

THE (BUSINESS) ETHICS OF AI REGULATION

Toni GIBEA^{a1}, Radu USZKAI^b

^{a, b} Bucharest University of Economic Studies, Romania; Romanian Young Academy, University of Bucharest, Romania

ABSTRACT

In the past couple of years, more and more companies have been trying to integrate, within their ethical infrastructure, a varying degree and amount of ethical concerns regarding the development and deployment of Artificial Intelligence (AI) systems. As a result, it would not be an exaggeration to say that we have witnessed an explosion of ethics codes concerning AI. The main purpose of this paper is to explore the way in which business organizations have dealt with such concerns. In particular, we first aim to analyze whether companies have a genuine interest in AI ethics or whether it is nothing more than a case of ethics washing, making this whole enterprise virtually useless. We will conclude by advancing an agenda regarding how AI ethical regulations could be empirically assessed, by highlighting a few of the downsides and explaining how different experimental methods could help us close the empirical knowledge gap.

KEYWORDS: *Artificial Intelligence, business ethics, code of ethics, ethical principles, longitudinal studies.*

DOI: 10.24818/IMC/2022/02.06

1. INTRODUCTION

While researchers in academia are debating the ethics of predictive and algorithmic policing (Moses & Chan, 2018; Alikhademi et al., 2022), or the use of facial recognition for mass surveillance, San Francisco voted, in late November 2022, to allow police to use killer robots (Wu, 2022). Discussions regarding the development and deployment of killer drones, robots, and Autonomous Weapons Systems are even older (Sparrow, 2007; Sparrow, 2015), but some have already been seen on the field of battle. Further issues have been raised with regards to other types of AI deployment. For instance, scoring systems can (potentially) pose a serious threat since they can identify and exploit various weaknesses that we have (and which we might not even be aware of) (Citron & Pasquale, 2014). Various decision-making algorithms used in HR Management can improve efficiency, but they can also reinforce racial prejudices and biases (Budhwar et al. 2022).

With constant progress being made in the field of computer science and AI research, it seems plausible that most (if not all) areas of human activity can, and will be affected by AI systems. The crux of the problem is that, since such systems are able to operate and take decisions instantaneously, automatically and autonomously, and at scale, our lives can be impacted (both positively and negatively) in the blink of an eye at "orders of magnitude and at speeds not previously possible" (Yeung, Howes & Pogrebna, 2019).

While AI tools and solutions have been developed and proposed as tools to help us fix various problems, the selection of examples we provided earlier stands as proof for the fact that "AI

¹ Corresponding author. E-mail address: toni.gibea@man.ase.ro

technologies have pushed the significance of dual-use to the extreme: whether we think about autonomous weapons, facial recognition technologies or already mundane decision-making software, all of these applications can be used for both good and bad purposes" (Vică et al., 2021). Even if the definition of what counts as dual-use (be it research, technologies or artifacts) raises some interesting philosophical questions (Forge, 2010), we consider that the idea that research in AI, the deployment of AI systems, or the use of AI artifacts raises dual-use types of issues is beyond any reasonable doubt. In fact, it is specifically because so many people have raised this issue that, lately, there has been a flurry of proposals for integrating AI ethics both at a national and international level, but also within business organizations. Since we are skeptical that moral diplomacies can prove a workable solution for AI ethics (Vică et al., 2021), we turn to exploring the way in which companies try to self-regulate their own AI research and deployment by compiling codes of ethics. The next section of this paper is dedicated to an overview of these self-regulatory attempts, with a focus on what some in the academic literature have labeled "ethics washing". We then proceed to explore why AI ethics seems to be largely useless in practice. Last but not least, we conclude with an extended discussion of how we can assess the efficiency of AI principles.

2. SELF-REGULATORY ATTEMPTS

Whenever companies either do, or are on the brink of doing something that raises moral concerns, there is a growing pressure on public authorities to strongly regulate them. This constantly raises some important philosophical and public choice concerns that have been previously discussed in depth by Munger (2014) which would be, unfortunately, beyond the scope of this paper. In other cases, there is a strong push from people from the fields of business ethics or ethics management for companies to self-regulate, by building on the capacities of their ethical infrastructure, especially on their codes of ethics.

While it might sound good on paper, having companies invested or interested in ethics is by no means a definite proof that they actually care about those issues, or that they are actively searching for ways to solve them. In fact, in some cases their interest in ethics amounts to nothing more than moral grandstanding (Voinea & Uszkai, 2020). In others, ethics seems to be used as a tool to either wash the image of the company, or to dodge potential harder regulation from the national governments and international organizations. It is for this reason that more and more scholars are talking about greenwashing (Marquis et al., 2016), pinkwashing (Carter, 2015) and sportswashing (Fruh et al., 2022). In a similar vein, some scholars have been arguing that, when it comes to AI ethics and companies drafting codes of ethics, what is really happening is "ethics washing".

The term was coined by Metzinger in 2019, when discussing the way in which the AI and tech industry has instrumentalized ethics. The purpose of this instrumentalization of AI ethics was, according to him, pretty simple: they do this in order to "distract the public and to prevent or at least delay effective regulation and policy-making" (Metzinger, 2019). Others have made a similar point, highlighting the gap between the perceived commitments from the codes of ethics and other similar documents, and their real intent and decisions or actions in real life (Bietti, 2020; Peukert and Kloker 2020; Vică et al., 2021). As a result, some have argued that "AI ethics" might be nothing more than a buzzword (Vică et al., 2021). However, what is clear is that the reason mentioned above, alongside many others, gave an incentive to different governing bodies, private sectors and academia to develop AI moral principles. The aggregated contributions can be examined on a website called "AI Ethics Guidelines Global Inventory" (Algorithmwatch, 2022) where 167 documents show us how multiple entities tackled normative challenges posed by the AI with formal institutional means. If we want proof that companies lack a genuine ethical commitment in the case of AI we should look no further beyond an overview of the way in which such documents and codes of ethics are generated and of their current form.

Most of them are principled-based and drafted within a deontological approach. Big ethical principles like autonomy, justice, beneficence or non-maleficence are effectively promoted in the health-care sector where a theory like principlism manages to specify and successfully apply its theoretical account to real-life situations. Sadly, principlism and its principles cannot be directly transmuted to AI for several different reasons. The most obvious one is that the AI industry shares no common goal. In medicine the common goal is to promote health while in AI you have legitimate different goals, for example one could make the case that AI could be used to assist managers when they want to make a profitable decision for the company (Mittelstadt, 2019). A further and related issue has to do with the fact that principled-based, deontological codes of ethics are not necessarily what the workers in the tech industry are fond of. For instance, a clear majority of the people employed in the tech industry (79% to be more precise) have a strong preference for clearer and more practical instructions instead of grand principles (Miller & Coldicott 2019). Moreover, such a focus on developing guidelines based on strong principles might very well be futile, since the nature of reality (i.e. the way in which AI systems are developed or deployed) is too complex to be reducible to such abstract ways of guiding behavior.

In light of such concerns, some have argued that AI ethics is nothing more than useless. Munn (2022), for instance, goes on to label the principles of AI ethics as being either “meaningless”, or “isolated” from the actual issues that they are devised to solve. By looking specifically at the way in which ethics codes are generated, Munn points out two serious problems of this process. First and foremost, if these codes are generated within a bottom-up approach, then we are faced with some important shortcomings. As many scholars have pointed out, the tech industry is marked by a flurry of moral issues. Furthermore, taking into account the lack of serious ethical training for computer or data scientists in both academic (while studying to get a degree) and non-academic settings (while being employed by a company), the outputs of such a process will exhibit, predictably, some of these issues. On the other hand, the top-down approach (which, we should add, is the preferred way of generating such codes and documents) has its own severe limitation (which we alluded to earlier): setting up a normative framework with some normative ideals might sound good in theory, but in practice they tend to be divorced from the day-to-day work experience of computer scientists, data analysts or managers who work with AI systems.

As an additional proof for the uselessness of AI ethics we have coined the ‘ought-is problem of AI ethics’. Reversing the classical Humean ‘is-ought’ problem, those who draft such codes and documents seem to think that, once generating a series of ‘oughts’ for employees or managers, the reality of their day to day working life will simply ‘be’ different. If something ought to be done with regards to AI, then it will be done. However, if we look at some recent empirical work on this issue, we can see how big the gap between intentions and reality actually is. For instance, McNamara et al. (2018) wanted to explore the degree to which priming people who work as software developers or who are students in IT with the ACM code of ethics might change their decision making as opposed to a control group. The results were pretty bleak: no difference was observed between the two groups. The two studies from Vakkuri et al. (2019; 2020) point towards a similar conclusion. Regardless of the fact that we look at small start-ups or big companies, the individuals involved in developing AI tools think that ethics is important but they acknowledge that, when faced with an important decision, they do not take it into consideration. However, as we will show in the next section, maybe more work is needed in order to arrive at such conclusions.

3. ASSESSING THE EFFICIENCY OF AI PRINCIPLES

The hypothesis that some companies might endorse certain ethical guidelines or codes for AI in order to avoid disagreement, controversy or liabilities is a sound and important case of ethics washing. There are multiple ways in which the actual interest of the companies for ethical problems is examined in the business ethics literature but only in relation to common business ethics codes (Kaptein, 2010;

Kaptein & Schwartz, 2008). It is harder to evaluate the effectiveness and how serious AI ethical principles are taken into account merely because of their nascent nature. There are notable attempts in the literature which focus on the content of the AI ethics codes and guidelines examining how normatively well formed they are based on well defined screening criterias (Schiff et al., 2020, 2021; Zhou et al., 2020). Also, Kelley (2022) examined the perceptions through 49 interviews of the employees after AI principles were formally implemented in their own organization. The target group of the study is composed from employees who work in the AI industry and who are in touch with specific moral principles in their own organization. Most of the interviewed employees agree that there are some new perspectives brought into their attention by the principles specified in the AI guidelines or codes, but they clearly stated that AI principles were only formally introduced to them. Some of them explicitly mentioned that they do not feel in charge nor do they believe that they could effectively tackle the specific moral issues or dilemmas raised in the context of new technologies, less so when it comes to AI related technologies. The study (Kelley, 2022) concludes, in a positive note, that the implementation of AI principles is effective despite its clear and imminent limitations. We believe that the empirical knowledge gap in the implementation of different ethical regulations is more profound than the author leads us to believe and it rests mainly on empirical methodological shortcomings.

From a methodological point of view, it is hard to assess in an objective way the impact of AI principles in the industry through interviews. And for that matter all qualitative studies are dealing with similar issues. Most of the employees are reluctant in expressing their own free opinions especially on topics regarding ethics. Also, a dozen of individual personal opinions can hardly make an objective overview. The researchers who evaluate the interviews are rarely biased free when they interpret the answers. Even more than that, it is challenging during the interviews to adopt a neutral or impartial attitude. This step is paramount because the researcher needs to know that she had little to no influence over the target group or that no personal suggestions were asserted. In another case, some participants might prefer to avoid being honest altogether only because they do not trust the interviewer. And it does not necessarily mean that they will prefer to say that everything is fine. Quite the opposite. A raging employee who plans to leave the company could act as an overzealous whistleblower, making the interviewer believe that the company is facing serious ethical problems. And the list of methodological shortcomings continues but it is important to keep a clear head about it. It does not mean that qualitative research is useless, it only means that we should look at the data with a healthy dose of skepticism and corroborate it with additional different types of data.

In general, quantitative research is deemed as being more objective than qualitative because it manages to use impartial and impersonal surveys which contain scales able to detect normatively significant characteristics and register small changes into the process of decision-making. In experimental ethics quantitative research is broadly used (Graham et al., 2009; Greene et al., 2009). Even more than that, recent advancements in this field showed us that ethical debates and discussions might have an impact upon the behavior of the students therefore extending their area of focus from moral decisions and judgments to moral behavior. For example, seminars on animal ethics lower the quantity of meat consumption at the cafeteria of the university (Schwitzgebel et al., 2020). There are also several studies which raise skepticism whether educational measures enhance the moral reasoning capacities or the behavior of the attendants. We know for a fact that ethicists' moral judgments are as prone to biases as non-ethicists' and that ethicists' behavior is not above the moral threshold of the non-ethicists (Schwitzgebel, 2009, 2014; Schwitzgebel & Cushman, 2012, 2015). It means that moral knowledge might have a limited motivational impact on moral behavior.

Another essential work on the topic of ethical behavior is the case of ethical education which presumably should enhance the behavior of future professionals (Brown et al., 2020), or participants' sense of justice (Sabbagh, 2021). According to a meta-analysis, seminars on ethics seem to be relatively effective if they are conducted in an interactive Socratic manner (Antes et al., 2009). Another study reported no difference at the level of ethical attitudes between students who graduated

an ethics course and those who didn't (Jewe, 2008). In yet another study, it is claimed that a course on ethics significantly improves the ethical judgment ability of accounting students (Mohamed Saat et al., 2010) and in an older study that courses on ethics substantially improved the ability to recognize ethical issues (Gautschi & Jones, 1998). Other researchers posed a more skeptical position, for example in a study it was showed that corporate ethics alongside corporate responsibility courses do not play a relevant role in students' ethical awareness (Tormo-Carbó et al., 2016) nor do standalone courses on ethics change students' attitudes or beliefs (Lord & Bjerregaard, 2003).

One potential issue, from a long list, is that all these empirical studies are not longitudinal studies in nature. Longitudinal studies are essential when we want to study the impact of a certain variable along a time series (i.e. Cowpertwait & Metcalfe, 2009). For example, when we try to discover whether smoking is unhealthy, the health status of smokers is screened over long periods of time and their case is compared with the case of non-smokers. A longitudinal study is an endeavor which would measure the moral status of an organization in relation to AI before implementing a code of ethics or a guideline and then compare the result with additional data after the guidelines and the codes were deployed, half a year, a year or over longer periods of time. This type of research would be able to close the empirical knowledge gap in this particular case because it would directly target the issues at hand in terms of efficiency and of actual moral change proving that AI principles are not just some attempt of 'ethics washing'.

To conclude, there are plenty of reasons why extensive longitudinal works are quite rare on the topic of ethical judgment or behavior. We will shortly summarize three of them which are also closely related. First, it is a well-known fact that behavioral and moral variables are hard to empirically control in an experimental study. Second, assuming that we would be able to closely control all the relevant variables, then we face serious ethical risks, like infringing on the rights of our human subjects. Third, all involved parties need to make a long term commitment in order to assess the impact of moral AI principles in organizations and fully disclose all the relevant aspects, conditions which could be hardly met without an incentive, if ever. Nevertheless, these are the challenges we need to overcome in order to narrow the empirical knowledge gap and if we want to quantify how efficient AI principles are in organizations.

ACKNOWLEDGMENT

Toni Gibea's contribution was supported by a postdoctoral grant from the Romanian Young Academy funded by Stiftung Mercator and hosted by the University of Bucharest (2021-2023); Radu Uszkai's contribution has benefited from the support of the Romanian Young Academy, which is funded by Stiftung Mercator and the Alexander von Humboldt Foundation for the period 2020-2022.

REFERENCES

- Alikhademi, K., Drobina, E., Prioleau, D., Richardson, B., Purves, D. & Gilbert, J.E. (2022). A review of predictive policing from the perspective of fairness. *Artificial Intelligence and Law*, 30, 1–17. <https://doi.org/10.1007/s10506-021-09286-4>
- Algorithmwatch (2022). AI Ethics Guidelines Global Inventory. Retrieved December 1st, 2022, from <https://inventory.algorithmwatch.org/>
- Antes, A. L., Murphy, S. T., Waples, E. P., Mumford, M. D., Brown, R. P., Connelly, S., & Devenport, L. D. (2009). A Meta-Analysis of Ethics Instruction Effectiveness in the Sciences. *Ethics & Behavior*, 19(5), 379–402. <https://doi.org/10.1080/10508420903035380>

- Bietti, E. (2020). From Ethics Washing to Ethics Bashing: A View on Tech Ethics from within Moral Philosophy. *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, (pp. 210–19). <https://doi.org/10.1145/3351095.3372860>.
- Brown, T., Isbel, S., Logan, A., & Etherington, J. (2020). Predictors of academic integrity in undergraduate and graduate-entry masters occupational therapy students. *Hong Kong Journal of Occupational Therapy*, 33(2), 42–54. <https://doi.org/10.1177/1569186120968035>
- Budhwar, P., Malik, A., De Silva, M. T. T. & Thevisuthan, P. (2022) Artificial intelligence – challenges and opportunities for international HRM: a review and research agenda. *The International Journal of Human Resource Management*, 33(6), 1065-1097.
- Carter M. (2015). Backlash against” pinkwashing” of breast cancer awareness campaigns. *BMJ: British Medical Journal*, 351, 1–2.
- Citron, D. K. & Pasquale, F. (2014). The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, 89(1), 1–33.
- Cowpertwait, P. S. P., & Metcalfe, A. V. (2009). *Introductory Time Series with R*. Springer Science & Business Media.
- Forge, J. (2010). A Note on the Definition of “Dual Use”. *Science and Engineering Ethics*, 16, 111–118. <https://doi.org/10.1007/s11948-009-9159-9>
- Fruh, K., Archer, A. & Wojtowicz, J. (2022). Sportswashing: Complicity and Corruption. *Sport, Ethics and Philosophy*, DOI: 10.1080/17511321.2022.2107697
- Gautschi, F. H., & Jones, T. M. (1998). Enhancing the Ability of Business Students to Recognize Ethical Issues: An Empirical Assessment of the Effectiveness of a Course in Business Ethics. *Journal of Business Ethics*, 17(2), 205–216. <https://doi.org/10.1023/A:1005740505208>
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5), 1029–1046.
- Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111(3), Article 3. <https://doi.org/10.1016/j.cognition.2009.02.001>
- Jewe, R. D. (2008). Do Business Ethics Courses Work? The Effectiveness of Business Ethics Education: An Empirical Study. *Journal of Global Business Issues*, 1–6.
- Kaptein, M. (2010). The Ethics of Organizations: A Longitudinal Study of the U.S. Working Population. *Journal of Business Ethics*, 92(4), 601–618. <https://doi.org/10.1007/s10551-009-0175-0>
- Kaptein, M., & Schwartz, M. S. (2008). The Effectiveness of Business Codes: A Critical Examination of Existing Studies and the Development of an Integrated Research Model. *Journal of Business Ethics*, 77(2), 111–127. <https://doi.org/10.1007/s10551-006-9305-0>
- Kelley, S. (2022). Employee Perceptions of the Effective Adoption of AI Principles. *Journal of Business Ethics*, 178(4), 871–893. <https://doi.org/10.1007/s10551-022-05051-y>
- Lord, V. B., & Bjerregaard, B. E. (2003). Ethics courses: Their impact on the values and ethical decisions of criminal justice students. *Journal of Criminal Justice Education*, 14(2), 191–211. <https://doi.org/10.1080/10511250300085751>
- Marquis C., Toffel M. & Zhou Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, 27(2), 483–504.
- McNamara, A., Smith, J. & Murphy-Hill, E. (2018). Does ACM’s code of ethics change ethical decision making in software development? *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (pp. 729–733).
- Metzinger, T. (2019, August 4). Ethics washing made in Europe. *Der Tagesspiegel*. Retrieved December 1st, 2022, from <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html>.

- Miller, C. & Coldicott, R. (2021). People, Power and Technology: The Tech Workers' View. Retrieved December 1st, 2021, from <https://doteveryone.org.uk/report/workersview/>
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), Article 11. <https://doi.org/10.1038/s42256-019-0114-4>
- Mohamed S., M., Porter, S., & Woodbine, G. (2010). The effect of ethics courses on the ethical judgement- making ability of Malaysian accounting students. *Journal of Financial Reporting and Accounting*, 8(2), 92–109. <https://doi.org/10.1108/19852511011088361>
- Moses, L.B. & Chan, J. (2018). Algorithmic prediction in policing: assumptions, evaluation, and accountability. *Policing and Society*, 28(7), 806-822.
- Munger, M. (2014, August 11). Unicorn Governance. *Foundation for Economic Education*. Retrieved December 1st, 2022, from <https://fee.org/articles/unicorn-governance>
- Munn, L. (2022). The uselessness of AI ethics. *AI and Ethics*. <https://doi.org/10.1007/s43681-022-00209-w>
- Peukert, C. & Kloker, S. (2020). Trustworthy AI: How Ethics Washing Undermines Consumer Trust. *Proceedings of the 15th International Conference on Wirtschaftsinformatik*. Potsdam, Germany. https://doi.org/10.30844/Wi_2020_j11-Peukert
- Sabbagh, C. (2021). Self-reported academic performance and academic cheating: Exploring the role of the perceived classroom (in)justice mediators. *British Journal of Educational Psychology*, n/a(n/a), e12433. <https://doi.org/10.1111/bjep.12433>
- Schiff, D., Biddle, J., Borenstein, J., & Laas, K. (2020). What's Next for AI Ethics, Policy, and Governance? A Global Overview. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 153–158. <https://doi.org/10.1145/3375627.3375804>
- Schiff, D., Borenstein, J., Biddle, J., & Laas, K. (2021). AI Ethics in the Public, Private, and NGO Sectors: A Review of a Global Document Collection. *IEEE Transactions on Technology and Society*, 2(1), 31–42. <https://doi.org/10.1109/TTS.2021.3052127>
- Schwitzgebel, E. (2009). Do ethicists steal more books? *Philosophical Psychology*, 22(6), 711–725. <https://doi.org/10.1080/09515080903409952>
- Schwitzgebel, E. (2014). The Moral Behavior of Ethicists and the Role of the Philosopher. In C. Luetge, H. Rusch, & M. Uhl (Eds.), *Experimental Ethics: Toward an Empirical Moral Philosophy* (pp. 59–64). Palgrave Macmillan UK. https://doi.org/10.1057/9781137409805_5
- Schwitzgebel, E., Cokelet, B., & Singer, P. (2020). Do ethics classes influence student behavior? Case study: Teaching the ethics of eating meat. *Cognition*, 203, 104397. <https://doi.org/10.1016/j.cognition.2020.104397>
- Schwitzgebel, E., & Cushman, F. (2012). Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind & Language*, 27(2), 135–153. <https://doi.org/10.1111/j.1468-0017.2012.01438.x>
- Schwitzgebel, E., & Cushman, F. (2015). Philosophers' biased judgments persist despite training, expertise and reflection. *Cognition*, 141, 127–137. <https://doi.org/10.1016/j.cognition.2015.04.015>
- Sparrow, R. (2007). Killer Robots. *Journal of Applied Philosophy*, 24(1), 62-77.
- Sparrow, R. (2015). Killer robots: Ethical issues in the design of unmanned systems for military applications. In K. P. Valavanis, & G. J. Vachtsevanos (Eds.), *Handbook of Unmanned Aerial Vehicles* (pp. 2965-2983). Springer.
- Tormo-Carbó, G., Oltra, V., Seguí-Mas, E., & Klimkiewicz, K. (2016). How Effective are Business Ethics/CSR Courses in Higher Education? *Procedia - Social and Behavioral Sciences*, 228, 567–574. <https://doi.org/10.1016/j.sbspro.2016.07.087>
- Vakkuri, V., Kemell, K.-K., Kultanen, J., Siponen, M. & Abrahamsson, P. (2019). *Ethically aligned design of autonomous systems: industry viewpoint and an empirical study*. <https://doi.org/10.48550/arXiv.1906.07946> (2019).

- Vakkuri, V., Kemell, K.K., Jantunen, M. & Abrahamsson, P. (2020). This is Just a Prototype": How Ethics Are Ignored in Software Startup-Like Environments. In Stray, V., Hoda, R., Paasivaara, M., Kruchten, P. (eds.) *Agile Processes in Software Engineering and Extreme Programming. XP 2020. Lecture Notes in Business Information Processing*, vol 383 (195-210). Springer. https://doi.org/10.1007/978-3-030-49392-9_13
- Voinea, C. & Uszkai, R. (2020). Do Companies Engage in Moral Grandstanding? *Paper presented at The 14th International Management Conference*. Bucharest, Romania. DOI: 10.24818/IMC/2020/05.13
- Vică, C., Voinea, C., & Uszkai, R. (2021). The emperor is naked: Moral diplomacies and the ethics of AI. *Információs Társadalom*, 21(2), 83–96. <https://dx.doi.org/10.22503/inftars.XXI.2021.2.6>
- Wu, D. (2022, November 30). Can police use robots to kill? San Francisco voted yes. *The Washington Post*. Retrieved December 1st, 2022, from <https://www.washingtonpost.com/nation/2022/11/30/san-francisco-police-robots-kill/>
- Yeung, K., Howes, A. & Pogrebna, G. (2019). AI Governance by Human Rights-Centred Design, Deliberation and Oversight: An End to Ethics Washing. *SSRN Scholarly*, Paper ID 3435011. <https://doi.org/10.2139/ssrn.3435011>
- Zhou, J., Chen, F., Berry, A., Reed, M., Zhang, S., & Savage, S. (2020). A Survey on Ethical Principles of AI and Implementations. *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 3010–3017. <https://doi.org/10.1109/SSCI47803.2020.9308437>