Sławomir Lewandowski
Akademia im. Aleksandra Gieysztora w Pułtusku
(filia Akademii Finansów i Biznesu Vistula)
ORCID: 0009-0006-5866-0257

doi.org/10.34765/sp.0223.a02

A CRITICAL ANALYSIS OF THE RELATIONSHIP BETWEEN HUMANS AND "THINKING MACHINES". POSSIBLE RESEARCH QUESTIONS IN THE CONTEXT OF INSTITUTIONAL ETHICAL INTERVENTIONS USING THE EXAMPLE OF THE EUROPEAN UNION'S "ETHICS OF CONNECTED AND AUTOMATED VEHICLES"[1]

Summary

The article proposes an analysis of ethical interventions made by European Union bodies in the form of ethical recommendations addressed to artificial intelligence (AI) developers – using the example of 'Ethics of connected and automated vehicles' issued by the European Commission. AI is a hitherto undefined precisely phenomenon involving automated data acquisition and processing. For the purposes of this article, AI isdefined as a technology that enables machines to „learn from experience, adapt to new information and perform human-like tasks" (www16).

[1] European Commission, Directorate-General for Research and Innovation, Ethics of connected and automated vehicles: recommendations on road safety, privacy, fairness, explainability and responsibility, Publications Office, 2020. The document was issued in the form of ‚recommendations', a specific regulation provided for by the Treaties, the formulation of which is discussed later in this article.

The proposed analysis would be semantic (examining the basic concepts used to justify interventions) and teleological (examining the intentions behind interventions). The recommendations addressed to AV developers and users, given the clearly defined scope included in the twenty theses, allow for a coherent analysis of the EU institutions attitude towards technology in its early stages of development.

Keywords: artificial intelligence, autonomous vehicle, ethics.
JEL code: O38

KRYTYCZNA ANALIZA RELACJI CZŁOWIEKA Z „MASZYNĄ MYŚLĄCĄ". MOŻLIWE PYTANIA BADAWCZE W KONTEKŚCIE INSTYTUCJONALNYCH INTERWENCJI ETYCZNYCH Z WYKORZYSTANIEM PRZYKŁADU *ETYKI POJAZDÓW POŁĄCZONYCH I ZAUTOMATYZOWANYCH* UNII EUROPEJSKIEJ

Streszczenie

Artykuł zawiera propozycję przeprowadzenia analizy interwencji etycznych podejmowanych przez organy Unii Europejskiej w formie rekomendacji etycznych kierowanych do twórców sztucznej inteligencji (SI) – na przykładzie dokumentu *Etyka pojazdów zautomatyzowanych i połączonych* wydanego przez Komisję Europejską. SI to dotąd niezdefiniowane precyzyjnie zjawisko obejmujące zautomatyzowane pozyskiwanie i przetwarzanie danych. Na potrzeby artykułu przyjęto jej definicję jako technologii umożliwiającej maszynom „uczenie się na podstawie doświadczeń, dostosowywanie się do nowych informacji i wykonywanie zadań podobnych do ludzkich" (www16).

Proponowana analiza miałaby charakter semantyczny (badanie podstawowych pojęć używanych do uzasadniania interwencji) oraz teleologiczny (badanie intencji stojących za interwencjami). Rekomendacje skierowane do twórców i użytkowników pojazdów autonomicznych (PA), biorąc pod uwagę jasno określony zakres zawarty w dwudziestu tezach, pozwalają na spójną analizę stosunku instytucji unijnych do technologii we wczesnych fazach jej rozwoju.

Słowa kluczowe: sztuczna inteligencja, pojazd autonomiczny, etyka.
Kody JEL: O38

„In the face of the instrumentalisation that prevails all around, which does not spare 'professional' thinking either, only a philosophising that is freed from the burden of utility and cannot be reduced to abstract formulas makes it possible to penetrate reality mentally".

(Kloc-Konkołowicz 2000)

Introduction – self-driving vehicles

Autonomous vehicle (AV) is the colloquial term for cars[2], that, thanks to AI, under certain circumstances, are able to move safely without driver supervision[3]. The most common rationale for developing AV is to make traffic safer and more efficient. It is widely accepted that more than 90 per cent of road accidents are caused by human error, so eliminating them will significantly improve road safety (COM(2018)283 final) (www25). At the same time, while the use of unmanned aircraft, not only for military tasks, has now become widespread and information about the use of autopilots during air travel no longer arouses much excitement, solutions aimed at ultimately relieving drivers of the driving task are still at a relatively early stage of development.

Automated driving systems (ADS)[4], i.e. systems that allow part or all of the planning and execution of journeys to be AI enables unsupervised driving delegated to machines, are being tested in many countries, but are only allowed for limited use (Threlfall 2020). Constrains come not only from lawmakers, algorithm developers, or OEMs marketing strategies. To take responsibility for navigation through the maze of vehicles moving on the complex road infrastructure, which is additionally used by a multitude of vulnerable road users, tools are needed that are even more precise than in the case of planes or drones. It takes time to develop ADSs, train them on pre-prepared data sets and have the prototypes drive millions of kilometres on test tracks and in real road conditions. Currently, by processing the data needed to control the vehicle and plan the route, AI enables unsupervised driving mainly in a pre-planned or fully

[2] In fact, it is not only 'cars', but motor vehicles in general, so also slow-moving vehicles, robots, special-purpose vehicles, etc.

[3] Regulation 2019/2144 art. 3(21)(22).

[4] ADS: Automated Driving Systems – a driving system made up of hardware and software to perform the various driving functions of a vehicle without the driver's intervention.

predictable environment. According to some experts, the full implementation of AV into traffic may not occur until around 30 to 40 years (Litman 2023). Optimistic predictions are 10 to 20 years. Be that as it may, it is a long way from the time when AV will become an accepted and sufficiently tame part of the road landscape. The development of the brain of automation – artificial intelligence – also remains a mystery: "much of the toxic AI facets […] are not yet readily apparent and have not yet garnered widespread public attention" (www1). It is also important to remember that AV is by no means, as announced even in the professional press, "a smartphone on wheels". It remains a speeding mass being a subject to the same physical laws as other vehicles.

The scale of the challenge facing developers and decision-makers is therefore really large. From a formal and technical point of view, its axis is responsibility for traffic incidents (dilemma: man or machine). However, in order to give the dreams of really autonomous vehicles come true, careful consideration must be given to equipping them with the ability to anticipate and react to events, including the sudden reactions of other road users in critical situations[5]. This leads directly to questions about the possible decision-making autonomy of AI-based machines and the potential for developing the latter.

Artificial iintelligence as a partner for humans

With today's state of the art, it is easy to fall into extremes while assessing the AI potential[6]. This is particularly true of its ability to make decisions in the field of human well-being. Alongside enthusiasm for its almost limitless data-processing capabilities, one hears voices full of sceptical doubt about any possibility for algorithms to act autonomously, especially when fed with unverified internet information. Indeed, at the current stage of development of AI-based technologies, its true potential is *not yet* known.

---

[5] In particular, it is about ensuring safety (the possibility of a multi-variant AV response) under conditions of mixed traffic, i.e. the co-participation of autonomous vehicles and human-driven ‚analogue' vehicles, often reacting spontaneously to traffic situations.

[6] AI is an as yet undefined precisely phenomenon involving automated data acquisition and processing. One proposal to deal with this problem is to try to define an „artificial intelligence system" – i.e. software developed using specific techniques that can (for given purposes defined by humans) generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with; – cf. Proposal for a Regulation...COM(2021)206.

Rather, attempts to regulate the creation and use of the products of modern technology are accompanied by a concern not to miss the point at which AI begins to bear socially problematic fruit. Sean O hEigearaigh, for example, argues that the fundamental difficulty will be to ensure co-decision in the human-machine loop. This is because machines require clear commands, and human values are "highly complex, often inconsistent, rarely universal" (Edmonds 2020)[7]. Pointing to the sources of such inconsistency, one speaks of the diametrically opposed operating model of "artificial" intelligence (based on syntactics, contextualisation of large sets) and the "real" one (guided by semantic analysis of concepts) (Chojnowski 2022). This was vividly explained over forty years ago by John Searl in his famous thought experiment called "Chinese room" showing the lack of necessity to understand a message built according to a strict formula: there is a man following instructions in his mother-longue for manipulating Chinese characters, who "produces a likeness of Chinese comprehension by following instructions for manipulating symbols, but does not thereby arrive at Chinese comprehension" (www26). The hackers' successful circumvention of the GPT chat-bot security, perfectly illustrates the ease of manipulating the machine's responses (www24). The problem found its confirmation by GPT's creator, Sam Altman, who had said that "The future of AI is both amazing and terrifying" (www23). His prominent competitors also express this emphatically, signing famous Future of Life Institute's open letter to halt work on AI until rules for its regulation are developed (www27)[8].

At the same time, research is ongoing into the possibility of using AI to correct itself. For example, "toxic artificial intelligence" is being tested, i.e. algorithms specifically taught to use unethical concepts (abuse and prejudice) to effectively detect such anomalies sewn into other algorithms (www1). There are also experiments underway to equip robots with some kind of self-knowledge giving them the ability to improve their own actions (www21). It is therefore also natural to speculate on the possibility of artificial intelligence acquiring a human-like consciousness. To take this problem to the extreme, it is impossible to exclude the possibility of humans overlooking the point at which decisions given over to machines begin to be created by them beyond the limits originally set for them (Hartman 2021). It is also possible to envisage perfect machines that abandon the consciousness they obtained earlier as something superfluous

---

[7] Translation based on Polish edition. Original edition: Philosophers Take On The World, Oxford University Press 2016.

[8] It should be noted that the high-profile appeal is part of a longer story, cf.: (www28).

in the face of the faultless repeatability of action (Schneider 2021). It should be noted, however, that the latter case would not apply to "intelligence", the essence of which includes learning and experimentation. Furthermore, the assumption of a perfectly repeatable result is contradicted by the bias-variance tradeoff theory. In a mathematical sense, the possibility of a perfectly repeatable result from a system operating in a complex environment cannot be proven. In a mathematical sense, the possibility of achieving a perfectly repeatable result by a system operating in a complex environment cannot be proven. However, it must be acknowledged that the conclusion regarding the randomness of the emergence of a conscious AI as depending on phenomena "as unpredictable as the whim of a single designer" (Schneider 2021)[9], cannot be easily dismissed. To summarise this thread, then, it might be assumed that AI consciousness, *if it does exist*, will do so in an environment we do not currently know how to define.

This is not to say that anticipating the development of AI should be guided by dystopian visions of human enslavement by machines. At the same time, it is difficult to deny their growing influence on human behaviour. The ubiquitous staring at smartphones ("smartphone zombies" or "sombies")[10] is just a small example of the side effects of interaction with attractive AI products. However, it demonstrates the ease of manipulation men or women using suitably configured tool. No revolt of the machines is needed. A negative scenario may occur not so much due to the uncontrolled self-development of AI, but due to human negligence or abuse in the area of its application.

Questions about the extent of the limits placed on experimentation with AI therefore become legitimate. An analysis of the fulfilment of the various roles played by AI, e.g. in describing the world around us (in journalism, art), in work systems (e.g. services) or in health treatment and prevention, is likely to be helpful. This kind of analysis, if carried out consistently with democratic rules (i.e. openly and with respect for the various social interests), should make it possible to rethink the long-term effects of the interaction between humans and machines and to arrange their relations flexibly.

---

[9] Translation based on Polish edition. Original edition: S. Schneider, Artificial You: AI and the Future of Your Mind, Princeton University Press 2019.

[10] Cf.: (www29) (www30) (www31).

Thinking machines

The Greek root of the term "autonomy" (αυτονομία) primarily meant the independence (discretion) of standing for one's rights (values): αὐτό – this one, one's own, νομος – custom/law. Thus, the very use of the adjective "autonomous" triggers the potential for association with an action no longer undertaken not only unsupervised, but also at one's own risk. Its use in relation to machines does not imply giving them subjectivity, but nevertheless anticipates the discussion on the right to make decisions currently reserved for humans (conscious and empowered individuals).

The discussion on the range of decisions that an autonomous vehicle could be empowered to make is colloquially reduced to an analysis of the so-called trolley dilemma, i.e. giving critical choices to the machine. However, the journalistic coverage fails to take into account the broad spectrum of problems to be solved (cf. Keeling 2020) and misses the point of the real problems involved in developing AV decision-making capabilities. From a practical point of view, it is much more interesting to gradually clarify the tools of the real-time interaction between humans and automation systems. This is primarily about planning when (and how) to effectively transfer control of the vehicle to each other, and setting boundary conditions for transferring this control to the machine at successive levels of performance. The consideration of the possibility of a AV not applying traffic rules when doing so would enhance the safety of its participants, is a special case of this issue[11]. Framing the dilemma concerning the limits of freedom granted to machines makes it possible to "calibrate" the meaning of the adjectives "thinking" and "autonomous" in relation to machines.

The process of thinking, defined as generalised cognition mediated by the senses and resulting in the ability to make decisions, is descriptively divided into three phases: problem perception, data analysis and solution verification. In the same way, the creation and operation of algorithms, which are models for artificial intelligence, are described. In the simplest terms, an algorithm is a description of an operation. Its effectiveness depends, i.a., on the accuracy of the set of rules to be followed, and on the actual possibility of applying it to different situations. Thinking machines are based on neural networks modelled on the functioning of the human brain, which allows the application of deep learning (DL). DL is the search for correlations and solutions to a given problem, requiring data sets (empirical and statistical) from which the machine creates

[11] For a discussion of this demand, see the concluding chapter of the article.

new cases that expand the initial cognitive base. One way of constructing the algorithms necessary for DL is through heuristics (creating models of situations devoid of some input data). As a result, solutions with different probabilities of occurrence are allowed to be obtained by the machine. Unsupervised machine learning models are used to extend the range of data and find solution paths beyond the training base. While there are not yet satisfactory algorithms for tasks where imagination is required, this does not prevent the employment of AI for activities described as creativity (as in the GPT case) (www22). Efforts are also being made to create robots using abstract intelligence (Czubenko, Kowalczuk 2019). The planning and construction of thinking machines raises a number of questions not only about the creative capabilities of the algorithms, but also about the possible implications of their machine processing of ethical assumptions. While the former are of interest mainly to engineers, the latter are on the minds of philosophers considering their social implications (www23).

Machine morality?

The discussion on how ethical norms emerge has been going on for years among philosophers, anthropologists, sociologists or neuroscientists. Modern science most often describes this process as a relationship between evolutionarily developed behaviour and a community-specific set of norms instilled in the individual during the socialisation process.

An analysis of the views presented in this area by Wittgenstein, points to the thesis of a gradual linguistic "calibration of perception" of mutual behaviour. Anthropology goes hand in hand here, pointing to the clarification of the norms description as the human species evolves. In sociology, on the other hand, the theory of social determinism clash with the theory of intuitive choice of the individual (Zarzycki 2021).

The evolutionary-behavioural approach is vividly supported by neuroscience, which has been developing over the last few decades. Our life choices appear to be conditioned reflexes, which are the resultant of evolutionary shaped factors and developed during socialisation. According to brain researchers, the ability to make decisions does not imply a "free will", because when we react to a given situation, we respond to it based on previously developed, often purely reflexive decision pathways. We therefore superimpose rules instilled during socialisation onto our evolutionary equipment, which means that the brain does

not work on the basis of constant reflection and abstract analysis, on the contrary, its functions are largely schematic and involve drawing on ready-made solutions (Sampolsky 2021).

The question of the "morality of machines" (bearing in mind the hypothesis-generating "*yet*" and "*if*" particles accompanying the consideration of AI potential) should be preceded by another one: can ethical recommendations be limited only to the ways in which modern technologies are used? If ethics is the result of the formation of decision-making patterns and operates under specific social conditions, and as we know from practice: "there is no complete and infallible moral knowledge" (Hołówka 2001), can we reject its modification in the course of interaction with thinking machines?

If it is possible for AI to mature, it must be assumed that what is at stake in the game of building interaction in the human-machine loop is the potential autonomy of machines and the limits of its social acceptance. Its boundary conditions must therefore be defined.

Consumer legal protection in the eu

an increasing number of people are not at all familiar with life before the era of instant messaging, fast search engines, networked broadcasting or real-time navigation. In those who still remember it, the convenience of functioning online is supplanting the once "analogue skills". It is not very clear today how anyone would cope if on-line services were missing from everyday life. It is also likely that an evolutionarily entrenched curiosity about the next developmental challenge: *the AI experiment*, outweighs anxiety about its outcome. Whatever the rationale, it is difficult to imagine the functioning of today's humans without the technological advances of the last two to three decades.

The obverse is the omnipresent surveillance, the greed of service providers for customers' data and their treatment as a resaleable commodity: a profiled marketing target. Internet users' well-known fatigue with formal legal obligations ("privacy fatigue") results in succumbing to the dominance of systems that transform privacy into a currency with which to pay for access to services. "In the multitude of decisions we make every day, taking care to protect data online is a tiresome barrier that slows us down in doing our jobs or other responsibilities. The reason many people are able to accept unfavourable contract terms or share data they don't really want to share is because there is no alternative available or

no time to make the right decisions and take action" (Grzeszak 2020). There is a progressive involuntary descent into a kind of enslavement resulting from the sacrifice of privacy for more or less tangible ad hoc benefits. Admittedly, such a situation has been growing for at least two generations, i.e. since the advent of television talk-shows, whose participants sacrificed their own and their loved ones' privacy for money and temporary fame. The resulting collapse of the taboo of intimacy is one of the side-effects of the prevailing model of consumption (Bauman 2011), but the current loss of control over private data is absolutely unprecedented and it is difficult to identify effective means of stopping it.

However, when it comes to values, it is not binary. Consumers want the freedom to use technological advances and, at the same time, expect control over the process of data use. A Polish study of the willingness to provide relevant personal information to potentially improve the quality of public services showed that "respondents who have a sense of control over data were more likely to declare their willingness to provide data, while respondents who have no control over data were significantly more likely to choose negative answers" (Grzeszak 2020). In contrast, a study by the University of Mannheim found that respondents would accept handing over even justice decisions to artificial intelligence, but (also) only under the condition of human control (www3). It seems that, having gained a certain kind of self-actualisation (or at least belonging) with technology, modern human still feels a lack of safety (like personal data threat). Diverse, but always profiled to attract as much user attention as possible, the products of modern digital technologies make it increasingly difficult to take care of the protection of an individual's sensitive data. This makes it (symptomatically for modern times: out of convenience and not necessarily out of deep conviction) that the concern for this protection is delegated to others, above all to institutions.

The European Union, being one of the addressees of such a delegation, is trying to develop rules for the safe deployment of artificial intelligence and its derivatives. The Explanatory Memorandum of the future AI law states that it will set out "policy options on how to achieve the twin objective of promoting the uptake of AI and of addressing the risks associated with certain uses of such technology" as well as "building an ecosystem of trust by proposing a legal framework for trustworthy artificial intelligence" (COM(2021)206 final). Earlier, with the implementation of the famous General Data Protection Regulation, the EU established the right of a citizen, i.a., "not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing" (Regulation 2016/679). At the same time, the European Court of Justice in one of its judgments, ordered Google to

provide its users with the right "to be forgotten". Similar rights are being created in terms of the entire system of data flows and processing[12].

In the framework of secondary law, making under Article 288 of the Treaty on the Functioning of the European Union, EU bodies may also enact so-called "soft law", e.g. opinions and recommendations. This is, incidentally, an example of an area of regulation developing on the basis of international law, which – being an expression of compromise or a certain intention of the parties – is to be taken into account when applying "hard law". "Ethics of automated and connected vehicles" (CAV Ethics), as a document issued in the form of recommendations, carries a disclaimer stating that its authors are solely responsible for the views contained therein, and that it "do not necessarily reflect the views of the European Commission". However, by virtue of a ruling of the Court of Justice of the EU, recognising that Member State authorities must take into account the content of acts issued under Article 288 TFEU in the interpretation of national law, stakeholders are not entirely free to apply the recommendations and should apply the Ethics – unless they wish to expose themselves to accusations of acting contrary to EU law.

There are also accusations of interpreting the treaties too freely and of the Union's bodies granting themselves "competence to establish competences", e.g. transforming the EU values enshrined in the Treaties into a source of power for the European Commission (Bainczyk 2020). An analysis of the CAV Ethics in this respect indicates that it cannot be subject to such an accusation. The recommendations, as support for the application of purposive interpretation of higher-order rules, may include content that goes beyond procedural comments.

The authors of the Recommendations state that the document "promotes the systematic integration of ethical considerations into the potential transition towards driverless mobility" and to support the integration of them into regulation. They also justify the document with a desire to support "alignment between technology and societal values and for the public to gain trust and acceptance of CAVs". As can easily be seen, this approach is primarily characterised by a concern for the usability of AI derivatives and the safety of their use. This is confirmed by an analysis of the concepts formulating the Recommendations subtitle: "Road safety, privacy, fairness, explainability and responsibility". A closer look at the content of the document, however, makes it possible to find elements

[12] Cf. Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act); Proposal for a Data act (COM(2022)68 final).

that go beyond this limitation and to formulate starting points for their application in the field of the creation of ethics common to humans and machines.

Summary – possible research questions

The authors of the legal analysis of CAV Ethics argue that it have failed to address the fundamental issue of the conflict between the necessary autonomy of human decision-making ("human will") and ensuring traffic safety by artificial intelligence. (Księżak, Wojtczak 2022) This analysis, however, gives perfunctory treatment to the proposed recommendation to consider exempting an autonomous vehicle from traffic regulations in certain circumstances, whereas it may be the element that represents a fundamental change in the approach to the human-machine relationship:

„Traffic rules are a means to road safety, not an end in themselves. Accordingly, the introduction of CAVs requires a careful consideration of the circumstances under which: (a) traffic rules should be changed; (b) CAVs should be allowed to not comply with a traffic rule; or (c) CAVs should hand over control so that a human can make the decision to not comply with a traffic rule" (Ethics of connected and automated vehicles, Recommendation 4).

The implementation of the above recommendation is sometimes jokingly portreyed as a possible reaction of an autonomous vehicle to its owner's treatment of it (e.g. Sampolsky 2021). In reality, asking the question of the AV's right to decide will initiate a process of creating a human-machine loop, or realising Eric Schwitzgebel's idea of "the design policy of the excluded middle": the creation of only those machines that clearly have no moral significance to us – or those that undoubtedly do (www23). For it is necessary to reject the fear of ethical speculation and to anticipate the products that are probably already being hatched in various laboratories.

The multitude of questions surrounding the development of the *AI experiment* and the need to define a framework for the use of its derivatives creates many dilemmas that cannot be resolved today. In the extreme case when thinking machines prove capable of creating solutions beyond human perception, or if neuroscience and cognitive science confirm the mechanical operation of the homo sapiens brain and eliminate "free will" from the vocabulary of our self-knowledge, autonomy will become a concept common to humans and machines. Thinking will then be equated with an algorithm, and the need

for a human-machine ethics will be obvious. For the time being, however, manipulation of artificial intelligence by those with the resources to apply it on a mass scale may be more likely[13]. The democratically uncontrolled use of technologies that give a significant advantage in the use of data may disrupt economic and social relations. Therefore, ethics should, first and foremost, safeguard civil rights.

At the same time, it should be borne in mind that free human experimentation with the environment, of which the climate crisis is one of the consequences, has led to a significant reduction in the margin of error of subsequent experiments. The ethical requirement of the moment, then, is to shoulder the responsibility of getting out of the blind alley of overly free actions (Sloterdijk 2021). This raises another ethical question: what are the limits of *AI experiment* today. Latour's conscious thought about man's necessary responsibility for his own creations, which must not be abandoned after a possible first disappointment, may be helpful in answering it. The right way is to correct and properly safeguard the technologies and products one has created. Doesn't the example of the human hybrid created by Dr Frankenstein, immediately abandoned by him for fear of being morally responsible for transgressing a taboo, fit like a glove to reflect on reactions to the potential of artificial intelligence? (Latour 2012).

The irreversible as it looks process of replacing human beings in their various roles by AI is giving rise to a discussion of its development dilemmas, which may turn out to be (given only basic intuitions):

a. merely a secondary reflection of human thinking errors,
b. a threatening humans autonomous intelligence, or
c. (most likely): a type of parallel to human, separate thinking that needs to be equipped with an element of ethical abstraction.

In this context, the question of whether ethical interventions by transnational institutions can help seems worth asking. A study of institutional positions towards a vision of the future realised with AI appears interesting. Given the axiological aspect of such an analysis, it should be helpful to draw on the strong theses of criticism of contemporary culture: from Nietzsche, through Horkheimer and Adorno's critique of "instrumental reason", to Heidegger's analysis of "technological thinking" and Sloterdijk's theses on the Anthropocene. Rawls's theory of justice, with its conception of society as "a system of cooperation over time,

[13] To be able to support the operation of a GDP-class chat-bot requires significant resources and, including energy.

from generation to generation" (Kędziora 2016) and Foucault's conception of the "moral subjectivity" should also be helpful. Finally, it will not be out of place to filter the institutional approach through the critical of late capitalism concepts presented by Bernard Stiegler and the Collectife Internation (Stiegler 2023), and by Andrzej Leder (Leder 2023).

An analysis structured in this way should provide an answer to the question of the possibility of integrating increasingly efficient thinking machines into the network of community action. Also, countering the unfavourable development of AI will only be possible through social persuasion. "Ethics of automated and connected vehicles" seem to provide a good starting point for such an outlined analysis. The conclusions that can be drawn from its study should serve to assess the potential for the effectiveness of EU (and similar) interventions and to evaluate the social credibility of such a practice.

Bibliography

Legislation and official documents

Regulation (EU) 2016/679 of The European Parliament and of The Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016, p. 1–88, https://eur-lex.europa.eu"https://eur-lex.europa.eu [accessed: 31.01.2023].

Regulation (EU) 2019/2144 of The European Parliament and of The Council of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users […], OJ L 325, 16.12.2019, p. 1–40, https://eur-lex.europa.eu [accessed: 31.01.2023].

Proposal for a Regulation of The European Parliament and of The Council on harmonised rules on fair access to and use of data (Data Act) (COM(2022)68 final), CELEX number: 52022PC0068, https://eur-lex.europa.eu [accessed: 31.01.2023].

Proposal for a Regulation of The European Parliament and of The Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)206), CELEX number: 52021PC0206, https://eur-lex.europa.eu [accessed: 31.01.2023].

Communication from The Commission to The European Parliament, the Council, the European Economic and Social Committee and The Committee of The Regions Sustainable and smart mobility strategy – putting European transport on track for the future (COM(2020)789 final), CELEX number: 52020DC0789, https://eur-lex.europa.eu [accessed: 31.01.2023].

Communication from The Commission to The European Parliament, the Council, the European Economic and Social Committee and The Committee of The Regions On the road to automated mobility: An EU strategy for mobility of the future (COM(2018)283 final), CELEX number: 52018DC0283, https://eur-lex.europa.eu [accessed: 31.01.2023].

European Commission, Directorate-General for Research and Innovation, Ethics of connected and automated vehicles: recommendations on road safety, privacy, fairness, explainability and responsibility, Publications Office, 2020, https://data.europa.eu/doi/10.2777/035239"https://data.europa.eu/doi/10.2777/035239 [accessed: 31.01.2023].

Ethics by design and ethics of use approaches for Artificial Intelligence, European Commission 2021, https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf"https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf [accessed: 31.01.2023].

United Nations Economic Commission for Europe (UNECE) – Working party on Automated/Autonomous and Connected Vehicles (GRVA), All you need to know about Automated Vehicles. Technical progress and regulatory activities, https://unece.org/sites/default/files/2022-01/Brochure%20Automated%20Vehicles.pdf"https://unece.org/sites/default/files/2022-01/Brochure%20Automated%20Vehicles.pdf [accessed: 31.01.2023].

Articles, books and monographs

Alfthan von K., Hyry J. (2020), *Sitra measuring wellbeing – citizen survey. Finland, Germany, The Netherlands and France total 2020 report*, https://www.sitra.fi/app/uploads/2020/10/measuring-wellbeing-citizen-survey-finland-germany-the-netherlands-and-france.pdf [accessed: 27.10.2022].

Bainczyk M. (2020), *Podział kompetencji między Unię Europejską a państwa członkowskie jako kluczowe wyzwanie dla przyszłości UE*, (w:) Domaradzki S., Zheltovsky V. (red.), *Racja stanu Polski w Unii Europejskiej*, Oficyna Wydawnicza Aspra, Warszawa.

Carr N. (2013), *Płytki umysł*, Helion, Gliwice.

Cesarz M. (2014), *Porządek prawny Unii Europejskiej*, OTO-Wrocław, http://www.repozytorium.uni.wroc.pl/Content/60497/PDF/07_Maciej_Cesarz.pdf [accessed: 23.01.2023].

Chojnowski M. (2022), *Etyka sztucznej inteligencji. Wprowadzenie*, Centrum Etyki Technologii Instytutu Humanites, https://ethicstech.eu/wp-content/uploads/2022/12/ESIW-v2.0_FINAL-2.pdf [accessed: 15.10.2022].

Choromański W. i in. (2020), *Pojazdy autonomiczne i systemy transportu autonomicznego*, Wydawnictwo Naukowe PWN, Warszawa.

Czubenko M., Kowalczuk Z. (2019), *Elementy psychologii w kontekście autonomii robotów*, Pomorskie Wydawnictwo Naukowo-Techniczne PWNT, Gdańsk.

Edmonds D. (red.) (2020), *Filozofia kontra świat*, Wydawnictwo Naukowe PWN, Warszawa.

Grzeszak J. i in. (2020), *Czy chcemy dzielić się prywatnymi danymi?*, Polski Instytut Ekonomiczny, https://pie.net.pl/wp-content/uploads/2021/04/PIE-Raport-Prywatne-dane.pdf [accessed: 01.10.2022].

Hartman J. (2021), *Etyka życia codziennego*, Wydawnictwo Naukowe PWN, Warszawa.

Hołówka J. (2001), *Etyka w działaniu*, Prószyński i S-ka, Warszawa.

Kamieński Ł. (2022), *Mimowolne cyborgi. Mózg i wojna przyszłości*, Czarne, Warszawa.

Keeling G. (2020), *Why Trolley Problems Matter for the Ethics of Automated Vehicles*, "Science and Engineering Ethics", No. 26, https://link.springer.com/article/10.1007/s11948-019-00096-1"https://link.springer.com/article/10.1007/s11948-019-00096-1 accessed: 31.01.2023].

Kędziora K. (2016), *John Rawls: uzasadnienie praktyczne koncepcji sprawiedliwości i refleksyjna równowaga*, „Studia Philosophiae Christianae", nr 3.

Kloc-Konkołowicz J. (2000), *Theodor Adorno wobec aporii nowoczesności. Krytyka „niemożliwa"*, „Etyka", nr 33.

Księżak P., Wojtczak S. (2022), *A human being must obey the commands of a robot! CAVs, Asimov's Second Law and the new ground-breaking ethics*, https://www.academia.edu/91633650/A_Human_Being_Must_Obey_the_Commands_of_a_Robot_CAVs_Asimov_s_Second_Law_and_the_New_Ground_Breaking_Ethics [accessed: 04.02.2023].

Latour B. (2012), *Love your monsters*, The Breakthrough Institute, https://thebreakthrough.org/journal/issue-2/love-your-monsters [accessed: 31.01.2023].

Leder A. (2023), *Ekonomia to stan umysłu*, Wydawnictwo Krytyki Politycznej, Warszawa.

Litman T. (2023), *Autonomous Vehicle Implementation Predictions. Implications for Transport Planning*, Victoria Transport Policy Institute, https://www.vtpi.org/avip.pdf [accessed: 31.01.2023].

Minkier K. i in. (2019), *Wprowadzenie do global studies*, Wydawnictwo Adam Marszałek, Toruń.

Napiórkowski M. (2022), *Naprawić przyszłość*, Wydawnictwo Literackie, Warszawa.

Sampolsky R.M. (2021), *Zachowuj się*, Media Rodzina, Poznań.

Sandel M.J. (2020), *Przeciwko udoskonalaniu człowieka*, Wydawnictwo Naukowe PWN, Warszawa.

Schneider S. (2021), *Świadome maszyny. Sztuczna inteligencja i projektowanie umysłów*, Wydawnictwo Naukowe PWN, Warszawa.

Sloterdijk P. (2021), *Co się zdarzyło w XX wieku?*, Wydawnictwo Aletheia, Warszawa.

Stiegler B. i Kolektyw Internacja (2023), *Konieczna bifurkacja*, Wydawnictwo Uniwersytetu Śląskiego, Katowice.

Szafrański D. (2023), *Źródła prawa Unii Europejskiej (po wejściu w życie Traktatu Lizbońskiego)*, Wydział Prawa i Administracji UW, https://wpia.uw.edu.pl/uploads/media/5da76c556fca7/zrodla-prawa-unii-europejskiej-dr-dariusz-szafranski.pdf [accessed: 23.01.2023].

Tanenblatt E.J. (2022), *Global Guide to autonomous vehicles 2022*, Denton, https://www.thedriverlesscommute.com/wp-content/uploads/2022/06/Global-Guide-to-Autonomous-Vehicles-2022.pdf [accessed: 09.11.2022].

Threlfall R. et al. (2020), *KPMG 2020 Autonomous Vehicles Readiness Index*, KPMG International, https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2020/07/2020-autonomous-vehicles-readiness-index.pdf [accessed: 09.11.2022].

Zarzycki R. (2021), *Etyka codzienności – czy można być dobrym po prostu?*, Collegium Civitas Press, https://www.civitas.edu.pl/wp-content/uploads/2015/03/roland-zarzycki-etyka-codziennosci.pdf [accessed: 31.01.2023].

Online articles

https://www.forbes.com/sites/lanceeliot/2022/06/15/ai-ethics-shocking-revelation-that-training-ai-to-be-toxic-or-biased-might-be-beneficial-including-for-those-autonomous-self-driving-cars/?sh=66a5cbc11baf [accessed: 28.09.2022].

https://wyborcza.pl/7,75400,28964959,czy-decyzje-podejmowane-przez-komputery-uznamy-za-sprawiedliwe.html [accessed: 01.10.2022].

https://www.cell.com/patterns/fulltext/S2666-3899(22)00209-4#S.embed_link-K.C-B.1-L.1.zw [accessed: 01.10.2022].

https://www.sztucznainteligencja.org.pl/duch-w-maszynie/ [accessed: 15.10.2022].

https://krytykapolityczna.pl/kultura/mistycyzm-popkulturowy-czy-sztuczna-inteligencja-zastapi-artystow/ [accessed: 17.10.2022].
https://www.deeptechnology.ai/zastosowania-ai-w-medycynie/?gclid=Cj0KCQjw166aBhDEARIsAMEyZh7GxwD7dYEE6zqLQZU-8UC_MHXGCEBd-hgg7P-nyGOp57iir4tmJFYaAnlXEALw_wcB [accessed: 17.10.2022].
https://amu.edu.pl/nauka/popularyzacja-nauki/artykuly-popularyzacja-nauki/badania-uam-jak-sztuczna-inteligencja-wplywa-na-ludzkie-zachowania [accessed: 17.10.2022].
https://mlodytechnik.pl/m-technik/28734-czy-da-sie-calkiem-zniknac-z-internetu [accessed: 17.10.2022].
https://www.itu.int/itu-d/reports/statistics/2021/11/15/mobile-phone-ownership/ [accessed: 17.10.2022].
https://cyfrowa.rp.pl/technologie/art17965591-polacy-maja-zaskakujacy-stosunek-do-sztucznej-inteligencji [accessed: 17.10.2022].
https://us06web.zoom.us/j/83089110226?pwd=OFBIbHlPTFlMRGxyU3lFSHV3YmxVQT09 [accessed: 18.10.2022].
https://wyborcza.biz/biznes/7,177150,29042054,polski-raper-stworzyl-klip-dzieki-sztucznej-inteligencji-technologia.html?disableRedirects=true [accessed: 20.10.2022].
https://www.projektpulsar.pl/srodowisko/2183046,1,marcin-ryszkiewicz-technologia-ocali-nas-i-matke-nasza-gaje.read [accessed: 21.10.2022].
https://data.euroAV.eu/data/datasets/s2228_92_4_503_eng?locale=en [accessed: 24.10.2022].
https://encyklopedia.pwn.pl/haslo/myslenie;4009251.html [accessed: 24.10.2022].
https://www.sas.com/pl_pl/insights/analytics/what-is-artificial-intelligence.html [accessed: 25.10.2022].
https://www.userconf.pl/ciekawostki/czlowiek-vs-maszyna-jaki-jest-wynik/ [accessed: 31.10.2022].
https://naukawpolsce.pl/aktualnosci/news%2C92876%2Cinteligentny-robot-zostal-seksista-i-rasista.html [accessed: 01.11.2022].
https://www.quora.com/If-robots-become-romantic-AVrtners-to-humans-will-they-program-them-to-disagree-or-will-the-relationship-be-perfect-and-drama-free [accessed: 01.11.2022].
https://cyfrowa.rp.pl/technologie/art37634921-rewolucja-chatgpt-poprosilismy-ai-o-napisanie-tego-artykulu-i-ilustracje [accessed: 28.12.2022].
https://www.nytimes.com/2023/01/06/science/robots-artificial-intelligence-consciousness.html [accessed: 16.01.2023].

https://www.jdpower.com/cars/shopping-guides/consumer-readiness-for-autonomous-vehicles-remains-low [accessed: 31.01.2023].
https://finance.yahoo.com/news/sam-altman-maker-chatgpt-says-110000987.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAABHP5sgFcHxNrqfieTPXT5uu-BeZYD9QdXxbEXCwoOaWKMeP9EfbQ7HWasfFVkaeNb21z61RZACkYXhuo4fHSMwnmTFQgkuUrCjI4lExrIH5pPWk-dfvyrDyHaS6Ga6OfxLGtgO1gOfQnix362n-Svw1pG-S0QC0DfeT-ykIS-yU [accessed: 01.02.2023].
https://antyweb.pl/chatgpt-na-kodach [accessed: 07.02.2023].
https://www.cbmclaw.com/what-percentage-of-car-accidents-are-caused-by-human-error/ [accessed: 17.02.2023].
https://pl.wikipedia.org/wiki/Chiński_pokój [accessed: 19.02.2023].
https://futureoflife.org/open-letter/ai-open-letter/ [accessed: 29.03.2023].
https://www.rp.pl/nowe-technologie/art12051181-naukowcy-ostrzegaja-przed-inteligencja-maszyn [accessed: 29.03.2023].
https://www.bbc.com/news/blogs-news-from-elsewhere-36576259 [accessed: 29.03.2023].
https://heyme.care/fr/blog/smombie-quand-ton-smartphone-te-transforme-en-zombie [accessed: 11.04.2023].
https://www.edeka-smart.de/news/smombie-was-ist-das-29 [accessed: 11.04.2023].

Afiliacja: **mgr Sławomir Lewandowski**
Akademia im. Aleksandra Gieysztora w Pułtusku
(filia Akademii Finansów i Biznesu Vistula)
ul. Mickiewicza 36 B
06-100 Pułtusk
e-mail: slawomir.lewandowski2000@gmail.com.