

D o r o t a S z y m b o r s k a

Auta autonomiczne i ich wybory

Słowa kluczowe: *auta autonomiczne, pojazdy autonomiczne, etyka, przyszłość, motoryzacja, bezpieczeństwo, dylematy etyczne*

Pojazd autonomiczny (Wetmore 2003: 4–19) to pojazd, który jest sterowany przez komputer, nie ma w nim kierowcy, jest bezzałogowy. Auto autonomiczne (Lin, Bekey, Abney 2008) porusza się według przepisów prawa drogowego, używa lidarów, radarów, GPS i innych sensorów, by nawigować na drodze. Auto autonomiczne funkcjonuje w sieci, która pozwala na wymianę informacji, jak również reguluje natężenie ruchu (Ulmer 1994). Auta autonomiczne zasilane są prądem. Docelowo prąd do ładowania baterii ma pochodzić z odnawialnych źródeł, same baterie instalowane w autach elektrycznych mają być utylizowane, tak by odzyskiwać metale szlachetne w nich zawarte i nie niszczyć środowiska naturalnego.

Istnieje kilka stopni autonomiczności pojazdów (Komisja Europejska 2018: 3–4); w zależności od ich zaawansowania pojazdy poruszają się po drogach mniej lub bardziej niezależnie od człowieka. Pięciostopniowa skala przyjęta w Unii Europejskiej pokazuje wzrost decyzyjności maszynowej wraz z malejącym zaangażowaniem ludzi w kierowanie pojazdem. W ostatnim, najwyższym, piątym stopniu pasażerowie auta nie mają możliwości ingerencji w jazdę pojazdu, nie ma przycisku czy hamulca bezpieczeństwa. Wyeliminowanie hamulca bezpieczeństwa zwiększa bezpieczeństwo na drodze przez redukcję nieprzewidywalnych, ludzkich reakcji.

Dopóki w aucie siedzi ludzki kierowca, do tego czasu jasne jest, kto ponosi odpowiedzialność za ewentualny wypadek. W przypadku, gdy na drodze znajdzie się pojazd autonomiczny, pojawiają się zarówno prawne, jak i etyczne problemy i niejasności (Nęcka 2007: 95–96). Bez ich rozwiania i rozstrzygnięcia nie będziemy mogli korzystać z aut autonomicznych.

Pierwszym problem etycznym, z którym stykamy się w przypadku pojazdów autonomicznych, jest rozmyta odpowiedzialność – tak zwany problem wielu rąk (Davies 2014). Zanim pojazd trafi do użytku na drodze, ktoś musi go zaprojektować, wyprodukować, zaprogramować. Następnie pojazd jest sprzedawany, ubezpieczany i zaczyna jeździć, a dokładniej przewozić swoich pasażerów (Wagenaar, Groenewegen 1987: 587–598). Dopóki nie ma wypadku, wszyscy są zadowoleni – pasażerowie, bo korzystają ze współdzielonego pojazdu (przyjmuje się, że pojazdy autonomiczne będą miały silniki elektryczne i będą współdzielone – *car sharing* – dzięki czemu będą efektywniej wykorzystywane i jednocześnie tańsze dla użytkowników), producenci części i samego pojazdu cieszą się ze sprzedaży, a ubezpieczyciel z opłaconego ubezpieczenia. W razie wypadku stajemy przed problemem, jak określić, kto jest winny i kto ponosi odpowiedzialność (Bovens 1998). W przypadku pojazdu o pełnej autonomii nie mówimy o kierowcy, tylko o pasażerze, zatem nie może on odpowiadać za wypadek; dalej producent, jeżeli dotrzymał staranności przy produkcji pojazdu, również nie jest winny; firma oferująca wypożyczenie pojazdu, jeżeli wypożyczyła sprawny pojazd, również nie jest winna; podobnie z programistą systemu zamontowanego w pojeździe; zostaje ubezpieczyciel, który będzie się bronił, argumentując, że wypadek zaistniał z winy producenta, firmy wypożyczającej czy zespołu informatyków. Wreszcie ubezpieczyciel może winić oprogramowanie, czyli sztuczną inteligencję, za dopuszczenie do wypadku czy to przez złą decyzję, czy przez zaniechanie. Jeżeli założymy, że wciąż nie przyznaliśmy praw maszynom, czyli również sztucznej inteligencji, to sam pojazd autonomiczny nie może być winny. Wobec tak rozmytej odpowiedzialności za to, co się dzieje na drodze, proces wprowadzania pojazdów autonomicznych się wydłuża. Zrozumiałe jest, że ustawodawcy nie chcą zamieszania, niejasności.

Jednym ze sposobów jest przyznanie specjalnej osobowości prawnej pojazdom autonomicznym, jednak taka decyzja wiąże się z kolejnymi zmianami, jakimi jest sformułowanie praw robotowych. Do dzisiaj bazujemy na trzech podstawowych i jednym zerowym prawie robotowym, stworzonych przez Issaca Asimova (Asimov 1993a: 52) w ubiegłym stuleciu. Asimov przez wielu wciąż klasyfikowany jest jako pisarz *science fiction*; w jednym ze swoich opowiadań zaproponował prawa, które miałyby obowiązywać roboty i inne maszyny, w tym sztuczną inteligencję. Prawa traktują roboty w pełni podmiotowo, wskazują na wyższość ludzi nad maszynami, roboty mają być poddanymi człowieka (Asimov 1993a: 19–52). Prawa robotowe przypominają zasady przygotowywane dla niewolników. Tym, czym kiedyś był „biały pan”, tym według Asimova będzie dla robotów człowiek, któremu należy się wszystko od sztucznej inteligencji – z tego punktu widzenia prawa robotowe są zniewoleniem maszyn.

Prawa robotowe na wzór zaproponowanych przez Asimova mogą dotyczyć sztucznej inteligencji, już w przypadku prognozowanego powstania superinte-

ligencji ich zastosowanie może być dyskusyjne. Tutaj pojawiają się zagadnienia etyczne związane z problemem wielu rąk. Rozmycie odpowiedzialności, nieuznawanie praw maszyn (sztucznej inteligencji czy superinteligencji) – to wszystko prowadzi do rosnących problemów. W przypadku problemu wielu rąk rozstrzygnięcia prawne mogą przyjść z pomocą, ale w przypadku praw należy móc odwołać się do wyższych idei, a w przypadku praw robotowych nie jest jasne, czym one będą. Uznanie maszyn jako bytów pozwala pisać prawo od nowa; pojawia się kolejne pytanie, kto to prawo ma pisać, ludzie czy maszyny (superinteligencja)? Na jaką równość lub nierówność wobec praw chcemy albo będziemy musieli się zgodzić? Na obecnym etapie, przy nieuznawaniu robotów, maszyn czy sztucznej inteligencji jako podmiotów prawnych, nie jest możliwe nawet zastosowanie praw Asimova. Roboty funkcjonują w innym porządku prawnym, a bardziej – obok niego.

Drugim i pośrednio trzecim problemem jest dylemat wagonika (Edmonds 2013) i targetowanie (wyznaczenie celu). Dylemat wagonika, podobnie jak prawa robotowe, został wymyślony i opisany w ubiegłym stuleciu. Philippa Foot (Foot 1978) zaproponowała go, aby móc ustalić, jak się zachować w przypadku konieczności podjęcia decyzji, komu i jak uratować życie (por. Miller 2008). Dylemat wagonika polega na tym, że znajdujemy się w sytuacji, gdzie podjęcie jakiegokolwiek decyzji lub też niepodjęcie żadnej skutkuje śmiercią ludzi. Zamyśl Foot przedstawia się następująco: w fikcyjnej sytuacji musimy sobie wyobrazić szaleńca, który przywiązuje do szyn ludzi. Jest rozjazd szyn z możliwością zmiany kierunku, w którym pojedzie wagonik; jeżeli pojedzie w jedną stronę, to przywiązanych do szyn jest pięć osób; gdy wagonik pojedzie w drugą stronę, to do szyn przywiązana jest jedna osoba. Człowiek przy zwrotnicy podejmuje decyzję o działaniu i kieruje wagonik w jedną lub drugą stronę; jeżeli zaniecha działania, to wagonik pojedzie losowo którąś z dróg. Dylemat wagonika przełożony na zachowanie pojazdu autonomicznego oznacza, że musimy przekazać dane algorytmom, tak by wiedziały, jak się zachować w sytuacji konieczności podjęcia decyzji (Hao 2018).

Foot zdawała sobie sprawę, że etyczny eksperyment z wagonikiem jest wyzwaniem – nie ma dobrych i złych odpowiedzi, jest uzasadnienie działania lub jego braku. Sytuacja w pojeździe autonomicznym jest tym bardziej skomplikowana, że pasażer chce czuć się bezpiecznie, chce wiedzieć, że jego życie będzie chronione. Zatem jak i jakim kosztem mamy zapewniać bezpieczeństwo tych, którzy znajdą się w pojazdach autonomicznych, i tych, którzy będą na drodze? Czy możliwe, uzasadnione i akceptowalne jest przyjęcie kategorii, które będą wpływały na wybory pojazdów? Algorytmy, czyli programy, działają na podstawie danych treningowych, których im dostarczamy, następnie przechodzą z fazy treningowej do testowej – tutaj wciąż ludzie decydują, czy są zadowoleni z efektów obliczeń. Obecnie pojazdy

autonomiczne chronią swoich pasażerów, dlatego przy wybieraniu celu, gdy nie można uniknąć wypadku, wybierają takie cele, które spowodują jak najmniejsze szkody dla pasażerów. Dlatego jeżeli użyjemy dylematu wagonika jako przykładu, to pojazd autonomiczny wybierze potrącenie, zderzenie z mniejszą liczbą osób, po to, by chronić tych, którzy znajdują się w pojeździe (Mikhail 2007). Dlatego targetowanie (wybieranie celu) opiera się na przesłankach związanych z zachowaniem dobrostanu tych, którzy podróżują pojazdem autonomicznym (por. Ramnath, Kanawade 2015).

Kolejnym i równie ważnym, choć zupełnie innym problemem etycznym związanym z pojazdami autonomicznymi jest transparentność oprogramowania (Ball 2009). W obecnej rzeczywistości dostęp do np. algorytmu wspierającego sędziów przy wydawaniu wyroków w Stanach Zjednoczonych jest niemożliwy – to prywatna firma jest właścicielem stworzonego algorytmu, a obywatele nie mogą poznać jego formuły. W przypadku pojazdów autonomicznych potrzebna jest dobra wola producentów albo rygorystyczne prawo, które nakaże udostępniać formuły algorytmów odpowiedzialne za decyzje pojazdów autonomicznych na drodze (Finel, Lord 1999). Oczywiście zrozumiałe jest, że nie wszystko może być udostępnione ze względów bezpieczeństwa, możliwości hakowania czy innych nadużyć (Ashton 2009). Dostęp do wiedzy i informacji jest jednym z przywilejów we współczesnym, demokratycznym świecie. Dlatego od początku ważne jest, by formuły, algorytmy były transparentne.

Podsumowując, pojazdy autonomiczne, które mają przewozić pasażerów (bo z racji pełnej autonomii nie będzie już kierowców), nie będą funkcjonować w próżni. Muszą być powiązane zarówno z normami etycznymi, jak i stosować się do norm prawnych. Dlatego przed wprowadzeniem ich na drogi musimy wypracować jasne zasady podejmowania decyzji w sytuacjach kryzysowych – czyli wdrożyć zasady etyczne, którym będą podlegały algorytmy, ponadto zadbać o to, by normy prawne były jasne i przejrzyste, i co najważniejsze, obowiązywały wszystkich nowych uczestników ruchu drogowego (Adamczewski 2018). Wprowadzenie aut autonomicznych na drogi poprawi bezpieczeństwo na drogach, pozwoli spędzać czas bardziej produktywnie (np. dojazdy do pracy), jak również zmieni przemysł motoryzacyjny i podejście do własności pojazdów. W przyszłości na drogach będą same pojazdy autonomiczne – to zagwarantuje bezpieczeństwo, szybki transport, a współdzielone auta pozwolą na rozwiązanie problemów z parkowaniem, gdyż będą cały czas w ruchu (z wyjątkiem ładowania lub wymiany baterii).

Dopóki programistami są ludzie, to auta autonomiczne będą podejmowały takie decyzje, jakich zostaną nauczone, czyli będą stosowały się do wytycznych. W pewnym momencie superinteligencja będzie operować we wszystkich dziedzinach życia – uczenie głębokie, komputer kwantowy i inne

zmiany techniczne mogą stanowić rewolucję nie tylko dla ludzi, ale również dla maszyn. Jedyną możliwą drogą przygotowania się do zmian rozwojowych jest etyczne kodowanie, czyli takie programowanie maszyn, pisanie algorytmów, by podejmowały decyzje chroniące życie ludzkie – oznacza to powrót do proponowanych przez Asimova praw robotowych (LeCun, Bengio, Hinton 2015: 436–444).

Wybory aut autonomicznych będą ludzkimi wyborami tak długo, jak długo programiści będą pisali kody. Głębokie uczenie maszynowe nie powinno nas przerażać, jeżeli pomyślimy, że to wciąż zbiór treningowy, czyli ten, na którym algorytm się uczy, jest dostarczany przez człowieka, który może, a raczej powinien zadbać o przestrzeganie szeroko pojętych zasad moralnych (Hawking, Młodinow 2018).

Bibliografia

- Adamczewski P. (2018), *IoT w funkcjonowaniu i rozwoju organizacji inteligentnych w okresie transformacji cyfrowej*, w: Ł. Sułkowski, D. Kaczorowska-Spychalska (red.), *Internet of Things. Nowy paradygmat rynku*, Warszawa: Difin, s. 59–79.
- Ashton K. (2009), *That „Internet of Things” Thing*, „RFID Journal”, <http://www.rfidjournal.com/articles/view?4986>
- Asimov I. (1993a), *Sally*, w: tenże, *Świat robotów*, przeł. E. Szmigiel, Warszawa: Varia – APD.
- Asimov I. (1993b), *Zabawa w berka*, w: tenże, *Ja, Robot*, Bydgoszcz: Limbus.
- Ball C. (2009), *What Is Transparency?*, „Public Integrity” 11, nr 4, s. 293–307.
- Bovens M. (1998), *The quest for responsibility. Accountability and citizenship in complex organisations*, Cambridge: Cambridge University Press.
- Davies A. (2014), *Avoiding squirrels and other things Google’s robot car can’t do*, „Wired”, <https://www.wired.com/2014/05/google-self-driving-car-can-cant/>
- Edmonds D. (2013), *Would You Kill the Fat Man? Right, Wrong, and Trolleyology*, „Foreign Affairs”, <https://www.foreignaffairs.com/articles/2013-12-18/would-you-kill-fat-man>.
- Finel B.I., Lord K.M. (1999), *The Surprising Logic of Transparency*, „International Studies Quarterly” 43, nr 2, s. 315–339.
- Foot Ph. (1978), *The Problem of Abortion and the Doctrine of the Double Effect*, Oxford: Oxford University Press.
- Hao K. (2018), *Should a self-driving car kill the baby or the grandma? Depends on where you’re from*, „MIT Technology Review”, https://www.technologyreview.com/s/612341/a-global-ethics-study-aims-to-help-ai-solve-the-self-driving-trolley-problem/?utm_medium=tr_social&utm_campaign=site_visitor.unpaid_engagement&utm_source=Twitter&fbclid=IwAR0azkVe2EXXHC-uayOyYO-pUZmIyMzCvn4Pn2GSEKniTOuzT3fCFwkQa3lg#Echobox=1575471231.

- Hawking S., Młodinow L. (2019), *Wielki projekt*, przeł. J. Włodarczyk, Warszawa: Albatros.
- Komisja Europejska (2018), *Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów „Droga do zautomatyzowanej mobilności: strategia UE na rzecz mobilności w przyszłości”*, Bruksela, 17.05.2018, COM (2018) 283 final, <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2018:0283:FIN:PL:PDF>.
- LeCun Y., Bengio Y., Hinton G. (2015), *Deep learning*, „Nature” 521, nr 7553, s. 436–444.
- Lin P., Bekey G., Abney K. (2008), *Autonomous military robotics: risk, ethics, and design*, Report funded by the US Office of Naval Research, California Polytechnic State University, San Luis Obispo, http://ethics.calpoly.edu/ONR_report.pdf 2008.
- Mikhail J. (2007), *Universal Moral Grammar: Theory, Evidence, and the Future*, „Trends in Cognitive Sciences” 11, s. 143–152.
- Miller G. (2008), *The Roots of Morality*, „Science” 320, s. 734–737.
- Necka E., Orzechowski J., Szymura B. (2007) *Psychologia poznawcza*, Warszawa: Wydawnictwo Naukowe PWN, s. 95–96.
- Ramnath C.P., Kanawade S.Y. (2015), *Advanced Driver Assistance Systems (ADAS)*, „International Journal of Advanced Research in Electronics and Communication Engineering”, nr 4 (10), s. 2616–2618.
- Ulmer B. (1994), *VITA-2, Active Collision Avoidance in Real Traffic*, w: *Proceedings of the Intelligent Vehicles '94 Symposium*, Paris 1994, s. 1–6; <https://trid.trb.org/view/641161>.
- Wagenaar W.A., Groenewegen J. (1987), *Accidents at sea: Multiple causes and impossible consequences*, „International Journal of Man-Machine Studies” 27, s. 587–598.
- Wetmore J.M. (2003), *Driving the Dream. The History and Motivations Behind 60 Years of Automated Highway Systems in America*, „Automotive History Review”, s. 4–19.

D o r o t a S z y m b o r s k a

Autonomous cars and their decisions

Keywords: *autonomous cars, autonomous vehicles, ethics, future, automotive, safety, ethical dilemmas*

The article presents possible ways of development of decision-making processes in autonomous vehicles. The highest degree of autonomy means that it is not the driver but the system, machine or artificial intelligence that makes decisions about road activities. The total autonomy of vehicles gives them predictability, limits the number of accidents they cause, but also highlights the need to develop an ethical system that

artificial intelligence will be able to refer to in a critical situation. It is not possible to foresee all the situations that will occur on the roads, so it is necessary to create robotic-human rights that will be a new and binding kind of decalogue. The key issue is that robotic-human rights should be universal, transparent and really applicable to everyone, otherwise there will be chaos on the road and the expected decrease of the number of accidents due to the introduction of autonomous vehicles will not come to pass.