Exploring Lessons from Cell Technology Ethics in AI based Transportation

Andreia Martinho

Delft University of Technology
The Netherlands
a.m.martinho@tudelft.nl

Background of Research

Artificial Intelligence (AI) brought a whole new set of ethical issues to the field of Transportation. Traditionally the ethics debates in Transportation revolved around cost-benefit analysis, sustainability, or fairness in pricing (Van Wee 2011) but the advent of autonomous driving gave rise to complex ethical issues related to this AI-based technology.

The vehicle at the center of these discussions is a machine designed to provide conveyance on public streets, roads, and highways with conditional, high, or full driving automation (Taxonomy 2014). Such machine is commonly known as Autonomous Vehicle (AV) and is expected to yield a multitude of social benefits to traffic safety, mobility, and accessibility(Anderson, Nidhi et al. 2014, Fleetwood 2017, Milakis, Van Arem et al. 2017).

The AV is considered here as a particular AI System, i.e. a regularly interacting or interdependent group of units forming an integrated structure that employs AI in any of its forms (learning, planning, reasoning, natural language processing, perception) separately or combined to perform a function while continuously interacting with the environment yet with limited expansion (Backlund 2000), as it relies heavily on AI to perform the driving function. For instance learning algorithms use data collected by sensors to interpret traffic situations in real time and select the appropriate action for the AV. The ethical issues of autonomous driving overlap greatly with more general ethics considerations of AI but there are also issues particular to AVs, such as the ethics of mundane traffic situations (Himmelreich 2018, Borenstein, Herkert et al. 2019).

To address the novel complexity in the field of Transportation I propose to seek guidance from Bioethics, which has a long-established tradition in dealing with convoluted ethical issues. Particularly I will focus on three elements of cell technology ethics ((i) The Ethics Process; (ii) Extreme Cases of Playing God; and (iii) Moratoriums). Our aim is to build on the lessons from cell technology ethics to frame and address relevant research questions related to AVs.

The Ethics Process: How to Mitigate the Risk of Speculative Ethics?

The Ethics process with respect to a particular disruptive technology is conceptualized as consisting of three major phases and for reference I use *In-vitro fertilization* (IVF) to illustrate this process. The first phase consists of initial research and development of the technology (proto-technology). For instance, even though the first baby conceived through IVF was born in 1978 the seminal work leading to this landmark event was published more than a decade earlier (Edwards, Donahue et al. 1966). It follows a phase of ethics debates about conflicting moral values, societal implications, and religious considerations associated with that technology. As a result of societal adjustment to a new technology and great scientific uncertainty (at this phase the technology is not yet fully developed), there is a substantial risk for speculative ethics. Illustratively, in the ethics literature of the 1970s, there are references to IVF as *ethically questionable biomedical research* and *unethical experiments on the unborn* which would lead to *unacceptable genetic manipulation* (Kass 1971, Cohen 1978). And finally a third phase of this process entails policy-making and regulation of the technology. Illustratively, in the Netherlands IVF technology was regulated under the Hospitals Act in the late eighties(Bos 1994).

A similar process is observed in the ethics of AV technology. The onset of autonomous driving provided fertile ground for ethics discussions in recent years but these debates have mainly revolved around different dimensions of the AV social dilemma, i.e., extreme situations in which the AV is required to make a difficult moral choice (Foot 1967, Thomson 1985, Bonnefon, Shariff et al. 2016). The overstatement of the AV social dilemma is problematic as it steers the attention from more realistic and pressing ethical issues regarding this technology.

In order to mitigate the risk for speculative ethics it is relevant to have a clear understanding of the ethical issues associated with this technology. For this purpose I relied heavily on a list of ethical issues in published guidelines for AI (Hagendorff 2019) later adjusted for AV technology and defined the following ethical issues: privacy; accountability, responsibility, liability; fairness, non-discrimination, justice; transparency; safety and cybersecurity; common good and sustainability; explainability and interpretability; human oversight, control, auditing; dual use problem and military; solidarity, inclusion, social cohesion; science-policy link; diversity; future of employment; public awareness and education about AVs and its risks; human autonomy; protection of whistleblowers; and hidden costs. Subsequently I proceeded with a pragmatic conceptualization of these ethi-

cal issues based on a literature review resulting from standard online searches on Google Scholar, Scopus, and Web of Science. It is expected that the list and conceptualization of the ethical issues of AVs guides researchers, industry, and policy-makers to a meaningful and non-speculative public debate on the ethics of autonomous driving.

Extreme Cases of Playing God: Is the AV Social Dilemma Relevant?

In the domain of cell technology there are extreme cases, such as somatic cell nuclear transfer (Wilmut, Beaujean et al. 2002), induced pluripotent stem cells (Takahashi and Yamanaka 2006), and clustered regularly interspaced short palindromic repeats (CRISPR) (Jinek, Chylinski et al. 2012), which metaphorically resemble Science playing God in the lab.

The AV technology has also brought up an extreme case to the Transportation domain. Because AVs will need to be programmed in advance to make decisions in collision situations there are heated debates about the underlying ethics concerning the distribution of harms in such extreme situations (Bonnefon, Shariff et al. 2016, Awad, Dsouza et al. 2018, Himmelreich 2018, Wolkenstein 2018). Concerns have been raised however about an overstatement of this social dilemma which could eventually cause a delay in the fruition of the social benefits of AVs(Wolkenstein 2018).

Here I investigate whether the AV Social Dilemma is relevant. I first looked into the literature and found contradictory arguments. Some authors consider that these extreme situations will take place in the future (Bonnefon, Shariff et al. 2016, Awad, Dsouza et al. 2018)while others consider such scenarios unrealistic (Goodall 2014, Himmelreich 2018). Subsequently I conducted a qualitative study in which I held thirteen semi-structured interviews with AV researchers working in academia, industry, and policy in The Netherlands. Once again the statements about the relevance of the AV social dilemma were contradictory as some interviewees consider it somewhat realistic (*the AV will always try to brake but in some situations there is no emergency braking time*), others consider it unrealistic (*Trolley problems are totally unrealistic…there is no such problem as AVs are programmed to always stop*), and one interviewee declared such problem as unavoidable. Finally I looked into the industry reports of companies with AV testing permit in California to provide insights that may alleviate the speculative nature of the social dilemma discussions. In this document review I did not find any reference to social dilemma issues as described in the scientific literature but nuanced yet relevant allusions to this issue were identified. More transparency is required from companies operating in the AV market.

Moratoriums: Should a Moratorium be Called for AV Technology?

When there is uncertainty about whether the risks of a disruptive technology outweigh its benefits or about its long-term effects often scientists call for moratoriums (Capron 2001). This was the case in 2015 when scientists called for a moratorium on clinical use of CRISPR, the gene editing tool mentioned above (McCarthy 2015).

There is still great uncertainty regarding the risks and benefits of AVs (Milakis, Van Arem et al. 2017) particularly with respect to convoluted collision situations (social dilemmas). In the semi-structured interviews study that I conducted in The Netherlands scientists emphasized that Ethics and Policy communities need time to adjust to AV technology and suggested weak versions of moratoriums such as skipping the deployment of levels 3 and 4 AVs which they consider more dangerous than fully autonomous level 5 AVs; limiting the scope of AVs for delivery of goods; and limiting the operational design domain.

Interestingly a key to avoid technology moratoriums is to focus on pressing and realistic ethical issues rather than speculative outlier cases (Mnyusiwalla, Daar et al. 2003, Nordmann and Rip 2009). Given the pragmatic nature of this work as a list of ethical issues and corresponding conceptualizations is offered as well as relevant empirical insights concerning the social dilemma, this research is expected to contribute to the mitigation of the risk of speculative AV ethics.

Outlook: AV as the Archetypal AI System

There are important lessons to be learned from Bioethics when it comes to complex ethical issues. Autonomous driving brought a novel complexity to the field of Transportation and we suggest seeking guidance from cell technology ethics when addressing AV ethics. Whereas this research is centered around autonomous driving technology, an AV is a particular AI System and as such the findings reported herein may be relevant for other AI Systems.

References

Anderson, J. M., K. Nidhi, K. D. Stanley, P. Sorensen, C. Samaras and O. A. Oluwatola (2014). Autonomous vehicle technology: A guide for policymakers, Rand Corporation.

Awad, E., S. Dsouza, R. Kim, J. Schulz, J. Henrich, A. Shariff, J.-F. Bonnefon and I. Rahwan (2018). "The moral machine experiment." Nature **563**(7729): 59.

Backlund, A. (2000). "The definition of system." Kybernetes **29**(4): 444-451.

Bonnefon, J.-F., A. Shariff and I. Rahwan (2016). "The social dilemma of autonomous vehicles." Science **352**(6293): 1573-1576.

Borenstein, J., J. R. Herkert and K. W. Miller (2019). "Self-driving cars and engineering ethics: the need for a system level analysis." Science and engineering ethics **25**(2): 383-398.

Bos, M. (1994). "Health care technology in the Netherlands." Health Policy **30**(1-3): 207-255.

Capron, A. M. (2001). "Placing a moratorium on research cloning to ensure effective control over reproductive cloning." Hastings LJ **53**: 1057.
Cohen, M. E. (1978). "The Brave New Baby and the Law: Fashioning Remedies for the Victims of In Vitro Fertilization." Am. JL & Med. **4**: 319.
Edwards, R. G., R. P. Donahue, T. A. Baramki and H. W. Jones Jr (1966). "Preliminary attempts to fertilize human oocytes matured in vitro." American Journal of Obstetrics and Gynecology **96**(2): 192-200.
Fleetwood, J. (2017). "Public health, ethics, and autonomous vehicles." American journal of public health **107**(4): 532-537.
Foot, P. (1967). "The problem of abortion and the doctrine of double effect."
Goodall, N. J. (2014). Machine ethics and automated vehicles. Road vehicle automation, Springer**:** 93-102.
Hagendorff, T. (2019). "The Ethics of AI Ethics--An Evaluation of Guidelines." arXiv preprint arXiv:1903.03425.
Himmelreich, J. (2018). "Never mind the trolley: The ethics of autonomous vehicles in mundane situations." Ethical Theory and Moral Practice **21**(3): 669-684.
Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). "A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity." science **337**(6096): 816-821.
Kass, L. R. (1971). "Babies by means of in vitro fertilization: unethical experiments on the unborn?" New England Journal of Medicine **285**(21): 1174-1179.
McCarthy, M. (2015). Scientists call for moratorium on clinical use of human germline editing, British Medical Journal Publishing Group.
Milakis, D., B. Van Arem and B. Van Wee (2017). "Policy and society related implications of automated driving: A review of literature and directions for future research." Journal of Intelligent Transportation Systems **21**(4): 324-348.
Mnyusiwalla, A., A. S. Daar and P. A. Singer (2003). "'Mind the gap': science and ethics in nanotechnology." Nanotechnology **14**(3): R9.
Nordmann, A. and A. Rip (2009). "Mind the gap revisited." Nature nanotechnology **4**(5): 273.
Takahashi, K. and S. Yamanaka (2006). "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." cell **126**(4): 663-676.
Taxonomy, S. (2014). "Definitions for terms related to on-road motor vehicle automated driving systems." Society of Automotive Engineers (SAE): Troy, MI, USA.
Thomson, J. J. (1985). "The trolley problem." The Yale Law Journal **94**(6): 1395-1415.
Van Wee, B. (2011). Transport and ethics: ethics and the evaluation of transport policies and projects, Edward Elgar Publishing.
Wilmut, I., N. Beaujean, P. De Sousa, A. Dinnyes, T. King, L. Paterson, D. Wells and L. Young (2002). "Somatic cell nuclear transfer." Nature **419**(6907): 583.
Wolkenstein, A. (2018). "What has the Trolley Dilemma ever done for us (and what will it do in the future)? On some recent debates about the ethics of self-driving cars." Ethics and Information Technology **20**(3): 163-173.