knowledge Societies. A Rights, Openness, Access, and Multi-stakeholder Perspective*, Paris: UNESCO, 2019, 61-66; Mark Coeckelberg, *AI Ethics*, London: MIT Press, 2020, 125-136.

[13]Cathy O'Neil cited from: Yves Demuth, "Die unheimliche Macht der Algorithmen," *Beobachter* 9 (2018) 14-18, 16.

3. Databased Systems with Ethics[14]

3.1 Ethical Databased Systems?

In view of the complexity of databased systems and of the complexity of ethical questions related to databased systems, it would be tempting from an ethical perspective for humans to delegate ethical responsibility to artificial intelligence and to trust in "moral technologies" and in "ethical" databased systems.[15] Beyond that, "because of their increased intelligence, autonomy, and interaction capabilities, AI systems are increasingly perceived and expected to behave as moral agents."[16] Terms such as "moral technologies" used in the current discourse on digital transformation and on the use of databased systems suggest such an option. They express the expectation that it would be possible to create "moral technologies." Moral capability is understood as one of the areas of human intelligence which databased systems could achieve. The main objective would be to prevent machines from harming people, for example. The term "moral technologies" is based on their abilities to follow ethical rules, take moral decisions, and perform acts based on these. The term "moral technologies" is also based on this and its desire to give machines ethical principles and norms. This would not be achieved through programming, but rather through learning.

Talking about "moral technologies" may be considered to be rationally questionable. Can technological inventions really be moral? Can databased systems be ethical? Can technological systems be trusted to have morals? Can one ascribe moral agency to them?[17] Or do they possess a limited but not full morality like a "functional morality"[18] allowing them to assess the ethical consequences of their actions, or a mindless morality without obtaining the characteristics

[14]See Peter G. Kirchschlaeger, "Die Rede von 'moral technologies': Eine Kritik aus theologisch-ethischer Sicht," in *feinschwarz.net* (2017).

[15]See Michael Anderson and Susan Anderson, "General Introduction," in Michael Anderson and Susan Anderson, ed., *Machine Ethics*, Cambridge: Cambridge University Press, 2011, 1-4.

[16]Virginia Dignum, *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*, Cham: Springer, 2019, 36.

[17]See John Sullins, "When Is a Robot a Moral Agent," *International Review of Information Ethics* 6 (2006) 23-30.

[18]Wendell Wallach and Colin Allen, *Moral Machines: Teaching Robots Right from Wrong*, Oxford: Oxford University Press, 2009, 39.

humans possess as a basis for their morality?[19] Or is it impossible to think of technologies with morality? In the following, the characterization as "moral technologies" possessing moral capabilities will be critically examined from an ethical standpoint.

3.2. Conscience[20]

A first challenge to "ethical" databased systems is based on the concept of conscience, which is central for humans and their morality. Conscience unites what is objectively required and what has been subjectively experienced in a specific and concrete situation, in a specific context, during a unique encounter with unique people. "Conscience is an active faculty that discovers and discerns the good within the complexity of each situation."[21] Conscience creates an authority within a person, which has an impact on an action a priori, but also a posteriori. Conscience does not act itself.[22] Conscience expresses a trust in the human individual. The human individual is expected to have this inner voice in moral questions, to be able to recognize it, listen to it, and then act responsibly. It is respected and upheld that the dignity of conscience belongs to the human individual.

It cannot be said that technologies have a conscience. Because the potentials that technologies possess in relation to ethical decisions and actions are nowhere close to the human conscience. They lack various levels of morality or duty, as well as the existence merged in the conscience in varying qualities, intensities and marked by individual development or social influence.[23] If conscience is understood as being essential for morality, the lack of conscience is a first argument against "ethical" databased systems.

---

[19]See Luciano Floridi and J.W. Sanders, "On the Morality of Artificial Agents," *Minds and Machines* 14, 3 (2004) 349-379; Moor, James H., "The Nature, Importance, and Difficulty of Machine Ethics," in *IEEE Intelligent Systems* 21, 4, (2006) 18–21.

[20]See Peter G. Kirchschlaeger "Gewissen aus moraltheologischer Sicht," *Zeitschrift für katholische Theologie* 139 (2017) 152–177.

[21]Linda Hogan, "Conscience in the Documents of Vatican II," in Charles E. Curran, ed., *Conscience: Readings in Moral Theology No 14,* New York: Paulist Press, 2004, 82-88, 86-87.

[22]See also Werner Wolbert, *Gewissen und Verantwortung. Studien zur Theologischen Ethik,* Freiburg: Academic Press Fribourg, 2008, 170.

[23] See Hanspeter Schmitt, *Sozialität und Gewissen. Anthropologische und theologisch-ethische Sondierung der klassischen Gewissenslehre*. Studien der Moraltheologie Bd. 40, Wien: Lit, 2008.

3.3. Freedom

A second question mark regarding "ethical" databased systems arises from freedom. Freedom is a *conditio sine qua non* for morality, because only freedom opens up the possibility to decide for or against the good or the right, respectively. Freedom means to be able to act according to one's own interests, preferences, wishes, and plans. It can encompass the freedom to want what one wants and the freedom to want what one doesn't want. The latter means that freedom can also mean to want the "required," i.e. the ethically demanded, even if this might not correspond to one's own wishes, needs, preferences, or interests. This opens up the social horizon of freedom.

Beyond that, freedom is the origin of science, research, and technology. This aspect must be emphasized in a time when some voices deny the existence of freedom at all.[24]

> That freedom, which is now denied, has made the developments of science, in the name of which it is now denied, possible. Indeed, there would never have been a science without the human mind's inherent ability to distinguish between false and true and to prefer the true to the false. False and true make no sense if not for a free mind capable of striving for one and rejecting the other. Without these essential prerequisites, any explanation remains merely a vociferous, meaningless act. For this reason it can be said with justification that science is the most glorious monument that freedom has erected for itself, and that scientific research is completely unthinkable without freedom.[25]

Machines lack freedom. Technologies are designed, developed and built by humans, meaning they are produced heteronomously. Therefore, the learning of ethical principles and norms are also guided

---

[24]See for this discourse Adrian Holderegger, Beat Sitter-Liver, Christian W. Hess, Günter Rager, ed., *Hirnforschung und Menschenbild. Beiträge zur interdisziplinären Verständigung*, Fribourg: Academic Press Fribourg & Basel: Schwabe, 2007; Helmut Fink and Rainer Rosenzweig, ed., *Freier Wille – frommer Wunsch? Gehirn und Willensfreiheit*, Paderborn: Mentis, 2006; Margot Fleischer, *Menschliche Freiheit – ein vielfältiges Phänomen. Perspektiven von Aristoteles, Augustin, Kant, Fichte, Sartre und Jonas*, Freiburg i. B.: Verlag Karl Alber, 2012; Walter Bloch, "Willensfreiheit? Neue Argumente in einem alten Streit. Hodos – Wege bildungsbezogener Ethikforschung," in *Philosophie und Theologie*, 11, Frankfurt a. M.: Peter Lang, 2011; Emmanuel J. Bauer, ed., *Freiheit in philosophischer, neurowissenschaftlicher und psychotherapeutischer Perspektive*, München: Wilhelm Fink Verlag, 2007; Wolfgang Achtner, *Willensfreiheit in Theologie und Naturwissenschaften. Ein historisch-systematischer Wegweiser,* Darmstadt: WBG, 2010; Barbara Guckes, *Ist Freiheit eine Illusion? Eine metaphysische Untersuchung*, Paderborn: Mentis, 2003.

[25]Jeanne Hersch, *Im Schnittpunkt der Zeit*, Zurich: Benziger, 1992, 60-61.

by humans. In a last consequence, machines would always be controlled from the outside. Metaphorically speaking, machines—even self-learning machines—will go back to a first line of code that always comes from humans. The lack of freedom is a second argument against the moral capabilities of databased systems.

3.4. Responsibility

The freedom to want what one does not want makes responsibility stand out.[26] Responsibility succeeds in understanding that one's own freedom is connected to the freedom of all other humans and to respect the human dignity of all humans. Responsibility enables freedom to go beyond one's own needs and interests to discover the horizon for the freedom of all other humans and for social tasks and objectives. "Responsibility opens up a freedom that is individualistic and concentrated on one's own needs and integrates it in a social framework, common tasks and objectives."[27] Responsibility is also a *conditio sine qua non* for morality. In order to carry or be given responsibility—i.e. to be able to be a subject of responsibility—freedom and rationality are necessary.

The query arises whether machines can assume responsibility. The answer would have to be negative, because machines cannot be a subject of responsibility due to the fact that they lack freedom—a third argument against "ethical" databased systems.

3.5. Autonomy

A fourth fundamental query concerning "ethical" databased systems arises from the autonomy proclaimed by humans. Humans are carriers of dignity, and therefore, may not be instrumentalized, because they, as rational beings, recognize common moral rules and principles for themselves, determine them for themselves and base their actions on them.[28] Human dignity is based on the ability of the human to set rules of reason for himself/herself. This means that moral rules and principles that the human being formulates in his/her autonomy must meet the following requirements of a critical, rational morality, which will guarantee their general applicability:

---

[26]See Peter G. Kirchschlaeger, "Verantwortung aus christlich-sozialethischer Perspektive," *ETHICA* 22 (2014) 29-54.

[27]Adrian Holderegger, "Art. Verantwortung," in Jean-Pierre Wils and Christoph Hübenthal, ed., *Lexikon der Ethik*, Paderborn 2006, 394-403, 401.

[28]See Immanuel Kant, *Grundlegung zur Metaphysik der Sitten*, Werkausgabe Weischedel, Vol. 7. Frankfurt a. M.: Suhrkamp, 1974, 74.

> A rational or critical morality is one that demands rational justifiability for its principles. Moral principles are rationally justified, if they are generally acceptable, i.e. acceptable for all concerned persons, under the precondition of their complete equality and ability for self-determination.[29]

Does the description of human autonomy, which can be expressed by humans, correspond to the potential of technologies to follow moral rules, to make moral decisions accordingly, and to carry out corresponding actions? There is a gap between technologies and ethics regarding the notion of "autonomy."[30] While humans recognize general moral rules and principles for themselves, set them for themselves and base their actions on them, this is not possible for technologies. Technologies are primarily made for their suitability and may set rules as a self-learning system, for example to increase their efficiency. But these rules do not contain any ethical qualities. Machines fail on the above-mentioned principle of generalizability. This negation is a fourth argument against "ethical" databased systems.

3.6. Databased Systems with Ethics

The above critique leads to the main consequence that humans are responsible for making ethical decisions,[31] for laying down ethical principles, ethical and legal norms, setting a framework, goals and limits of databased systems, as well as defining their use besides examining, analysing, evaluating, and assessing technology-based innovation from an ethical perspective. It is up to humans to program and train databased systems with ethical principles and norms—even though technologies do not recognize the ethical quality of them.

Databased systems might cause good and evil but they do it without recognizing, knowing about or being aware of the ethical quality of it. The exclusive ability of humans to recognize the ethically legitimate, to decide accordingly, and to act upon it comprises also the possibility to decide not even to create, design, produce, disseminate, or to use

---

[29]Peter Koller, "Die Begründung von Rechten," in Peter Koller, Csaba Varga and Ota Weinberger, ed., *Theoretische Grundlagen der Rechtspolitik. Ungarisch-Österreichisches Symposium der internationalen Vereinigung für Rechts- und Sozialphilosophie*, ARSP 54, Stuttgart: Franz Steiner Verlag, 1990, 74-84, 75.

[30]See Ko Insok, "Can Artificial Intelligence Be an Autonomous Entity?" *Korean Journal of Philosophy* 133 (2017) 163-187.

[31]See Deborah Johnson, "Computer Systems: Moral Entities but not Moral Agents," *Ethics and Information Technology* 8, 4 (2006) 195-204; Roman V. Yampolski, "Artificial Intelligence Safety Engineering: Why Machine Ethics Is a Wrong Approach," in Vincent C. Mueller, ed., *Philosophy and Theory of Artificial Intelligence*, Cham: Springer, 2013, 289-296.

technologies as objects correspondingly the possibility to decide to abolish or to destroy technologies.

4. Ethics is Not Democracy

While living up to this ethical responsibility, humans are confronted with the complexity of ethics. This complexity of ethics is a final argument why this responsibility must remain in the hands of humans instead of them being able to trust it to databased systems.

In the first, the complexity of ethics consists in something that remains a fundamental conceptual challenge (e.g., for ethics committees),[32] namely: ethics as a science is not democratic. A democratic process *per se* does not guarantee legitimacy. It is conceivable that a democratic opinion-forming and decision-making process may also lead to results that are morally wrong. Ethics in a rational and critical way needs to satisfy the following requirements:

> A rational or critical morality is one that claims for itself rational justifiability for its principles. Moral principles are rationally justified if they are generally endorsed by, that is to say acceptable to, all affected persons, given their full equality and effective self-determination.[33]

5. Ethics Beyond Principles and Norms

Databased systems are able to follow moral rules and make moral decisions based on these and act accordingly. Databased systems can be programmed or trained with ethical rules in order to come to ethically legitimate decisions and perform ethically legitimate actions as a machine.[34] In order to do justice to the complexity of reality, it takes much more than rules like "Asimov's Law:"

> 1. A robot may not injure a human being or, through inaction, allow a human being to come to harm. 2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law. 3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.[35]

---

[32]See Claude Huriet, "Ethics Committees," in Henk A.M.J. ten Have and Michele S. Jean, ed., *The UNESCO Universal Declaration on Bioethics and Human Rights. Background, Principles and Application*, Paris: UNESCO, 2009, 265-270.

[33]Peter Koller, "Die Begründung von Rechten," in Peter Koller, Csaba Varga and Ota Weinberger, ed., *Theoretische Grundlagen der Rechtspolitik. Ungarisch-Österreichisches Symposium der internationalen Vereinigung für Rechts- und Sozialphilosophie*, ARSP 54, Stuttgart: Franz Steiner Verlag, 1990, 74, 84, 75.

[34]See on this Catrin Misselhorn, *Grundfragen der Maschinenethik*, Ditzingen: Reclam, 2018, 70-135.

[35]Isaac Asimov, *Meine Freunde, die Roboter*, München: Heyne Verlag, 1982, 67.

Isaac Asimow was probably aware of this himself, which he expressed, among other things, by choosing the genre of a short story to publish these rules.

One could think that databased systems could just simulate humans in the domain of ethics. "If the AI can understand human morality, it is hard to see what is the technical difficulty in getting it to follow that morality."[36] The first counter-argumentation line against this position embraces that meaning is created by humans.[37] Finding a purpose for one's own existence and getting an idea of a "good life" and a "right life" is possible for humans and is limited to humans because humans develop moral capabilities that machines are lacking for the above-mentioned reasons.

The second counter-argumentation line embraces the "value alignment problem" or the "value-loading-problem"[38] identified by Stuart Russell[39] emphasizing the complex contexts with which databased systems interact. If databased systems were created with principles, norms, and values, the challenge raises that these principles, norms, and values must be and remain aligned with the principles, norms, and values of humans.

The third counter-argumentation line against this reductionist view of human morality[40] as something that could just be translated into the langue of mathematics and simulated by databased systems consists in the necessary interaction between ethics and technology embracing reciprocally enriching exchange. E.g., technology-based innovation does not just provoke a normative reaction but possesses itself already an ethical dimension that is influencing the decision to launch an innovation-process. Reciprocally, technological progress has an impact

---

[36]Ernest Davis, *Ethical Guidelines for A Superintelligence*, New York: New York University Press, 2014, 1-5, 3; see also Nicholas Agar, "Don't Worry about Superintelligence," *Journal of Evolution and Technology* 26, 1 (February 2016) 73-82; Nate Soares and Benya Fallenstein, *Agent Foundations for Aligning Machine Intelligence with Human Interests: A Technical Research Agenda*, Machine Intelligence Research Institute Technical Report 2014/8, 1-14.

[37]See John R. Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3, 3 (1980) 417-457; Margaret A. Boden, *AI: Its Nature and Future*, Oxford: Oxford University Press, 2016.

[38]See Nik Bostrom, "The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents," *Minds and Machines* 22, 2 (2012) 71-85.

[39]See Stuart Russell, "Will They Make Us Better People?" contribution to the Annual Question 2015 on edge.org, http://www.edge.org/response-detail/26157.

[40]See also Mark Graves "Shared Moral and Spiritual Development Among Human Persons and Artificially Intelligent Agents," *Theology and Science* 15, 3 (2017) 333-351.

on ethics as well, e.g., leading to certain ethical questions due to the existence of technological possibilities and realities.

The fourth counter-argumentation line against this reductionist view of ethics honours that ethics goes beyond principles, norms, and rules. "With a computer we can turn almost all human problems into statistics, graphs, equations. The really disturbing thing, though, is that in doing so we create the illusion that these problems can be solved with computers."[41] In order to do justice to the complexity of ethics, mathematical or digital ethics need not be used. It is important to note that ethics in its complexity and in its entirety is not translatable into the language of mathematics and programming,[42] because of its sensitivity to the rule-transcending uniqueness of the concrete. That is why, among other things, ethics is not casuistry. Certain aspects of ethics can be programmed or trained as rules for databased systems. Some ethical elements, though, cannot be reached by digital instruments. For example, databased systems can learn the principle of human dignity for all, human rights, and ethical guiding principles (prohibition of lying, stealing, etc.). However, even within the realm of possibilities, it should not be neglected that databased systems can learn and follow these rules, yet they follow the rules without knowing about the *ethical* quality of those rules. In other words, databased systems would respect non-ethical or unethical rules in the same manner.

> I do not think that they will end up with a moral or ethical robot. For that, we need to have moral agency. For that, we need to understand others and know what it means to suffer. The robot may be installed with some rules of ethics but it won't really care. It will follow a human designer's idea of ethics.[43]

Databased systems cannot pass these limitations. "AI will not share these human traits unless we specifically create them to do so. They operate on a task and goal-oriented manner."[44] These limitations are part of databased systems because they rely exclusively on data

---

[41]Naief Yehya, *Homo cyborg. Il corpo postumano tra realtà e fantascienza*, Milano, 2005, 15.

[42]See Michał Klincewicz, "Challenges to Engineering Moral Reasoners," in Patrick Lin, Ryan Jenkins and Keith Abney, ed., *Robot Ethics 2.0 from Autonomous Cars to Artificial Intelligence*, New York: Oxford University Press, 2020, 244-257.

[43]Noel Sharkey, cited from Patrick Tucker, "Can the Military Really Teach Robots Right from Wrong?" *The Atlantic*, May 14, 2014.

[44]Gonenc Gurkaynak, Ilay Yilmaz and Gunes Haksever, "Stifling Artificial Intelligence: Human Perils," *Computer Law & Security Review* 32, 5 (2016) 749-758, 756.

without a theory; they accept a solution without addressing the question "why" and while neglecting the search for reasons. This also applies to ethical rules.[45] This causes the problematic consequence, from an ethical point of view, that it can be necessary to convey to databased systems ethical values, principles, and norms, which one thinks are false just because they are not able to handle the right ones (e.g., consequentialist instead of deontological approach because consequentialist values, principles, and norms are simpler to be translated into the language of mathematics).

Transferring ethics to mathematics or programming becomes difficult or even impossible when guiding principles diverge or collide. Through the increasing complexity of everyday reality, humans are challenged to find insights into norms that are adequate to reality, and to consider in a more differentiated and better manner what would be expecting too much of databased systems due to their lack of moral capability. In situations and cases where in humans the virtue of *epikeia* and conscience come into play, translating ethics into the language of mathematics, programming, and digitalization is impossible. "Epikeia is the rectification of the law where there are gaps due to its general formulation."[46] *Epikeia* means that "an independent practical judgement records the moral demands of a concrete situation in the light of moral principles and standards."[47] *Epikeia* consists in "the search for greater justice,"[48] it is "to stimulate and to maintain the search for the justice of meaning."[49] *Epikeia* accounts for the truth that in a concrete encounter with concrete persons in a concrete situation rules reach their limit, because the concrete in its uniqueness outranks the rule. "The general, concrete ethical, the positive legal and many other norms that are generally applicable, although indispensable, are not sufficient to guarantee the basic humanity which, in the face of diversity, will save this society from disintegration and the terrible

---

[45]See Brian Talbot, Ryan Jenkins and Duncan Purves, "When Robots Should Do the Wrong Thing," in Patrick Lin, Ryan Jenkins and Keith Abney, ed., *Robot Ethics 2.0 from Autonomous Cars to Artificial Intelligence*, New York: Oxford University Press, 2020, 258-273.

[46]Aristoteles, *Nikomachische Ethik* V, Ditzingen: Reclam, 14.

[47]Eberhard Schockenhoff, *Grundlegung der Ethik. Ein theologischer Entwurf*, Freiburg i. B.: Herder, 2014, 601.

[48]Kerstin Schlögl-Flierl, "Die Tugend der Epikie im Spannungsfeld von Recht und Ethik," in Paul-Chummar Chittilappilly, ed., *Horizonte gegenwärtiger Ethik*. Freiburg i. B.: FS Josef Schuster, 2016, 29-39, 29.

[49]Kerstin Schlögl-Flierl, "Die Tugend der Epikie im Spannungsfeld von Recht und Ethik," in Paul-Chummar Chittilappilly, ed., *Horizonte gegenwärtiger Ethik*. Freiburg i. B.: FS Josef Schuster, 2016, 29-39, 29-30.

consequences which result from it. It is inevitable that we have to cross norms in certain situations in order to act humanely, but this does not mean that we deny the need for norms in general or refute that they are generally applicable."[50] Ethical and legal norms and their validity are of course not questioned by *epikeia*. *Epikeia* "not only directs one to apply norms, but to recognize the more urgent ones."[51] They are reconfirmed by this virtue striving for justice. At the same time, *epikeia* ensures that the ethical and legal norms serve humans and not vice-versa.[52] "With the help of *epikeia*, it is possible to act in a way that is appropriate to the situation and useful to people."[53] *Epikeia* requires, though, ethically critical and constructive participation[54] by "a human as a responsible person who is able to consider and interpret standards and laws creatively."[55]

In this context, humans are expected to take responsibility for designing norms, something that is unattainable for databased systems because they lack moral capability. This responsibility for designing norms aims at continuously having to critically question these rules, and in the service of a prospective, ethical improvement, they are adapted by humans.

This prospective, creative level also contains a human responsibility to create standards. "Perceiving the moral claim does not mean to merely read normatively defined factual and meaningful behavior, but is always a creative process of seeing and discovering. The process of seeing and discovering becomes creative, because the human is called upon to risk in his phantasy new meaningful moments for their lifestyle, which did not occur in the previous system of rules. The moral goodness of the person urges him to develop the correct thing, from a human perspective, in the form of models."[56] The responsibility to create standards goes far beyond what can be translated into the

---

[50]Günter Virt, *Damit Menschsein Zukunft hat*, Würzburg: Echter Verlag, 2007, 42-43.

[51]James F. Keenan, *A History of Catholic Moral Theology in the Twentieth Century. From Confessing Sins to Liberating Conscience*, New York: Continuum, 2010, 155.

[52]See Kerstin Schlögl-Flierl, "Die Tugend der Epikie im Spannungsfeld von Recht und Ethik," in Paul-Chummar Chittilappilly, ed., *Horizonte gegenwärtiger Ethik*. Freiburg i. B.: FS Josef Schuster, 2016, 29-39, 39.

[53]Kerstin Schlögl-Flierl, "Die Tugend der Epikie im Spannungsfeld von Recht und Ethik," 29-39, 39.

[54]See Klaus Demmer, Bedrängte Freiheit. Die Lehre von der Mitwirkung—neu bedacht, SThE 127, Fribourg: Academic Press Fribourg, 2010, 110-113.

[55]Kerstin Schlögl-Flierl, "Die Tugend der Epikie im Spannungsfeld von Recht und Ethik," 29-39, 39.

[56]Günter Virt, *Damit Menschsein Zukunft hat*, Würzburg: Echter Verlag, 2007, 43.

language of mathematics or programming and, therefore, cannot be transferred to databased systems.

A transferability to mathematics and programming also excludes ethics of virtue with its focus on character traits and attitudes, because the moral capability of human cannot be digitalized. Based on his paradox, Hans Moravec would probably not be surprised that the complexity of ethics leads to unreachable areas for databased systems. Moravec's Paradox can be summarized: "The hard problems are easy, and the easy problems are hard."[57] The following example can illustrate its core meaning:

> Thinking several moves ahead in a game of chess is difficult for a human. It has been unexpectedly easy to program computers to do this. Chess computers now beat the best human players. Practical tasks, especially those connected with sensorimotor abilities, the kinds of tasks that humans perform effortlessly, have proved very challenging.[58]

6. Ethical Responsibility of Humans for Databased Systems

As ethics is not easy for humans, but they possess moral capability and therefore are able to meet ethical challenges, an adapted version of Moravec's Paradox is able to highlight the main point regarding the complexity of ethics: The unresolvable problems are resolvable for databased systems, and the resolvable problems (ethics) are unresolvable for databased systems. Humans, therefore, have and continue to have a responsibility not to lose oneself in the illusion of *ethical databased systems* but to create, design, produce, and use *databased systems with ethics*.[59]

---

[57]Steven Pinker, *The Language Instinct: How the Mind Creates Language*, New York: Collins Publishers, 2007, 190.

[58]Nicholas Agar, "Don't Worry about Superintelligence," *Journal of Evolution and Technology* 26, 1 (February 2016) 73-82, 76.

[59] For a detailed discussion on the topic, see Peter G. Kirchschlaeger, *Digital Transformation and Ethics: Ethical Considerations on the Robotization and Automatization of Society and Economy and the Use of Artificial Intelligence,* Baden-Baden: Nomos-Verlag, 2021.