

Krzysztof Michalski
Politechnika Rzeszowska

Autonomizacja techniki i niepożądane skutki eliminowania człowieka jako źródła błędów

Wprowadzenie

Autonomizacja techniki i zastępowanie człowieka przez inteligentne maszyny jest jednym z wariantów zapewniania niezawodności systemów technicznych. Pod tym względem autonomiczne systemy techniczne konkurują m.in. z *human enhancement*¹ i neuroinżynierią². Eliminowanie człowieka z systemów technicznych jest

¹ Od połowy lat 90. XX w. trwa międzynarodowa debata wokół idei udoskonalenia człowieka dzięki zdobyczom współczesnej nauki i technologii, szczególnie postępowi w inżynierii genetycznej i biotechnologii, farmakologii i chirurgii, ale także neuroinżynierii. Nowe technologiczne możliwości modyfikacji organicznych i psychicznych potencjałów wydajnościowych człowieka sięgają od chirurgii estetycznej, poprzez doping w sporcie, wspomaganie psychofarmakologiczne (np. popularne „dopalacze”), terapię syndromu ADHD u dzieci, wydłużanie życia i spowalnianie procesów starzenia się, aż po neuroprotezy i głęboką elektryczną lub magnetyczną neurostymulację, a także futurystyczne projekty radykalnych wariantów „human enhancement” oparte na kognitywistyce i konwergencji nanotechnologii, bioinżynierii, cybernetyki i telekomunikacji, mające udoskonalić potencjały intelektualne człowieka, jego zdolności poznawcze, pamięć, inteligencję, kreatywność, wyobraźnię, ale także pewne cechy psychiczne (np. opanowanie i odporność na stres, motywację itp.), a nawet kompetencje społeczne (umiejętności komunikowania się, siłę perswazji, zdolności przywódcze itp.). Enhancement tym różni się od działań terapeutycznych, że interwencjom poddawani są zupełnie zdrowi ludzie. Trudno jednak uznać takie kryterium rozróżnienia za jednoznaczne ze względu na wielość pojęć zdrowia w samej medycynie. Enhancement jest źródłem wielu społecznych i etycznych kontrowersji, w tym także kontrowersji znanych z debat dotyczących eugeniki. Por. M.C. Roco, W.S. Bainbridge (eds.), *Converging Technologies for improving human performance. Nanotechnology, Biotechnology, Information Technology and Cognitive Science*, Kluwer Academic Publishers, Dordrecht 2003; B. Schöne-Seifert, D. Talbot (Hrsg.), *Enhancement. Die ethische Debatte*, Paderborn 2009.

² Terminu „neuroinżynieria” użyto w znaczeniu możliwie szerokim na określenie ogółu procedur technicznych łączących metody neurobiologii, informatyki i nauk technicznych, mających na celu poszerzenie możliwości zewnętrznego oddziaływania na ludzi w zakresie samoświadomości, doświadczania świata, wartościowania i emocjonalności oraz innych aspektów związanych z indywidualnym trybem życia. Neuroinżynieria obejmuje zarówno zastosowania inwazyjne głęboko penetrujące (np. BMI – interfejs mózg-maszyna lub BCI – interfejs mózg-komputer stosowane w sensorycznej lub motorycznej neuroprotezyce i implantologii, wymagającej głębokiej stymulacji mózgu lub rdzenia kręgowego, jak w przypadku implantów ślimakowych przywracających osobom niesłyszącym zmysł słuchu czy neuroprotezy motorycznych pierwszej generacji umożliwiających oso-

również coraz częściej podyktowane względami ekonomicznymi. Dynamiczny rozwój zaawansowanych technologii, wzrost popytu na inteligentne rozwiązania oraz konkurencja wśród dostawców przyczyniają się do spadku cen inteligentnych maszyn do takich poziomów, przy których człowiekowi coraz trudniej konkurować z technicznymi substytutami w aspekcie kosztowo-wydajnościowym. Zastępowanie człowieka tańszymi i wydajniejszymi inteligentnymi maszynami stało się dla krajów wysoko uprzemysłowionych szansą na przeciwdziałanie powszechnemu w globalnym biznesie zjawisku offshoringu oraz związanemu z nim odpływowi kapitału, przemieszczaniu miejsc pracy do krajów biednych oraz utratą wpływów z podatków.

Jednak podyktowane względami ekonomicznymi i wzrastającymi wymogami bezpieczeństwa procesy autonomizacji techniki z wielu ważnych powodów wymagają poddania większej społecznej kontroli niż inne procesy rozwojowe we współczesnym technicznie zainstrumentowanym świecie. Inteligentniejsze i bardziej autonomiczne systemy techniczne mają większe zdolności do „życia własnym życiem”, grożą nieodwracalną utratą kontroli, a nawet instrumentalizacją człowieka. Dlatego niepożądane skutki uboczne, paradoksy i pułapki, zagrożenia i potencjały konfliktogenne procesów skutkujących wyrugowaniem człowieka z wielu obszarów jego dotychczasowej działalności należy rozpoznać odpowiednio wcześniej, zanim procesy autonomizacji staną się faktycznie nieodwracalne.

Główne obszary zastosowań autonomicznych systemów technicznych

Automatyka przemysłowa i roboty fabryczne

Zgodnie z międzynarodowymi standardami normalizacyjnymi robotem jest każdy uniwersalny wielofunkcyjny, wielozadaniowy manipulator wyposażony w co najmniej trzy niezależne osie i nadający się do dowolnego wielokrotnego za-

bom sparaliżowanym uzyskanie kontroli nad ramieniem za pomocą mikroelektrod zaimplantowanych w korze mózgowej), jak i zastosowania nieinwazyjne, oparte na elektroencefalograficznym obrazowaniu aktywności neuronalnej u osób niezdolnych do komunikowania się z własnym ciałem lub otoczeniem zwyczajnymi kanałami (np. przezczaszkowa stymulacja magnetyczna, przezczaszkowa bezpośrednia stymulacja stałoprądowa, metody diagnostyki obrazowej takie jak MRT, PET, CT oraz metody analizy procesów korowych oparte na elektro- lub magnetoencefalografii). Do nieinwazyjnych zastosowań neuroinżynierii zalicza się również tzw. neuroimaging, czyli techniki wizualizacji korelacji występujących między procesami neuronalnymi a zjawiskami mentalnymi. Aktualnie na gruncie neuroinżynierii poszukuje się głównie nowych rozwiązań umożliwiających kompensowanie dysfunkcji systemu nerwowego dzięki wykorzystaniu najnowszych zdobyczy mechatroniki w celach czysto terapeutycznych, a także wypróbowuje się możliwości głębokiej stymulacji mózgu w terapii chorób psychicznych, ciężkich depresji i leczeniu uzależnień, ale nie brakuje również pomysłów, aby z pomocą systemów interakcyjnych wywierać techniczny wpływ na procesy neuronalne w ludzkim mózgu w celach optymalizacyjnych (*Human Enhancement*). Szerzej na ten temat zob. M.R. Bennet, P.M.S. Hacker, *Philosophical Foundations of Neuroscience*, Malden, Oxford 2003; D. Sturma, *Neurotechniken*, [w:] A. Grunwald (Hrsg.), *Handbuch Technikethik*, Metzler, Stuttgart – Weimar 2013, s. 343-348.

programowania pod kątem poruszania przedmiotami po zaprogramowanych, zmiennych torach³.

Gwałtowny rozwój robotyki zapoczątkowało rozłożenie przemysłowych procesów wytwórczych na pojedyncze czynności robocze nadające się do programowania. Automatyzacja takich czynności w połączeniu z wykorzystaniem podajników taśmowych spowodowała w przemyśle przełom z punktu widzenia polityki wydajności i zapoczątkowała proces stopniowego eliminowania tzw. czynnika ludzkiego z przemysłowej wytwórczości.

Robotyka przemysłowa stała się obecnie tak powszechna, że już w roku 2009 szacowano całkowitą liczbę robotów przemysłowych na świecie na ponad milion⁴.

Niektóre nowoczesne roboty fabryczne są „autonomiczne”, ale jedynie w ściśle określonych warunkach. Ze względu na specyficzne środowisko pracy robotów produkcyjnych nie programuje się tych urządzeń z myślą o zbyt dużej swobodzie działania. Autonomia takich maszyn ograniczona jest przeważnie do samodzielnego lokalizowania przedmiotów obróbki, rozróżniania ich rodzajów i rozpoznawania własnych zadań, orientacji w przestrzeni, samoczynnego poruszania się, rozpoznawania sytuacji niebezpiecznych dla siebie i ludzi, a także uczenia się nowych metod wykonywania zadań i dostosowywania się do zmieniających się warunków otoczenia.

Roboty serwisowe

Robotami serwisowymi nazywa się roboty przeznaczone do zadań niezwiązanych bezpośrednio z przemysłowymi procesami produkcyjnymi. Wśród komercyjnych zastosowań urządzeń sprzedanych do roku 2010 dominowały rodzaje działalności związane z obronnością, ratownictwem i bezpieczeństwem (ok. trzydziestopięciodziesiąt procentowy udział) oraz rolnictwem (25%)⁵. Roboty serwisowe można zwykle scharakteryzować jako formę rozszerzenia ludzkich zdolności operacyjnych. Na przykład roboty dozorujskie wykorzystuje się do monitorowania terytoriów leżących poza zasięgiem ludzkich możliwości percepcyjnych (zbyt duży obszar, miejsce zbyt odległe, zasłonięte, spowite ciemnościami lub zbyt niebezpieczne dla człowieka itp.). Automatyczne dojarki są w stanie „obsłużyć” w tym samym czasie o wiele więcej krów niż systemy manualne. Roboty dostarczające (dostawcze) transportują przesyłki bez przerw potrzebnych człowiekowi na odpoczynek, a autonomiczne pojazdy obsługiwane przez roboty nie tylko pozwalają na osiągnięcie większej płynności ruchu przy wysokim wskaźniku bezkolizyjności, ale także umożliwiają podróżującym wykonywanie czynności, które byłyby nie do pogodzenia z wymaganiami związanymi z kierowaniem pojazdem (np. sen lub spożywanie alkoholu). Dzięki funkcjonalnościom robotów serwisowych udaje się nie tylko sporo zaoszczędzić

³ Por. ISO-Standard 8373, 1994.

⁴ Por. M. Decker, *Robotik*, [w:] A. Grunwald (Hrsg.), *Handbuch Technikethik*, Metzler, Stuttgart – Weimar 2013, s. 355.

⁵ Por. *ibidem*, s. 356.

zczędzić na tzw. kosztach osobowych, ale przede wszystkim znacząco zwiększyć wydajność usług.

Autonomiczne pojazdy i statki powietrzne (drony)

Autonomiczny samochód, znany również jako auto bez kierowcy, auto kierowane przez robota, to bezzałogowy pojazd naziemny – ostatni „krzyk mody” na rynku motoryzacyjnym. Osoby niezorientowane mogą być wprowadzane w błąd przez medialne rewelacje o pladze samochodów bez kierowców na drogach w USA i Europie Zachodniej, tymczasem żadne autonomiczne auto dopuszczone do tej pory do ruchu po drogach publicznych nie było w pełni autonomiczne i pozbawione ingerencji człowieka. Zawsze w pojeździe był człowiek w razie potrzeby gotowy do przejęcia kontroli nad pojazdem. Ze względów bezpieczeństwa pojazdy autonomiczne wyposażane są w systemy awaryjnego zatrzymania w razie wystąpienia zdarzeń nieprzewidzianych w programie pracy. W odróżnieniu od łazików terenowych i dronów autonomiczne pojazdy konstruowane z myślą o uczestnictwie w ruchu drogowym muszą być nie tylko wyposażone w skomplikowane systemy nawigacyjne umożliwiające orientację w terenie, ale także muszą mieć zdolność rozpoznawania oznakowania, która jest oparta m.in. na umiejętności tzw. arbitralnego dekodowania. Większości producentów aut nie zależy w pierwszej kolejności na stworzeniu całkowicie bezzałogowego auta seryjnego, a testy drogowe z autonomicznymi pojazdami służą głównie udoskonalaniu rozwiązań technicznych wspomagających ludzkiego kierowcę. Mimo to nie należy bagatelizować skutków społecznych – zarówno dobroczynnych, jak i złoczynnych – jakie mogą się pojawić z chwilą osiągnięcia przez te technologie odpowiedniej dojrzałości rynkowej. Analitycy szacują, że za ok. 15 lat samochody z autopilotem wejdą do produkcji seryjnej⁶.

Roboty społeczne

Robotami społecznymi nazywa się autonomiczne urządzenia projektowane pod kątem złożonych interakcji z człowiekiem w środowisku społecznym, imitujące wygląd i zachowania człowieka, wyposażone nie tylko w określone kompetencje operacyjne, ale także w kompetencje komunikacyjne⁷. Kompetencje komunikacyjne robotów społecznych obejmują przede wszystkim:

- (1) zdolność do rozpoznawania obecności człowieka;
- (2) zdolność do zachowań dotykowych;
- (3) zdolność do gestykulacji i zachowań ruchowych;
- (4) mowę i zdolność do konwersacji oraz

⁶ Por. M. Ramsey, *Self-Driving Cars Could Cut Down on Accidents, Study Says*, „The Wall Street Journal” 5.03.2015 (<https://www.wsj.com/articles/self-driving-cars-could-cut-down-on-accidents-study-says-1425567905>) [08.11.2017].

⁷ Por. S. Breazeal, A. Takanishi, T. Kobayashi, *Social Robots that Interact With People*, [w:] B. Siciliano, O. Khatib (eds.), Springer Handbook of Robotics, Springer – Verlag, Berlin – Heidelberg 2008, s. 1350.

(5) rozpoznawanie i wyrażanie emocji⁸.

Kompetencje komunikacyjne robotów umożliwiają wbudowane systemy sensoryczne imitujące zmysł wzroku, słuchu i dotyku, rozwiązania z zakresu sztucznej inteligencji, a także systemy komunikacyjne, takie jak wyświetlacze lub generatory mowy. Z punktu widzenia komunikacyjnego nie bez znaczenia są również humanoidalne kształty robotów, umożliwiające wykorzystanie wielu kanałów komunikacyjnych z obszaru mowy ciała, głównie sygnałów pozycyjnych, proksemiki, gestykulacji czy sygnałów dotykowych, których realizacja wymaga rozwiązań imitujących funkcje ludzkich kończyn. Roboty społeczne już dawno przestały być jedynie atrakcją targów przemysłowych i gadżetem dostarczającym rozrywki. Znajdują one obecnie coraz więcej sensownych praktycznych zastosowań w psychoterapii osób z zaburzeniami rozwojowymi (np. autyzmem), osób uzależnionych, w resocjalizacji, pielęgnacji osób starszych i chorych, asystując osobom z deprywacjami sensorycznymi (np. niewidomym).

Autonomia robotów – mity i rzeczywistość

Autonomia w przypadku urządzeń technicznych jest cechą względną i stopniowalną. Podobnie jak w przypadku osób ludzkich, wśród których nie brakuje jednostek o ograniczonej samoświadomości (noworodki, osoby z poważnymi uszkodzeniami mózgu itp.) czy jednostek o ograniczonych możliwościach samodzielnego poruszania się (osoby sparaliżowane), również roboty wyposażone są w różne stopnie i zakresy autonomii. Autonomiczność maszyn i systemów technicznych jest głównym wykładnikiem ich inteligencji. W przypadku robotów fabrycznych autonomia jest zwykle ograniczona. Istnieje możliwość wyposażania takich robotów w zdolności do samotrzymania ruchu dzięki systemom propriocepcyjnym umożliwiającym rozpoznanie własnego statusu (samoocena poprawności działania, monitoring stanu naładowania akumulatorów, czujniki termiczne, optyczne, dotykowe i elektryczne), a także systemom nawigacyjnym i napędem umożliwiającym samodzielne „żerowanie” (zdolność robotów do wyszukiwania i pobierania potrzebnych zasobów: energii poprzez wyszukiwanie stacji ładowania i dokowanie, a także części zamiennych). Wyższe wymagania dotyczące autonomii stawia się robotom serwisowym. Zautomatyzowane kosiarki do trawników mają zdolność dostosowywania swojego oprogramowania do tempa wzrostu trawy, a roboty odkurzające mają wbudowane detektory zanieczyszczeń i są w stanie dostosowywać czas pracy w określonym miejscu do ilości zabrudzeń. Jeszcze wyższy stopień autonomii jest związany z wykonywaniem zadań warunkowych. Roboty-strażnicy są wyposażane w zdolność do identyfikacji osób

⁸ Por. H. Li, J.J. Cabibihan, Y.K. Tan, *Towards an Effective Design of Social Robots*, “International Journal of Social Robotics”, 2011 (3), s. 333.

na podstawie analizy ich zachowań oraz reagowania w określony sposób w zależności od lokalizacji intruza. Nowoczesne roboty wyposażone w triangulacyjne systemy nawigacyjne same wytwarzają sobie mapy laserowe budynków, dzięki którym uzyskują zdolność do samodzielnego poruszania się po otwartych przestrzeniach i korytarzach oraz rozpoznawania i omijania przeszkód. Początkowo automatyczne systemy nawigacyjne opierały się na dalmierzach laserowych wykonujących pomiary w jednej płaszczyźnie, ale najnowsze systemy nawigacyjne posługują się już triangulacyjnymi systemami integrującymi sygnały pozycyjne i nawigacyjne z wielu czujników, samodzielnie eliminując informacje najmniej wiarygodne. Dzięki supernowoczesnym systemom nawigacyjnym i cyfrowym systemom kontroli dostępu, niewymagającym mechanicznego otwierania drzwi, takie inteligentne roboty wykorzystywane w systemach bezpieczeństwa obiektów mogą pod nieobecność ludzi swobodnie poruszać się po wszystkich kondygnacjach budynków, otwierać elektronicznie drzwi, uruchamiać windy itp. – to wszystko pomimo nadal zasadniczo nierozwiązanego problemu poruszania się robotów po schodach i wykonywania złożonych operacji manualnych. Autonomiczne pojazdy i drony są obecnie w stanie wykonywać skomplikowane zadania i misje bez jakichkolwiek ingerencji człowieka. Szczególnie autonomiczne pojazdy naziemne muszą mierzyć się z poważnymi trudnościami w poruszaniu się po terenie, związanymi m.in. z trójwymiarowością terenu, dysproporcjami w gęstości powierzchniowej czy też niekorzystnymi warunkami atmosferycznymi. Roboty-łaziki potrafią dokonywać oceny bezpieczeństwa obszarów znajdujących się w ich polu „widzenia”, wybierać na jej podstawie optymalną, najbardziej wydajną i bezpieczną drogę do miejsca przeznaczenia i podążać tą drogą, powtarzając procedurę w razie potrzeby.

Problemy na gruncie teorii pracy

Pomimo wysokiego stopnia zaawansowania technologicznego procesów produkcyjnych tylko nieliczne takie procesy udało się do dzisiaj w pełni zautomatyzować, głównie ze względów ekonomicznych. Większość współczesnych procesów wytwórczych opiera się więc w praktyce na różnych kombinacjach czynności wykonywanych przez ludzi i czynności wykonywanych przez roboty. Automatyzowane są przeważnie te czynności, których programowanie jest opłacalne, czyli przede wszystkim czynności powtarzalne, niezłożone, wymagające szybkich lub nadzwyczaj precyzyjnych operacji albo dużego nakładu siły (wysiłku fizycznego). Pozostałe czynności pozostają domeną człowieka, choć wraz z postępującym potaniem rozwiązań wynikającym z coraz większego upowszechnienia automatyki i robotyki domena człowieka w przemysłowych procesach wytwórczych będzie się prawdopodobnie coraz bardziej kurczyć. Daleko idące wyeliminowanie czynnika ludzkiego z procesów przemysłowych zakłada zresztą aktualnie forso-

wana koncepcja czwartej rewolucji przemysłowej (Przemysł 4.0)⁹. W tej sytuacji staje się zrozumiałe, dlaczego główna uwaga w publicznych debatach dotyczących automatyzacji i robotyki zogniskowana jest na zmianach, jakie rozwój i upowszechnianie tych technologii powoduje w świecie pracy.

Ponieważ już tylko niektóre ogniwa procesu produkcyjnego związane są z pojedynczymi czynnościami wykonywanymi przez człowieka, w odniesieniu do zmian w świecie pracy można zaobserwować dwa przeciwstawne efekty. Z jednej strony spada zapotrzebowanie na złożone kompetencje po stronie pracowników wymagające specjalistycznego wyższego wykształcenia, a wzrasta zapotrzebowanie na proste umiejętności, do których w zupełności wystarcza wykształcenie zawodowe na poziomie podstawowym (*down-skill-effect*). Z drugiej zaś strony wzrasta liczba i złożoność zadań związanych z technicznym dozorem nad pracą systemów automatycznych, co powoduje gwałtowny wzrost wymagań kompetencyjnych wobec operatorów takich systemów, których określa się jako „Click-Workers” (*up-skill-effect*)¹⁰. Postępujące wypieranie człowieka z procesów wytwórczych przez systemy automatyczne i roboty ma również inne, jeszcze bardziej paradoksalne konsekwencje z punktu widzenia teorii pracy, które Lisanne Bainbridge¹¹ określa mianem „ironii automatyzacji”. Odbierając człowiekowi

⁹ Pojęcie „Industry 4.0” wprowadził do debaty publicznej w roku 2011 Henning Kagermann, a do jego popularyzacji przyczyniły się w 2013 r. Międzynarodowe Targi w Hanowerze, które idee czwartej rewolucji przemysłowej uczyniły swoim hasłem przewodnim. Głównym przejawem „kwantowego skoku” w przemyśle jest inteligentne usieciwienie produkcji, dzięki któremu „inteligentne fabryki bez ludzi” we w pełni zautomatyzowanych, samoorganizujących się i samooptimalizujących się procesach cyberfizycznych (Cyber Physical Systems, CPS – złożone systemy z wbudowanym oprogramowaniem i elektroniką, które są połączone z zewnętrznym światem fizycznym za pośrednictwem sensorów pobierających informacje i udostępniających je usługom opartym na sieci oraz elementów napędu, które mogą bezpośrednio oddziaływać na procesy w świecie fizycznym) będą w stanie opanować każdą złożoność zadań związanych z wytwarzaniem zindywidualizowanych produktów przy możliwie minimalnym udziale człowieka. Przemysł maszynowy krajów Europy Zachodniej – z wielkimi koncernami motoryzacyjnymi na czele – podpadający pod wpływem przemian wymuszonych polityczno-ekonomiczną integracją świata i procesami globalizacyjnymi, odczłowieczył wizję przyszłości przemysłu, w której przedsiębiorstwa, wykorzystując nowoczesne technologie informacyjno-komunikacyjne, połączą swoje maszyny, systemy magazynowe, zasoby informacyjne i kompetencyjne oraz środki utrzymania ruchu w globalną sieć CPS. W jej ramach jednostki produkcyjne będą samodzielnie wymieniały się informacjami, inicjowały działania i same sobą wzajemnie sterowały bez ingerencji ze strony człowieka, co ma poprawić konkurencyjność i odbudować pozycję geoekonomiczną zachodnioeuropejskiego przemysłu, które ten w ostatnim ćwierćwieczu utracił z powodu zbyt wysokich kosztów produkcji (przede wszystkim tzw. kosztów osobowych). Ponieważ procesy stopniowego rozszerzania Unii Europejskiej o kraje gospodarczo zacofane lub zdewastowane przez transformacje ustrojowe spowodowały jedynie przejściową poprawę koniunktury, poszukiwano innych środków zapobiegawczych, które powstrzymałyby odpływ kapitału i miejsc pracy z pogrążonej w stagnacji Europy Zachodniej do krajów szybko się rozwijających. Antidotum na problemy ekonomiczne Europy Zachodniej ma być inteligentna usieciwiona produkcja wykorzystująca najnowsze zdobycze IT, takie jak BigData, technologie chmurowe (CC), rzeczywistość rozszerzona, technologie przyrostowe (3D, AM), autonomiczne roboty, technologie symulacyjne, internet rzeczy, systemy zintegrowane oraz najnowsze rozwiązania z zakresu inżynierii (cyber-)bezpieczeństwa. Por. Acatech (National Academy of Science and Engineering), *Securing the future of German manufacturing industry. Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final report of the Industrie 4.0 Working Group*, Frankfurt am Main 2013.

¹⁰ Por. M. Decker, *op. cit.*, s. 355.

¹¹ Por. L. Bainbridge, *Ironies of Automation*, „Automatica” 19 (1983) 6, s. 775-779; L. Bainbridge, *Ironies of Automation*, [w:] J. Rasmussen, K. Duncan, J. Leplat (eds.), *New Technology and Human Error*, Chichester, J. Wiley, New York 1987, s. 271-283.

łatwiejszą część jego pracy (czynności powtarzalne, wykonywane rutynowo bez większego wysiłku umysłowego), automatyzacja uczyniła trudniejszą część zadań ludzkiego operatora jeszcze trudniejszą. Praca operatora złożonych systemów wymaga wysokiej koncentracji, podzielności uwagi, ciągłej gotowości do szybkiego reagowania i utrzymywania w stałej gotowości oraz ciągłego uzupełniania zasobów specjalistycznej wiedzy potrzebnych do rozwiązywania złożonych problemów. Taki profil zadaniowy w połączeniu z nieregularnym trybem życia związanym z pracą na trzy zmiany stanowi czynnik przeciążenia człowieka, przyczynę problemów zdrowotnych i częstszych absencji oraz dodatkowe źródło błędów, których wyeliminowanie było w zamyśle powodem zastępowania człowieka inteligentnymi maszynami. Zbyt wysokie kognitywne obciążenie pracą – zarówno monotonne, jak i zmienne – może powodować przeciążenia psychiczne, fizyczne, sensoryczne, błędy w działaniu oraz krótkotrwałe lub trwałe problemy ze zdrowiem. Natomiast zbyt niskie kognitywne obciążenie pracą powoduje dekwalfikację pracowników, jest źródłem problemów motywacyjnych oraz przyczyną często nieodwracalnego zaniku kompetencji. Całkowite wyeliminowanie człowieka z przemysłowych procesów wytwórczych póki co jest utopią, bowiem nawet najbardziej zautomatyzowany system potrzebuje ludzi do nadzorowania funkcji systemu i reagowania na jego nadzwyczajne, nietypowe zachowania. Projektanci autonomicznych systemów próbują wyeliminować czynnik ludzki jako źródło błędów, ale z jednej strony oni sami również są „czynnikiem ludzkim” i potencjalnym źródłem błędów jeszcze bardziej brzemiennych w skutkach, z drugiej nie wszystkie elementy procesów poddają się automatyzacji. Części każdego procesu, co do których projektanci nie wiedzą, jak można je zautomatyzować, muszą być więc nadal sterowane przez ludzkiego operatora.

Problemy na gruncie teorii bezpieczeństwa

Obok problemów na gruncie teorii pracy i ich konsekwencji socjoekonomicznych dla jednostki i społeczeństwa podstawowe znaczenie z punktu widzenia społecznych oddziaływań i skutków procesów autonomizacji techniki mają również aspekty bezpieczeństwa. Skoro jednym z głównych powodów zastępowania człowieka inteligentnymi maszynami w określonych obszarach działalności (wytwórczość, usługi, społeczne interakcje) są wzrastające obecnie wymagania bezpieczeństwa, to należałoby się bliżej przyjrzeć rzeczywistym zyskom w zakresie bezpieczeństwa oraz ocenić ich sensowność w perspektywie zagrożeń występujących w innych obszarach życia jednostki i społeczeństwa, w których z różnych względów nie da się w takim samym stopniu wyeliminować czynnika ludzkiego. Może się bowiem okazać, że podniesienie przewidywalności i standardów bezpieczeństwa w pewnych ogniach systemu zarządzania bezpieczeństwem wcale w sposób istotny nie poprawi ani ogólnego poziomu

bezpieczeństwa, ani społecznego poczucia bezpieczeństwa, jeśli bezpieczeństwo systemowe będą obniżały inne słabe ogniwa.

Kwestie poprawy bezpieczeństwa są powszechnie podnoszone przez zwolenników autonomicznych systemów transportowych (autonomicznych pojazdów). W świetle dotychczasowych badań i szacunków autonomizacja transportu znacząco ograniczy liczbę kolizji drogowych oraz wynikających z nich szkód i obrażeń oraz kosztów ich likwidacji, a co za tym idzie – obniży koszty ubezpieczeń komunikacyjnych¹². Ponieważ większość zdarzeń drogowych jest wynikiem błędów człowieka (niewłaściwa ocena sytuacji i własnych umiejętności, niedostosowanie prędkości do warunków drogowych, brawura, dekoncentracja, opóźniony czas reakcji, osobiste porachunki między kierującymi, jazda pod wpływem alkoholu lub środków odurzających itp.), wyeliminowanie człowieka wydaje się antidotum na problemy ryzyk w transporcie.

Z punktu widzenia inżynierii bezpieczeństwa bezpieczeństwo infrastruktury, działań lub wytworów sprowadza się w sensie ogólnym do wykluczenia (poważnych) zagrożeń dla dóbr podlegających ochronie prawnej lub świadomego obcowania, radzenia sobie z takimi zagrożeniami¹³. Automatyzacja procesów technicznych, zdanie się na nawet najinteligentniejsze roboty z pewnością przyczynia się do podniesienia przewidywalności zachowań systemów technicznych i wykluczenia wielu zagrożeń, ale nie wnosi niczego istotnego do świadomego, zreflektowanego obcowania z zagrożeniami i radzenia sobie z niewiedzą. Bezpieczne stany z udziałem ludzi mają zupełnie inne właściwości niż bezpieczeństwo oparte wyłącznie na aparaturach technicznych. Na poziomie technicznych systemów rzeczowych problemy bezpieczeństwa determinowane są przez zależności systemowe o charakterze dynamicznym (interakcyjnym: wzajemne oddziaływania między podsystemami technicznymi oraz nimi a ludźmi) oraz statycznym (sprzężeniowym: stosunki między podsystemami technicznymi). Poszczególne interakcje mogą mieć strukturę liniową (zdarzenie występuje w spodziewanym i znanym przebiegu procesu lub jest dobrze obserwowalne dla operatora nawet wtedy, gdy występuje w sposób poplanyowy) lub nieliniową, złożoną (zdarzenie jest albo zaplanowane, ale nieznanie operatorowi, albo nieplanowane i nieoczekiwane, dla obsługi albo niewidoczne, nieprzejrzyste, albo w inny sposób niekontrolowalne). Sprzężenia mogą zaś mieć albo charakter silny, ścisły albo słaby, luźny. O sprzężeniach silnych mówi się wtedy, kiedy między dwoma wzajemnie połączonymi elementami nie występują luzy ani elastyczności, co oznacza, że każda zmiana zachowania jednego ele-

¹² Według szacunków amerykańskiej firmy doradczej McKinsey & Company, upowszechnienie autonomicznych pojazdów może wyeliminować nawet 90% wypadków drogowych w USA, zapobiec nawet 190 miliardom dolarów odszkodowań i kosztów opieki zdrowotnej rocznie oraz uratować tysiące istnień ludzkich. Por. M. Bertonecello, D. Wee, *Ten ways autonomous driving could redefine the automotive world*, „McKinsey&Co. Automotive&Assembly Quarterly” June 2015.

¹³ Por. G. Banse, *Industrie 4.0 aus der Sicht der Technikphilosophie und der Technikfolgenabschätzung*, [w:] G. Banse, U. Busch, M. Thomas (Hrsg.), *Digitalisierung und Transformation. Industrie 4.0 und digitalisierte Gesellschaft*, Trafo, Berlin 2017, s. 131n.

mentu skutkuje bezpośrednio zmianą zachowania drugiego elementu. Natomiast sprzężenia luźne umożliwiają określonym elementom systemu funkcjonowanie zgodnie z własną logiką bez destabilizacji systemu. W przypadku systemów technicznych charakteryzujących się silnymi sprzężeniami i nieliniowymi, złożonymi interakcjami wypadki i katastrofy są czymś normalnym¹⁴. Inteligentne obcowanie z zagrożeniami i radzenie sobie z niewiedzą trudno w sposób jednoznaczny zalgorytmizować, aby umożliwić inteligentnym systemom maszynowe uczenie się. Przy okazji wypadków i niebezpiecznych incydentów uwarunkowanych technicznie człowiek doświadcza utraty kontroli nad zależnościami, co do których wcześniej zakładał pełne panowanie nad sytuacją. Ale maszynowe uczenie się na zasadzie prób i błędów nie byłoby właściwe zwłaszcza w przypadku systemów, których niezawodność i ciągłość pracy jest warunkiem przetrwania jednostki, ładu społecznego, stabilności gospodarki lub trwałości państwa (np. krytyczne infrastruktury, systemy wytwarzające leki, aparatury ratujące życie itp.).

Dodatkowe komponenty w systemie człowiek-technika wcale nie muszą (w sposób automatyczny) podnosić bezpieczeństwa, mogą bowiem stwarzać potencjalnie nowy problem bezpieczeństwa za sprawą wzrostu złożoności. Prawidłowe, niezawodne działanie nowoczesnych systemów autonomicznych w dużej mierze zależy od bezpieczeństwa cybertechnologii, a ono z kolei wymaga spełnienia wielu trudno wykonalnych warunków¹⁵. W praktyce wobec wzrastającej złożoności systemów i nieprzejrzystości interakcji między ich pojedynczymi elementami, niezamierzonych lub nieprzewidzianych interakcji człowiek-technika, problemów autoryzacyjnych wynikających z pseudonimowości lub anonimowości usług sieciowych, chmur obliczeniowych, autonomicznych dynamik oraz potencjałów destabilizacyjnych związanych z wrażliwością systemów, możliwością pozbawienia zasilania¹⁶, awariami sprzętowymi, niepożądanymi incydentami z udziałem człowieka i wieloma innymi zdarzeniami, zapewnianie bezpieczeństwa autonomicznych systemów opartych na automatyce i nowoczesnych technologiach informacyjno-komunikacyjnych stanowi coraz poważniejsze wyzwanie.

¹⁴ Zob. Ch. Perrow, *Normale Katastrophen. Die unvermeidbaren Risiken der Großtechnik*. Springer, Campus Verlag, Frankfurt am Main, New York 1987.

¹⁵ Por. D. Balzer, *Die gegenwärtige und zukünftige Rolle der Automatisierungs- und Kommunikationstechnik in der Sicherheitswirtschaft*, [w:] I. Oleksiewicz, K. Stępień (red.), *Zagrożenia i wyzwania bezpieczeństwa współczesnego świata. Wymiar polityczno-społeczny*, Rambler, Warszawa 2016, s. 238.

¹⁶ Widmo paraliżu „inteligentnych fabryk” spowodowanego przerwaniem zasilania centralnych serwerowni, regionalnych centrów danych itp., z usług których korzysta miliony aplikacji sieciowych wykorzystywanych w przemyśle 4.0, przestało być teoretycznym konstruktem 9 listopada 2017 r. W ten czwartkowy poranek doszło do równoczesnego odłączenia dwóch niezależnych linii zasilających jedną z największych serwerowni w Europie, należąca do firmy OVH – dużego francuskiego dostawcy globalnych zintegrowanych usług hostingowych i chmurowych, którego ponad 260 tys. serwerów na całym świecie obsługuje ok. 18 mln serwisów i aplikacji sieciowych. Ta bezprecedensowa awaria systemu zasilania sparaliżowała pracę kilku zachodnioeuropejskich centrów danych należących do firmy oraz spowodowała poważną awarię znacznej części zależnej sieci szkieletowej, przerywając na wiele godzin pracę serwerów w połowie Europy, wskutek czego przestały działać dziesiątki tysięcy serwisów internetowych, a duża część klientów OVH utraciła dostęp do usług zlokalizowanych w centrach danych na terytorium Wielkiej Brytanii, Niemiec i Francji.

Problemy etyczne

Możliwości zastąpienia człowieka przez autonomiczne systemy można analizować pod kątem technicznej wykonalności, ekonomicznej opłacalności, bezpieczeństwa albo prawnej lub etycznej dopuszczalności. Ten ostatni aspekt otwiera wiele wymiarów problemowych wyznaczonych przede wszystkim przez zmieniające się warunki możliwości odpowiedzialności w świecie coraz bardziej wypełnionym autonomicznymi systemami technicznymi¹⁷. Ze społecznego i etycznego punktu widzenia głównymi problemami związanymi z rozwojem i upowszechnianiem automatyzacji i robotyzacji są:

- bezpośrednie zagrożenia dla bezpieczeństwa jednostek lub zbiorowości, w tym zwłaszcza zagrożenia dóbr chronionych (życie, zdrowie, mienie *etc.*) wynikające z nieprzewidywanych lub niekontrolowanych zachowań systemów technicznych,
- niebezpieczeństwa związane z instrumentalizacją człowieka, w tym szczególnie groźba instrumentalizacji w sferze pracy,
- społecznie niesprawiedliwe rozkłady szkód i korzyści wynikających z rozwoju i upowszechniania automatyzacji i robotyzacji, zwłaszcza w odniesieniu do sfery zatrudnienia,
- wprowadzanie nowych rozwiązań na społecznie nieuczciwych warunkach, a więc bez zapewniania wszystkim grupom interesariuszy równych szans udziału w procesie decyzyjnym.

Upowszechnienie robotów przemysłowych w procesach wytwórczych spowodowało skutki, które są odmiennie postrzegane i oceniane przez różne podmioty. Dla przykładu w przedsiębiorstwie zupełnie inaczej zapatrują się na tego typu zmiany technologiczne kadry kierownicze odpowiedzialne za organizację produkcji czy osoby odpowiedzialne za zapewnianie i kontrolę jakości, a inaczej przedstawiciele pracowników i związkowcy. Procesy automatyzacji produkcji i upowszechnianie robotów pozostawiają pracownikom w praktyce tylko dwie możliwości: albo korzystając z ofert doksztalcania zawodowego, odpowiednio podniosą swoje kwalifikacje do poziomu wymaganego na wyższych stanowiskach operatorskich, albo zadowolą się niskopłatną, niepełnowartościową pracą, niewymagającą żadnych kwalifikacji zawodowych. Dzieje się tak pod wpływem wspomnianych wcześniej przeciwstawnych trendów związanych z jednej strony z obniżającymi się wymaganiami kwalifikacyjnymi stawianymi pracownikom fizycznym i wzrastającymi wymaganiami kwalifikacyjnymi stawianymi pracownikom na stanowiskach operatorskich i kontrolnych z drugiej. Innowacje technologiczne w sektorze produkcji mają swoich wygranych i przegranych również w innych aspektach. Z punktu widzenia wysoko uprzemysłowionych krajów Zachodu i Północy wykorzystanie

¹⁷ Por. A. Kiepas, *Przemysł 4.0 – nowe wyzwania dla odpowiedzialności*, [w:] L. Karczewski, H.A. Kretek (red.), *Etyczne i społeczne uwarunkowania biznesu, gospodarki i zarządzania*, Oficyna Wydaw. Politechniki Opolskiej, Opole 2015, s. 29-42.

robotów produkcyjnych na coraz większą skalę wiąże się co prawda ze społecznie relewantnymi przesunięciami w sferze pracy, ale jako takie gwarantuje w ogóle zachowanie w gospodarkach tych krajów przynajmniej części stanowisk pracy. Bez tych innowacji ze względu na koszty z dużym prawdopodobieństwem zostałyby one przemieszczone do krajów biednych, w których siła robocza jest znacznie tańsza niż w krajach wysoko uprzemysłowionych¹⁸. Jeśli zaś chodzi o zagrożenia związane z instrumentalizacją człowieka, to występują one szczególnie tam, gdzie czynności wykonywane przez człowieka sprowadzają się do tzw. robót przejściowych, których robotyzacja jest ekonomicznie nierentowna. Detronizacja człowieka poprzez sprowadzenie go do rangi „tańszego wariantu rozwiązania” może w określonych okolicznościach naruszać etyczny nakaz poszanowania godności osoby i być formą etycznie niedopuszczalnej instrumentalizacji człowieka, ale etycznemu osądowi można poddawać tylko konkretne praktyki po odpowiedniej interpretacji i analizie okoliczności (w tym także porównaniu wszystkich innych dostępnych wariantów działania). W konkretnym przypadku może się bowiem okazać, że każdy z dostępnych wariantów decyzyjnych godzi w zasadę etyczną tej samej lub wyższej rangi, co może istotnie wpływać na normatywną siłę osądu (kategoryczność, warunkowość). Nie bez znaczenia dla oceny etycznej jest również to, czy w danej sytuacji istnieją możliwości adekwatnej kompensacji szkód moralnych spowodowanych domniemaną instrumentalizacją człowieka.

Z zastosowaniami robotów serwisowych wiąże się jednak pewne problemy natury społeczno-etycznej, które nie występują w przypadku omówionych powyżej przemysłowych zastosowań robotów w procesach wytwórczych. Problemy te wynikają z faktu, że roboty serwisowe – w przeciwieństwie do robotów produkcyjnych w przemyśle – nie są zamknięte w pomieszczeniach zabezpieczonych przed „wizytami nieproszonych gości”, lecz są wykorzystywane w miejscach, w których w każdej chwili mogą pojawić się osoby postronne. Jeśli prace sprzętowe na polu uprawnym wykonuje autonomiczny traktor bez kierowcy, to nigdy nie da się wykluczyć, że w zasięgu jego oddziaływania znajdzie się rowerzysta lub osoba spacerująca na obrzeżach pola. Zawsze może się więc zdarzyć, że osoby postronne zostaną poszkodowane przez robota i problem ten otwiera interesującą perspektywę analityczną dotyczącą winy i odpowiedzialności za niebezpieczne zdarzenia. W świetle etyki odpowiedzialności podział zadań między człowieka i robota jako taki nie znosi problemu odpowiedzialności (moralnej, prawnej) za następstwa działania, lecz rozkłada tę odpowiedzialność na osoby w różny sposób zaangażowane w dane działanie proporcjonalnie do ich udziału, pozycji i znaczenia w kolektywnym lub korporacyjnym akcie sprawczym¹⁹. W przypadku robotów serwisowych sprawa jest pod względem strukturalnym dość skomplikowana, bo

¹⁸ Zob. T. Christaller, M. Decker, J.M. Gilsbach, G. Hirzinger, K.W. Lauterbach, E. Schweighofer, G. Schweitzer, D. Sturma, *Robotik. Perspektiven für menschliches Handeln in der zukünftigen Gesellschaft*, Berlin – Heidelberg 2001, s. 21n.

¹⁹ Zob. M. Maring, *Kollektive und korporative Verantwortung*, LIT Verlag, Münster 2001.

odpowiedzialnością za skutki działania robota dzielą się w różny sposób producent i operator urządzenia. W przypadku inteligentnych robotów uczących się, adaptujących się do konkretnych zadań rozkład odpowiedzialności na producenta i operatora nie może opierać się na rutynowych procedurach i standardach eksploatacyjnych, utrwalonych w toku dotychczasowych doświadczeń z infrastrukturami przemysłowymi. Społecznie akceptowalna eksploatacja nowej generacji robotów serwisowych wymaga więc zmiany obowiązujących przepisów zarówno w zakresie prawa cywilnego, jak i prawa publicznego²⁰.

Dodatkowe trudności z przypisaniem odpowiedzialności dochodzą w przypadku robotów serwisowych wykorzystywanych na użytek prywatny przez osoby, które nie mają odpowiedniego przeszkolenia oraz robotów socjalnych interagujących z osobami prywatnymi nie tylko niewykwalifikowanymi, ale także niemającymi pełni ludzkich dyspozycji, jak to się dzieje w przypadku robotów pielęgnujących chorych, seniorów, osoby z demencją lub osoby niepełnosprawne oraz robotów wykorzystywanych w pracy z dziećmi (np. robot Charlie wykorzystywany w terapii autyzmu²¹).

Głównym wyznacznikiem funkcjonalności robotów wykorzystywanych w przestrzeni prywatnej osób niemających specjalistycznego wykształcenia w zakresie programowania takich urządzeń są zdolności adaptacyjne i umiejętności uczenia się robotów. Na przykład robot do odkurzania musi mieć zdolności do autonomicznego rozpoznania nowego otoczenia i orientacji w nim. Proces samoczynnej pierwszej instalacji takiego robota w nowym otoczeniu może trwać nawet kilka dni, ale nie wymaga udziału ani obecności człowieka. W tym przypadku procesy uczenia się nie mają charakteru ciągłego i ograniczają się do otoczenia fizycznego nadającego się do parametrycznej ilościowej charakterystyki i penetracji sensorycznej. O wiele trudniej zaprogramować procesy uczenia się w przypadku robotów społecznych ze względu na konieczność dostosowywania się robota do różnych użytkowników, zwłaszcza pod kątem cech tych osób i wymagań wymykających się metodom uczenia maszynowego. Proces uczenia się w przypadku takich zastosowań ma charakter ciągły, nieskończony i wymaga nie tylko długotrwałych, intensywnych interakcji robot-użytkownik, ale także częstych interwencji programisty. Taka sytuacja stawia jednak w zupełnie innym świetle problem społecznego rozkładu odpowiedzialności za działania realizowane przez robota i ich skutki. Producent, który dostarcza inteligentnego robota uczącego się, w toku procesu samoadaptacji robota do nowych zadań ma coraz mniejsze możliwości trafnego przewidywania jego zachowań, nie znając historii procesu uczenia się ani nie mogąc go wstecz zrekonstruować nawet w przypadku algorytmów uczących opartych na sztucznych sieciach neuronowych. Z tego względu producent będzie zwykle dążył do ograni-

²⁰ Por. S. Beck, *Roboter, Cyborgs und das Recht. Von der Fiktion zur Realität*, [w:] T.M. Spranger (Hrsg.), *Aktuelle Herausforderungen der Life Science*, LIT Verlag, Berlin – Münster 2010, s. 95-120.

²¹ Zob. H. Li, J.J. Cabibihan, Y.K. Tan, *Towards an Effective Design of Social Robots*, "International Journal of Social Robotics", 2011 (3), s. 333.

czenia swojej odpowiedzialności za ewentualne niepożądane zachowania robota. Tymczasem użytkownik ma jeszcze bardziej ograniczone możliwości zapobiegania takim niepożądanym zachowaniom złożonych i samoczynnie zmieniających się systemów technicznych, takich jak roboty uczące się. Pod względem prawnym obowiązki użytkownika systemów technicznych sprowadzają się do przestrzegania instrukcji obsługi i właściwego, terminowego przeprowadzania czynności serwisowych przewidzianych w takiej instrukcji. Jednak w przypadku inteligentnych, uczących się robotów społecznych standardowe wymagania kontroli nad procesem uczenia się, która mogłaby stanowić podstawę do obarczenia użytkownika odpowiedzialnością za błędy robota, nie mogą być w praktyce zrealizowane. Skutkuje to powstaniem tzw. „szarej strefy” między odpowiedzialnością producenta a odpowiedzialnością użytkownika²². Również cena, jaką płaci się obecnie np. za zwiększenie samowystarczalności pacjentów i osób starszych dzięki robotom pielęgnacyjnym, które umożliwiają takim osobom przebywanie w domu w przyjaznym otoczeniu społecznym, w którym takie osoby czują się bardziej komfortowo niż w domach opieki, jest źródłem poważnych problemów społecznych związanych z rozkładem odpowiedzialności. Nie ma dotąd zgody co do tego, czy systemy zabezpieczenia społecznego powinny brać na siebie tak wysokie koszty pielęgnacyjne i rozkładać je na szerokie barki społeczne na zasadzie solidaryzmu. Dopóki ten problem nie zostanie rozwiązany w sposób zadowalający, roboty pielęgnacyjne pozostaną źródłem konfliktów społecznych związanych z tym, że beneficjentami takich innowacji w społeczeństwie są wyłącznie osoby o określonym profilu dochodowym lub majątkowym²³.

Wylimitowanie człowieka i co dalej?

Postępy w sieciowej integracji egzystencjalnie fundamentalnych procesów i czynności wykonywanych na przemian przez człowieka i maszyny będą wymagały poddania dotychczasowych indywidualistycznych pojęć etycznych i wyobrażeń o odpowiedzialności, ufundowanych na kognitywistycznym przekonaniu o zasadniczej przewidywalności skutków działania, gruntownej rewizji. Ze względu na zasadniczą niewiedzę i niepewność co do powodzenia lub niepowodzenia i ewentualnych przyszłych następstw inicjowanych obecnie innowacyjnych procesów nowym kryterium odpowiedzialności działania stanie się zapewne właściwy poziom zaufania do systemów technicznych²⁴. Wobec ograniczonej wiedzy poszczególnych podmiotów zaangażowanych w proces sieciowych interakcji interpersonalnych,

²² Por. A. Matthias, *The responsibility gap: Ascribing responsibility for actions of learning automata*, „Ethics and Information Technology” 6 (2004), s. 175-183; M. Decker, *op. cit.*, s. 357.

²³ Por. M. Decker, *op. cit.*, s. 357.

²⁴ Por. A. Kiepas, *Filozofia techniki w dobie nowych mediów*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2017, s. 156.

interakcji człowiek – maszyna i interakcji część – całość, odnoszącej się do ostatecznych całościowych rezultatów tego typu złożonych zbiorowych oddziaływań, zapewnianie koniecznych warunków możliwości odpowiedzialnego działania przemiesci się z tradycyjnej płaszczyzny intencjonalności jednostek na płaszczyznę technicznej optymalizacji i niezawodności, które będą wymagały inteligentnego radzenia sobie z niepewnością i niewiedzą. Ze względu na złożoność i nieprzejrzystość stosunków i interakcji we współczesnych systemach społeczno-technicznych, krzyżowe wpływy, zwrotne sprzężenia, efekty kumulacyjne, kaskadowe, synergiczne, bumerangowe i rykoszetowe nie ma możliwości podzielenia się odpowiedzialnością wzdłuż liniowych łańcuchów kauzalnych. Ale oddziaływania sieciowe i systemowe można nadal z powodzeniem opisywać za pomocą analiz skutków i ich społecznych rozkładów. Ze względu na hipotetyczny charakter szacunków odnoszących się do przyszłych skutków działania i ich społecznego rozkładu w świecie wszechobecnych autonomicznych systemów technicznych społeczna i etyczna (współ-) odpowiedzialność sprowadza się z konieczności do umiejętnego zarządzania ryzykiem. Wymaga ono nie tylko możliwie pełnej, trafnej identyfikacji możliwych wariantów działania, wczesnego, kompletnego (komprehenzywnego), kognitywnie spolegliwego rozpoznania szans i zagrożeń związanych z wyborem poszczególnych wariantów, trafnego oszacowania prawdopodobieństwa wystąpienia ewentualnych szkód lub korzyści oraz analizy ich społecznego rozkładu, ale także społecznie wiarygodnej oceny poszczególnych wariantów działania oraz kolektywnego podejmowania decyzji na społecznie uczciwych warunkach (tj. przy zapewnieniu wszystkim potencjalnym interesariuszom prawa do współdecydowania). Ze względu na to, że o percepcji, ocenie i akceptacji ryzyka decydują głównie czynniki natury subiektywnej²⁵ (czynniki biograficzne, indywidualne poczucie sensu życia i ocena jego wartości, indywidualne drzewo wartości itp.), podjęcie określonego ryzyka, a zwłaszcza narażenie innych na ryzyko nieuchronnie związane z każdym działaniem sieciowym, można społecznie legitymizować tylko pod warunkiem, że osoby, w które te ryzyka najbardziej uderzą, wyrażą na nie zgodę. Stanie się tak wtedy i tylko wtedy, gdy niepodjęciu takiego ryzyka sprzeciwią się społecznie o wiele ważniejsze powody, a potencjalnie uszkodzonym zagwarantuje się adekwatną, akceptowalną dla nich kompensację. Praktyczna realizacja tych postulatów wymaga upowszechnienia w społeczeństwie kultury komunikacyjnej i obywatelskiego zaangażowania oraz rozległego wbudowania elementów partycypacyjnych w procesy decyzyjne zarówno na szczeblu polityki państwa, jak i na szczeblu przedsiębiorstw²⁶.

²⁵ Por. A. Kiepas, *Człowiek wobec dylematów filozofii techniki*, Gnome, Katowice 2000, s. 53n.; K. Michalski, *Filozofia a ryzyko. Filozoficzne aspekty percepcji, oceny i akceptacji ryzyka*, [w:] R. Klamut, K. Michalski (red.), *Percepcja, ocena i akceptacja ryzyka*. Wybrane zagadnienia, Eikon Plus, Kraków 2006, s. 49-70.

²⁶ Por. K. Michalski, *Uczestnictwo obywateli w publicznych procesach decyzyjnych jako forma aktywności obywatelskiej na przykładzie partycypacyjnej oceny technologii*, [w:] R. Klamut, H. Sommer, K. Michalski (red.), *Aktywność obywatelska we współczesnym społeczeństwie demokratycznym*, Seiton, Kraków 2010, s. 59-110; K. Michalski, *Partycypacyjna ocena technologii w demokratycznej polityce technologicznej (jako przykład za-*

Gdy w kontekście nowoczesnej automatyki i robotów mowa jest o „zastępowaniu człowieka”, to oczywiście nie chodzi o całkowitą, abstrakcyjną substytucję i wyeliminowanie człowieka ze wszystkich procesów konstytuujących współczesny świat, a jedynie o niektóre czynności, które roboty są w stanie wykonywać taniej, w sposób bardziej wydajny lub bez zastrzeżeń natury etycznej. Jeśli właściwie rozumie się działanie swoiście ludzkie i interpretuje je w świetle takich atrybutów, jak: samoświadomość, zdolność do refleksji nad sobą, emocjonalność, odpowiedzialne podejmowanie decyzji, pokierowane wolą dobra i roztropnością (praktyczną mądrością), czy zdolność do argumentacyjnego ugruntowania, uprawomocnienia takich decyzji na mocy ich uogólnialności (powszechnej, intersubiektywnej ważności), wówczas nie ma obawy o to, że przynajmniej w perspektywie krótkoterminowej systemy techniczne będą w stanie symulować człowieka w sposób nie do odróżnienia. Możliwe są jednak – i z całą pewnością w nieodległej przyszłości należy się spodziewać tego typu prób – wycinkowe lub peryferyjne przypadki substytucji człowieka przez roboty zwłaszcza na obszarach, w których już od pewnego czasu dominują niepełne, redukcjonistyczne koncepcje człowieka. Wbrew obawom osób awersyjnie nastawionych do zdobyczy automatyki i robotyki to nie sama robotyzacja świata życia jest źródłem problemów i zagrożeń dla człowieka, ale raczej pewne redukcjonistyczne wizje człowieka, zacierające kategoriałne różnice między ludzkimi dyspozycjami a zdolnościami systemów technicznych, np. między rozumnością ludzi a inteligencją maszyn. Jeśli ludzkość właściwie zadba o to, aby we wszystkich sferach życia społecznego oraz życia jednostki społecznie chronione i wymagane były konieczne warunki możliwości rozwoju i zachowania pełnej ludzkiej podmiotowości, to niezależnie od stanu zaawansowania technologii przyszłość człowiek pozostanie w swoich najważniejszych, substancjalnych dyspozycjach niezastąpiony.

Krzysztof Michalski

The Autonomization of Technology and the Undesirable Consequences of Eliminating Humans as a Source of Errors

Abstract

The article discusses selected social and ethical aspects of the dissemination of the intelligent autonomous systems, which result in the progressive elimination of human from technical processes. The first part presents the most important modern applications of autonomous systems in industry, transport and logistics, agriculture, safety engineering and services. Next, the current possibilities and limitations of the

autonomy of machines and technical systems are overallly assessed. The next two parts discuss, from the point of view of the theory of employment and the theory of security, the most important consequences of the autonomization of technology. In the fifth part, the most important ethical problems generated by autonomous technical systems are cataloged. In summary, the author outlines a vision of the future, in which people are not threatened with the processes of autonomization of technological systems themselves, but rather by reductionist concepts of man, misorienting these processes.

Keywords: Ethics of Technology, Technology Assessment, Foresight & Forecasting, Automatization, Robotics, Industry 4.0.

