

Modele i warunki rozwoju etyki sztucznej inteligencji w biznesie

Robert Sroka

Akademia Leona Koźmińskiego

rsroka@kozminski.edu.pl

Biznes odgrywa dominującą rolę w rozwoju produktów i usług wykorzystujących sztuczną inteligencję. Wyzwaniem jest zapewnienie, aby takie produkty i usługi były etyczne. Trudność zmierzenia się z tym zadaniem wynika z faktów, że (a) proces ich rozwoju jest nietransparentny; (b) nie zawsze firmy chcą brać udział w dyskusji na temat etyki sztucznej inteligencji oraz stosować się do jej wyników; (c) nowość zagadnienia powoduje, że rezultaty tych dyskusji nie powstają szybko i nie są praktyczne.

Celem artykułu jest wskazanie różnych modeli etyki sztucznej inteligencji rozwija-

nych przez biznes, a także warunków i problemów, jakie łączą się z tym zadaniem. Pokazano w jaki sposób biznes odpowiada na wyzwania etyczne spowodowane przez zjawisko rozwoju sztucznej inteligencji. Omówiono również odpowiedź, jakiej na te wyzwania udzieliła Grupa Robocza Wysokiego Szczebla powołana przez Komisję Europejską do wypracowania kodeksu etyki sztucznej inteligencji.

Słowa kluczowe: etyka sztucznej inteligencji, etyka biznesu, AI, sztuczna inteligencja

Wprowadzenie

Etyczne wyzwania związane ze sztuczną inteligencją są obecnie jednym z najczęściej podejmowanych tematów przez naukowców zajmujących się tym zagadnieniem zarówno od strony technologicznej, jak i filozoficznej. Można wskazać na dwie grupy ośrodków, w których prowadzone są prace nad rozwojem sztucznej inteligencji. Jedną są uniwersytety i instytuty badawcze, drugą prywatne

podmioty, poczynawszy od niewielkich start up'ów na korporacjach międzynarodowych skończywszy. Można założyć, że badania nad sztuczną inteligencją w ośrodkach naukowych od strony oceny etycznej są prowadzone transparentnie i podlegają w pewnym zakresie regulacjom i kontroli czy to przez sam *ethos* naukowca, czy poprzez różnego rodzaju komisje etyczne. Są również oceniane czy kontrolowane przez podmioty przyznające granty na tego typu badania. Natomiast prace nad sztuczną inteligencją prowadzone przez firmy, wydają się poza jakąkolwiek kontrolą społeczną od strony oceny etycznej. Sondaż przeprowadzony wśród firm zainteresowanych sztuczną inteligencją wskazuje, że 72% z nich wdrożyło już rozwiązania z tego obszaru lub jest w trakcie wdrożenia¹. Projekty z zakresu sztucznej inteligencji na ogół nie są przedstawiane na konferencjach naukowych a społeczeństwo, np. klienci, doświadczają ich rezultatów w momencie komercjalizacji już w postaci usługi czy produktu. Dlatego wielu naukowców, w tym Paweł Łupkowski, zauważa, że w badaniach nad sztuczną inteligencją mamy do czynienia z niebezpieczeństwem braku społecznej kontroli [Łupkowski 2005: 241].

Produkty czy usługi bazujące na sztucznej inteligencji nie są czymś, co nas czeka w przyszłości. Duże zbiory danych, dzięki którym sztuczna inteligencja może być rozwijana, już teraz są używane przez przedsiębiorstwa. Dostarczycielami tych danych jesteśmy my wszyscy. Codziennie jesteśmy poddawani inwigilacji, chociażby używając przeglądarek internetowych, telefonów komórkowych, robiąc zakupy, będąc z wizytą u lekarza czy zgłaszając szkodę ubezpieczycielowi. Skoro tak, to warto zastanawiać się nie tylko nad etycznymi dylematami, z którymi będziemy mieć do czynienia

¹ *AI Momentum, Maturity and Models for Success*, SAS Institute, 2018, https://www.accenture.com/t20180919T202227Z___w___us-en/_acnmedia/PDF-86/Accenture-AI-Momentum-Final.pdf [dostęp 2.04.2019]. Badanie przeprowadzono wśród 305 przedstawicieli biznesu z całego świata. Ponad połowa z nich pracuje na stanowisku *chief information officer* (dyrektora ds. informatyki), *chief technology officer* (dyrektora ds. technologii) oraz *chief analytics officer* (dyrektora ds. analityki).

w przyszłości, a które są powiązane z tak zwaną głęboką sztuczną inteligencją, ale również tymi, z którymi mamy do czynienia już w chwili obecnej. Zatem prowadzone w tym artykule rozważania nie dotyczą przyszłości, lecz procesu, którego jako ludzkość, jako biznes czy jako klienci już doświadczamy.

Problem, przed jakim stoi etyka biznesu, to możliwość opracowania standardów etycznych dla badań i komercjalizacji produktów i usług bazujących na sztucznej inteligencji, które byłyby respektowane przez biznes. W artykule wskazano różne modele etyki sztucznej inteligencji rozwijane przez biznes, a także warunki i problemy, jakie łączą się z tym zadaniem. Przedstawiono syntetycznie odpowiedzi biznesu na wyzwania etyczne, jakie niesie ze sobą zjawisko rozwoju sztucznej inteligencji. Omówiono również odpowiedź, jaką na te wyzwania udzieliła Grupa Robocza Wysokiego Szczebła powołana przez Komisję Europejską do wypracowania kodeksu etyki sztucznej inteligencji.

Obecnie szybki rozwój sztucznej inteligencji wpływa na wiele dziedzin życia poprzez oferowane przez firmy, ale również przez administrację produkty i usługi korzystające z efektów tego rozwoju. Dlatego dla powodzenia w opracowaniu i wdrożenia standardów etyki sztucznej inteligencji ważne jest, aby myślenie największych firm technologicznych w tym obszarze było zbieżne z kodeksem przygotowanym przez Grupę Roboczą Wysokiego Szczebła. Wydaje się, że tylko spójne podejście prywatnych podmiotów i administracji, która może przyjąć na siebie rolę regulatora, pozwoli na wypracowania zasad etycznych dla sztucznej inteligencji, które będą respektowane przez przedsiębiorstwa, które, jak już zostało wspomniane, rozwijają tę dziedzinę poza zewnętrzną kontrolą.

Sztuczna inteligencja – definicja i rodzaje

Zanim zostanie przedstawione podejście największych korporacji zajmujących się wykorzystaniem sztucznej inteligencji w swoich produktach i usługach, ważne jest zdefiniowanie omawianego ter-

minu. Istnieją dziesiątki różnych definicji sztucznej inteligencji (od angielskiej nazwy *artificial intelligence* powszechnie używa się skrótu AI). Nie jest celem tego artykułu wnikliwe analizowanie samego pojęcia sztucznej inteligencji. Dlatego dla uproszczenia posłużmy się definicją wypracowaną przez Grupę Roboczą Wysokiego Szczebla Komisji Europejskiej i przedstawioną w dokumencie *A definition of AI: Main Capabilities and Scientific Disciplines* [2019]. Czytamy w nim, że sztuczna inteligencja odnosi się do systemów, które wykazują inteligentne zachowanie, analizując swoje środowisko i podejmując działania – z pewnym stopniem autonomii – w celu osiągnięcia określonych celów [s. 2]. Zachowania inteligentne w tym przypadku to zdolność do takiego przetwarzania informacji, które prowadzi do kreowania nieoczywistych połączeń między posiadanymi danymi. Systemy zbudowane z wykorzystaniem sztucznej inteligencji mogą być oparte wyłącznie na oprogramowaniu i działać w świecie wirtualnym (np. asystenci głosowi, oprogramowanie do analizy obrazu, wyszukiwarki, systemy rozpoznawania mowy i twarzy) lub mogą być wbudowane w urządzenia sprzętowe (np. zaawansowane roboty, autonomiczne samochody, drony lub aplikacje dla internetu rzeczy ang. *Internet of Things*) [s. 2].

Podsumowując, systemy sztucznej inteligencji (AI) to systemy oprogramowania (bądź urządzeń) zaprojektowane przez ludzi, które, biorąc pod uwagę złożony cel, działają w wymiarze fizycznym lub cyfrowym. Ich środowiskiem działania jest gromadzenie i interpretowanie zebranych ustrukturyzowanych lub nieustrukturyzowanych danych, które są logicznie przetwarzane w oparciu o dostępną wiedzę, aby podjąć decyzję prowadzącą do osiągnięcia danego celu. Systemy sztucznej inteligencji mogą wykorzystywać reguły symboliczne, uczyć się modelu numerycznego, a także dostosowywać swoje zachowanie, analizując, w jaki sposób ich poprzednie działania wpływały na środowisko w jakim działają oraz w kontekście celu jaki mają osiągnąć [s. 6].

Dla podejmowanych w tym artykule rozważań ważne jest również wprowadzenie rozróżnienia wąskiej (inaczej słabej) i ogólnej

(inaczej silnej) sztucznej inteligencji. System bazujący na ogólnej sztucznej inteligencji to system bardzo zaawansowany, który może wykonywać większość czynności, w tym abstrakcyjne myślenie, które mogą wykonywać ludzie. Wąska sztuczna inteligencja to systemy, które mogą wykonywać jedno lub kilka określonych zadań [s. 5]. Obecnie stosowane systemy sztucznej inteligencji są przykładami wąskiej sztucznej inteligencji. Dalsze rozważania będą odnosiły się do wąskiej sztucznej inteligencji. Warto jednak zasygnalizować, że istnieje wiele wyzwań etycznych, naukowych i technologicznych, związanych z ogólną sztuczną inteligencją. Należą do nich takie zagadnienia jak myślenie zdroworozsądkowe, samoświadomość i zdolność maszyny do zdefiniowania własnego celu.

Warunki badań w zakresie etyki sztucznej inteligencji

Prowadzenie rozważań etycznych związanych ze sztuczną inteligencją i wypracowywanie konkretnych zasad i wytycznych już na tym etapie jej stosowania jest tym bardziej uzasadnione, że wielu badaczy tego zagadnienia [por. Bostrom 2016] nie zakłada linearnego przebiegu rozwoju historii i nie projektuje jej dalszych etapów na podstawie znajomości praw rozwoju dziejów [Przybysz 2017: 202]. Dlatego w badaniach nad etyką sztucznej inteligencji rozwijanej przez biznes ważne jest wypracowanie pewnego systemu ich prowadzenia. Powinien on, w opinii autora artykułu, spełniać trzy następujące warunki: (1) dostęp do aktualnej wiedzy o realizowanych projektach, (2) udział w debacie odpowiednich osób, (3) praktyczne i szybkie efekty działań.

Warunek pierwszy, czyli dostęp do aktualnej wiedzy o realizowanych projektach, jest trudny do spełnienia ze względu na fakt, że firmy nie udzielają wystarczającej ilości informacji, zasłaniając się tajemnicą przedsiębiorstwa. Wydaje się to zupełnie zrozumiałe ze względu na interes samych firm. I to ten warunek stanowi największe praktyczne wyzwanie w badaniach nad etyką sztucznej inteligencji.

Warunkiem drugim jest udział odpowiednich osób w debacie nad etyką sztucznej inteligencji. Powinny to być osoby mające dostęp do wiedzy o realizowanych w tym zakresie projektach, zwłaszcza jeśli dyskusja odnosi się do projektów realizowanych przez prywatny biznes. Jednocześnie powinny być to osoby decyzyjne i mające szeroką perspektywę, a niekoniecznie te, które rozwijają wąski fragment sztucznej inteligencji, ponieważ w tej dziedzinie należy dostrzegać zjawisko skrajnych specjalizacji. Ci, którzy pracują nad wąskim wycinkiem, niekoniecznie mogą zdawać sobie sprawę ze skutków, jakie mogą przynieść ich prace w połączeniu z innymi elementami. Twórcy algorytmu nie muszą być świadomi skutków jego komercjalizacji.

Warunek trzeci to praktyczne i szybkie efekty badań nad etyką sztucznej inteligencji. Należy go rozumieć następująco. Badania nad etyką sztucznej inteligencji, która jest rozwijana przez prywatne przedsiębiorstwa, należą do etyki biznesu i dlatego powinny spełniać cele stawiane etyce biznesu rozumianej jako dziedzina etyki stosowanej. Inaczej mówiąc, ich wyniki powinny proponować rozwiązania mogące znaleźć zastosowanie, a nie być tylko etyką postulatyczną. A taką pozostaną, jeśli do rozmów nie będą zaproszeni twórcy sztucznej inteligencji, lecz będą one prowadzone jedynie w gronie świadków, którzy dostrzegają niebezpieczeństwa, jednak, których głos jest niesłyszany, to znaczy nie ma wpływu na kształt i wynik procesów rozwoju produktów czy usług bazujących na sztucznej inteligencji. Czyli jeśli nie będzie spełniony powyżej przedstawiony warunek drugi.

W dokumencie Komisji Europejskiej *Ethics Guidelines for Trustworthy Artificial Intelligence* etyka stosowana została określona jako ta, która zajmuje się rzeczywistymi sytuacjami, w których decyzje muszą być podejmowane pod presją czasu, a często również z ograniczoną liczbą informacji. Etyka sztucznej inteligencji ogólnie postrzegana jako przykład etyki stosowanej koncentruje się na kwestiach normatywnych wynikających z projektowania, rozwoju, wdrażania i stosowania sztucznej inteligencji [*Ethics Guidelines...*, 2019: 37].

Spełnienie trzech opisanych warunków pozwala na prowadzenie takich badań nad etyką sztucznej inteligencji, które mogą skutkować zastosowaniem ich wyników w praktyce. A wyzwanie dla etyki stosowanej w obszarze sztucznej inteligencji jest ogromne. Jednym z nich jest pogląd, że dzięki sztucznej inteligencji chcemy się pozbyć odpowiedzialności przy podejmowaniu trudnych decyzji. Odpowiedzialność spoczywałaby na algorytmach znajdujących się na przykład w samochodach niewymagających kierowców, robotach eksploracyjnych i ratowniczych, systemach militarnych (np. dronach), robotach medycznych, ale również seksrobotach, które będą zyskiwały coraz większą autonomię. Należy także zadać sobie pytanie, czy jesteśmy wystarczająco gotowi technologicznie i etycznie, aby przekazać maszynom decyzje dotyczące czyjegoś życia lub śmierci? Ponadto, w przypadku wąskiej sztucznej inteligencji algorytmy wciąż piszą ludzie. Łatwo popełnić w nich różne rodzaje błędów poznawczych, nie przewidzieć wszystkich scenariuszy. Dla przykładu Cathy O'Neil opisała w książce *Broń matematycznej zagłady* [2017] wiele przykładów źle zaprojektowanych systemów. Potrafią one dyskryminować różne grupy ze względu na płeć i kolor skóry, wysyłać niegroźnych więźniów na recydywę oraz tworzyć enklawy dla uprzywilejowanych osób, np. białych, bogatych mężczyzn. Są też produkty napędzane sztuczną inteligencją, które z założenia zostały zaprojektowane do działań nieetycznych. Należą do nich systemy wspomagające manipulowanie ludźmi, sianie dezinformacji i strachu. Należy pamiętać, że mówiąc o algorytmach, mówimy również o ich twórcach lub o wytycznych, jakimi się kierowali.

Etyka sztucznej inteligencji – perspektywa biznesu

Etyka sztucznej inteligencji w biznesie powinna przede wszystkim odnosić się, do tych, którzy są dostawcami rozwiązań technologicznych. To oni są na początku wypracowania rozwiązań na niej opartych. I to ta grupa ma tendencję do unikania odpowiedzialności za swoje produkty czy usługi, twierdząc, że jej członkowie nie powinni

odpowiadać za przypadki zastosowania ich produktów czy usług do złych celów. Argumentują przez analogię, mówiąc, że dostarczane przez nich rozwiązania są jak nóż. Może służyć do krojenia chleba, ale też może stać się narzędziem zbrodni. Powstaje pytanie, czy aby na pewno taka argumentacja jest uzasadniona? Czy chociażby w jednoznacznych sytuacjach, gdy sprzedawca, wiedząc, że nóż będzie użyty do popełnienia zbrodni, nie powinien odmówić jego sprzedaży?

Wydaje się jednak, że porównanie z nożem nie jest adekwatne. W przeciwieństwie do producentów noży, w kontekście sztucznej inteligencji w większości przypadków nie wiadomo, nad czym pracują firmy. Często znane są dopiero już gotowe rozwiązania. A implementacja tych rozwiązań ma coraz częściej zasięg globalny, międzysektorowy oraz przenika różne sfery: biznesową, prywatną, publiczną. Powstaje pytanie, jak te działania mogą być kontrolowane już na etapie rozwoju? Na etapie, do którego opinia publiczna ani żadne ciała kontrolne nie mają dostępu.

Jednym z niewielu przykładów otwartej, światowej dyskusji etycznej jest zastosowanie sztucznej inteligencji w pojazdach autonomicznych, czyli takich, które poruszają się po drogach bez udziału kierowcy i które podejmują autonomiczne decyzje również w przypadku zaistnienia niebezpiecznych sytuacji. Zdaniem Tomasa Kwarciańskiego [2018] niewykluczone, że zaprogramowane przez producentów pojazdy zmienią nie tylko dotychczasową organizację ruchu ulicznego, lecz także nasze myślenie na temat dobra i zła. Kwarciański powołuje się na opublikowane w czasopiśmie „Nature” wyniki badań świadczące o tym, że opinie moralne na temat tego, jak w konkretnej sytuacji powinien zachować się samochód autonomiczny, są zróżnicowane kulturowo. Zadaje pytanie, która z nich powinna zatem zostać zaimplementowana w oprogramowaniu samochodów autonomicznych? Opinia najpopularniejsza czy najlepiej uzasadniona? Według Kwarciańskiego, trzeba zdawać sobie sprawę z tego, że standardy etyczne wprowadzane w samochodach autonomicznych w postaci algorytmów będą wpływać na

intuicje moralne ludzi. W algorytmach „zaszyte” są bowiem opinie samych programistów. I dodaje, że analiza etyczna sztucznej inteligencji staje się więc nieodzowna. Co jednak się stanie, gdy pojazd autonomiczny zdoła przewidzieć kolizję, ale nie będzie w stanie jej uniknąć? Czyje życie i zdrowie powinno być w wówczas chronione: pasażerów czy przechodniów? Jak zauważają autorzy projektu *Moral Machine* po raz pierwszy w historii stajemy przed koniecznością oddania decyzji moralnych w gestię maszyn sterowanych sztuczną inteligencją². Co więcej, wydarzy się to nie gdzieś daleko na jakiejś nieznanej wojnie, lecz tuż za rogiem, na ulicy, przy której bawią się nasze dzieci. Kto powinien decydować o ewentualnych scenariuszach [Kwarciański 2018]?

W przypadku zastosowania sztucznej inteligencji w samochodach naukowcy z Massachusetts Institute of Technology zbudowali wspomnianą wyżej stronę internetową „Moral Machine”, na której każdy internauta mógł podjąć własne decyzje, kto powinien przeżyć, a kto zginąć w kolizji pojazdu autonomicznego. Na platformie zgromadzono 40 milionów decyzji, podjętych w 10 językach, od ludzi mieszkających w ponad 200 krajach. Szczegółowe omówienie tych wyników to temat na inny artykuł. Ale sam mechanizm pokazuje, jak już dziś sztuczna inteligencja jest powiązana z „zaszytymi” w niej poglądami na temat dobra i zła.

Chociaż w powyższym opisanym przypadku proces uzgadniania zasad etycznych jest transparentny, to nie wiemy, jakie rozwiązania finalnie przyjmą producenci samochodów autonomicznych. I czy będą decydowały czynniki etyczne, czy może jednak opinie jakichś grup interesów, na przykład ubezpieczycieli. Podobnych dyskusji nad etyką sztucznej inteligencji jest niewiele, a produktów czy usług ją wykorzystujących powstają tysiące. Większość z nich jest skierowana do szerokiej rzeszy klientów.

Zdaje się jednak, że firmy są świadome elementu etycznego ich prac nad zastosowaniem sztucznej inteligencji. Jak wynika z bada-

² Źródło: <http://moralmachine.mit.edu/>

nia SAS, Accenture Applied Intelligence, Intela i Forbes Insights³, 70% firm wdrażających projekty z zakresu sztucznej inteligencji organizuje szkolenia dla pracowników, których celem jest zapewnienie, że projekty AI są realizowane z zachowaniem zasad etyki. Aż 92% liderów w tej dziedzinie, określających prowadzone projekty jako sukces lub duży sukces, szkoli pracowników z etyki sztucznej inteligencji. Organizacje, w których inicjatywy związane ze sztuczną inteligencją są mniej zaawansowane, realizuje takie szkolenia tylko w 48% przypadków. Pomimo że z założenia sztuczna inteligencja działa niezależnie od człowieka, badania wskazują, że nadzór nad tą technologią jest niezbędny. 74% najbardziej zaawansowanych technologicznie firm, które brały udział w badaniu, wskazało, że przywiązuje dużą wagę do ewaluacji projektów AI. Yinyin Liu, dyrektor w firmie Intel AI Products Group twierdzi, że możliwość zrozumienia, w jaki sposób systemy sztucznej inteligencji podejmują decyzje, buduje zaufanie i czyni możliwym skuteczny nadzór człowieka⁴. 60% organizacji, które wdrożyły projekty z zakresu sztucznej inteligencji lub planuje to zrobić, wskazało, że obawia się wpływu decyzji podejmowanych przez maszyny na zaangażowanie klientów, na przykład mogą one nie wykazać się dostateczną empatią w kontaktach z kupującymi⁵.

Aby poradzić sobie z wyzwaniem etycznymi związanymi ze sztuczną inteligencją, firmy podejmują działania nie tylko wewnątrz własnej struktury, ale również inicjują współpracę z zewnętrznymi partnerami. Celem takich działań jest między innymi ochrona wła-

³ *AI Momentum, Maturity and Models for Success*, SAS Institute.

⁴ *Liderzy biznesu szkolą pracowników z etycznych aspektów wykorzystania sztucznej inteligencji*, www.sas.com/pl_pl/news/informacje-prasowe-pl/2018/liderzy-biznesu-szkola-pracownikow-z-etycznych-aspektow-wykorzystania-ai.html [dostęp 16.06.2019].

⁵ *Organizations Are Gearing Up for More Ethical and Responsible Use of Artificial Intelligence*, <https://newsroom.accenture.com/news/organizations-are-gearing-up-for-more-ethical-and-responsible-use-of-artificial-intelligence-finds-study.htm> [dostęp 16.06.2019]. W nawiązaniu do badań przeprowadzonych przez SAS, Accenture Applied Intelligence, Intel i Forbes Insights w lipcu 2018 roku zawiera opinie 305 liderów biznesu z Ameryki Północnej i Południowej, Europy oraz regionu Azji i Pacyfiku.

snej wiarygodności. Dostrzegając możliwość etycznie negatywnych konsekwencji funkcjonowania sztucznej inteligencji, ponad pięćdziesiąt przedsiębiorstw w Niemczech wprowadziło wspólny znak jakości dla sztucznej inteligencji. Znak jakości Federalnego Zrzeszenia Sztucznej Inteligencji opiera się na przestrzeganiu podstawowych wartości etycznych oraz na bezpiecznej i zgodnej z zasadami ochronie prywatności w obchodzeniu się z wykorzystywanymi danymi. Oznacza to między innymi konieczność zapewnienia bezstronności wykorzystywanej bazy danych oraz posiadanie przejrzystej procedury rozwoju sztucznej inteligencji. Zgodnie z wymaganiami tego znaku firmy zobowiązują się w rozwoju i stosowaniu sztucznej inteligencji dbać o dobro ludzi oraz respektować podstawowe europejskie wartości godności człowieka⁶.

Innym sposobem firm w zakresie pracy nad etyką sztucznej inteligencji, który z jednej strony zapewnia poufność informacji, a z drugiej uwiarygadnia wysiłki rozwoju i przestrzegania standardów etycznych, jest współpraca z ośrodkami badawczymi. Facebook, jedna z najbardziej znanych firm wykorzystujących rozwiązania sztucznej inteligencji, zobowiązała się przekazać 7,5 miliona dolarów na badania etyki sztucznej inteligencji w celu przeanalizowania problemów i pozyskaniu wskazówek dla rozwoju etycznych rozwiązań. Facebook nawiązał współpracę z Uniwersytetem Technicznym w Monachium (TUM) i zamierza wesprzeć utworzenie niezależnego centrum badań etyki sztucznej inteligencji. Instytut Etyki w Sztucznej Inteligencji (The Institute for Ethics in Artificial Intelligence) tej uczelni ma zbadać podstawowe problemy wpływające na wykorzystanie sztucznej inteligencji oraz wypracować wskazówki dla społeczeństwa, przemysłu, ustawodawców i decydentów w sektorze prywatnym oraz publicznym. Zajmie się również kwestiami, takimi jak bezpieczeństwo, prywatność, uczciwość i przejrzystość sztucznej

⁶ *Podstawowe wartości etyczne dla sztucznej inteligencji*, deutschland.de 2019, <https://www.deutschland.de/pl/news/podstawowe-wartosci-etyczne-dla-sztucznej-inteligencji> [dostęp 2.04.2019].

inteligencji. W opinii przedstawicieli Facebooka sztuczna inteligencja oferuje wiele korzyści, jednak konieczne jest kontrolowanie, czy traktuje użytkowników uczciwie, chroni od niebezpieczeństwa i nie narusza prywatności. Firma Facebook wykorzystuje obecnie sztuczną inteligencję m.in. do wykrywania potencjalnych samobójców⁷.

Inny globalny potentat technologiczny, firma Google, prowadzi samodzielnie prace nad dylematami etycznymi i społecznymi związanymi z rozwojem sztucznych inteligencji. Realizuje to poprzez należącą do Google firmę DeepMind, w ramach której powstała specjalna grupa badawcza. Przedmiot badania to między innymi ustalenie priorytetów w programowaniu pojazdów autonomicznych w razie wypadku oraz zbadanie problemu uprzedzeń rasowych, zawartych w algorytmach stosowanych przez amerykański system sądownictwa.

Zatem rola, jaką na siebie przyjmuje biznes w zakresie rozwoju standardów etycznych, przybiera różnorodne formy od wewnętrznych wytycznych w zakresie etyki sztucznej inteligencji i mechanizmów kontrolnych, poprzez inicjatywy branżowe, szeroką partycypację, współpracę z instytucjami naukowymi czy tworzenie własnych ośrodków badawczych w zakresie etyki sztucznej inteligencji.

Powyższe przykłady wskazują na świadomość firm rozwijających produkty i usługi w oparciu o sztuczną inteligencję w zakresie wyzwań etycznych. Należy wskazać jednak na kilka problemów. Przede wszystkim są to działania woluntarystyczne. Firmy w zdecydowanej większości nie informują szerokiej opinii publicznej o prowadzonych pracach i ich rezultatach, zatem sam proces jest nietransparentny. W przedstawionych powyżej modelach trudno uzyskać możliwość jednoczesnej realizacji opisanych wcześniej trzech warunków etycznej oceny produktów czy usług działających na bazie sztucznej inteligencji. O ile inicjatywy wewnętrzne firm

⁷ M. Kuchta, *Facebook przekaże 7,5 miliona dolarów na badania etyki sztucznej inteligencji*, Social Press 2019, <https://socialpress.pl/2019/01/facebook-przekaze-75-miliona-dolarow-na-badania-etyki-sztucznej-inteligencji> [dostęp 1.04.2019].

pozwalają na dostęp do aktualnej wiedzy, to już takiej pewności nie ma, gdy stosowane jest rozwiązanie wypracowania oceny etycznej w ramach organizacji branżowej czy przez ośrodek badawczy. Natomiast inicjatywy wewnętrzne nie zapewnią udziału w debacie odpowiednich osób, na co jest większa szansa w modelu organizacji branżowych czy instytutów badawczych. Warunek praktycznych i szybkich efektów jest trudny do spełnienia w modelu szerokiej partycypacji, natomiast może być zrealizowany w pozostałych modelach, o ile by sprawnie funkcjonowały.

Trzeba podkreślić, że nie wszystkie firmy analizują swoje działania z perspektywy etycznej, a nawet jeśli to robią, to niekoniecznie stosują się do rekomendacji wynikających z takich analiz. Dlatego istotne wydaje się wypracowanie uznanego, przyjętego przez administrację i uznanego przez biznes kodeksu etyki sztucznej inteligencji.

Etyka sztucznej inteligencji w biznesie – w poszukiwaniu kodeksu

Od kilku lat w różnych ośrodkach naukowych, politycznych, pozarządowych trwają niezależne prace nad opracowaniem zasad etycznych dla sztucznej inteligencji. Obok wskazanych wyżej wysiłków prywatnych korporacji, tego typu prace prowadzone są również pod auspicjami rządów. Należy do nich między innymi australijski *Artificial Intelligence: Australia's Ethics Framework* [Dawson et al. 2019] czy rekomendacje Grupy roboczej ds. etyki i prawa sztucznej inteligencji zawarte w dokumencie polskiego Ministerstwa Cyfryzacji zatytułowanym *Założenia do strategii AI w Polsce. Plan działań Ministerstwa Cyfryzacji z 2018 roku* [2018].

Najbardziej kompleksowe prace w tym zakresie były prowadzone na poziomie Unii Europejskiej i dlatego można z dużym prawdopodobieństwem założyć, że to ich rezultat będzie osią dalszych badań nad etyką sztucznej inteligencji ujętą w pewnego rodzaju kodeks.

8 kwietnia 2019 roku Komisja Europejska opublikowała ostateczną wersję dokumentu *Ethics Guidelines for Trustworthy AI* przy-

gotowanego przez Grupę Ekspertów Wysokiego Szczebla ds. Sztucznej Inteligencji. Zespół pracujący nad tym dokumentem składał się z ponad pięćdziesięciu ekspertów (polityków, naukowców, praktyków) z różnych europejskich krajów. Prace nad wytycznymi miały charakter partycypacyjny. W pierwszej fazie nad ich treścią pracowali eksperci wchodzący w skład Zespołu. 18 grudnia 2018 roku opublikowali oni wersję dokumentu przygotowaną do otwartych konsultacji społecznych [*Draft Ethics Guidelines...*, 2018]. Po otrzymaniu ponad 500 uwag i komentarzy przedstawili ostateczną treść dokumentu.

W tym dokumencie etyka sztucznej inteligencji rozumiana jest jako zgodność działań rozwojowych, wdrożeniowych oraz w zakresie użytkowania sztucznej inteligencji z normami etycznymi, w tym prawami podstawowymi jako szczególnymi uprawnieniami moralnymi, zasadami etycznymi i powiązanymi wartościami podstawowymi [*Ethics Guidelines...*, 2018: 37].

Autorom dokumentu przyświecało przede wszystkim podejście, w którym człowiek stoi w centrum refleksji etycznej (*Human-Centric AI*). Z treści jednak nie wynika, że autorzy inspirowali się filozofią personalistyczną czy filozoficznym nurtem antropocentrycznym, które łączone są z koncepcjami filozoficznymi, dla których człowiek stanowi punkt odniesienia. Podstawowym źródłem wartości dla autorów są traktaty Unii Europejskiej oraz Karta Praw Podstawowych [s. 37].

W tytule omawianego dokumentu zostało zamieszczone słowo wiarygodność (*trustworthy*). Zdaniem autorów to właśnie wiarygodność jest warunkiem wstępnym dla ludzi i społeczeństw do opracowania, wdrożenia i korzystania z systemów sztucznej inteligencji. Bez systemów sztucznej inteligencji, które są ewidentnie godne zaufania, mogą wystąpić niepożądane konsekwencje i ich wykorzystanie może być utrudnione, uniemożliwiając realizację potencjalnie ogromnych korzyści społecznych i ekonomicznych, które mogą przynieść [s. 5].

Zgodnie z *Ethics Guidelines...* wiarygodna sztuczna inteligencja powinna być: (1) zgodna z prawem, (2) etyczna, czyli zgodna z zasa-

dami i wartościami etycznymi oraz (3) odporna, rzetelna, wytrzymała w sensie społecznym i technicznym. Te ostatnie charakterystyki wyrażono za pomocą angielskiego wyraz *robust*, który nie posiada adekwatnego tłumaczenia w języku polskim. Słowo to użyto do podkreślenia, że produkty i usługi bazujące na sztucznej inteligencji powinny być tak tworzone, aby nie powodować – nawet niezamierzonej – szkody. Każdy z trzech warunków jest zdaniem autorów omawianego dokumentu konieczny do zbudowania etycznej, a raczej używając terminologii dokumentu, wiarygodnej sztucznej inteligencji. Żaden nie jest jednak wystarczający i jedynie dzięki zestawieniu ich razem możliwe jest osiągnięcie pożądanego rezultatu, jakim jest tworzenie godnej zaufania sztucznej inteligencji [s. 2]. Godnej zaufania głównie z punktu widzenia pojedynczych użytkowników i społeczeństwa jako całości.

Intencją autorów nie było przygotowanie dokumentu zawierającego listę zasad etycznych dla sztucznej inteligencji, lecz pokazanie jak zasady etyczne powinny być stosowane w obszarze społecznym i technicznym związanym ze sztuczną inteligencją. W podzielonym na trzy części dokumencie, w pierwszej omówiono prawa podstawowe i zasady etyczne istotne z punktu widzenia sztucznej inteligencji. Druga część zawiera siedem podstawowych wymagań, które sztuczna inteligencja powinna spełnić, aby być wiarygodną. Zostały one podzielone na metody techniczne i nietechniczne. W trzecim rozdziale zaproponowano narzędzie do oceny jakości zastosowania wymienionych w rozdziale drugim siedmiu wymagań.

Jak wspomniano, w rozdziale pierwszym autorzy proponują, aby sztuczna inteligencja była rozwijana, wdrażana i wykorzystywana w sposób zgodny z takimi zasadami etycznymi jak poszanowanie ludzkiej autonomii, zapobieganie szkodom, sprawiedliwość oraz transparentność, rozumiana jako możliwość wyjaśnienia przyszłych i obecnych zdarzeń związanych z zastosowaniem sztucznej inteligencji. Autorzy dokumentu są świadomi, że na poziomie zastosowania pomiędzy tymi zasadami może dojść do konfliktu. Jeśli z jednej strony należy być wiernym zasadzie zapobiegania szkodom, wyini-

kającym chociażby z zachowania prawa do prywatności, to nie zawsze będzie możliwe zachowanie zasady możliwości wyjaśniania zdarzeń. Nie dają jednak gotowej odpowiedzi na to, jak z takim konfliktem należy postępować.

W realizacji tych zasad szczególną ochroną powinny być objęte jednostki i grupy bardziej niż inne bezbronne i narażone na niepożądane skutki. Tymi grupami mogą być dzieci, mniejszości narodowe czy osoby z niepełnosprawnościami [s. 2]. Dla zapewnienia etycznej sztucznej inteligencji konieczne jest całościowe i systemowe podejście angażujące wszystkich aktorów i procesy w kontekście społecznym i technicznym w całym cyklu „życia” produktu czy usługi [s. 5].

Budowa odpowiedniego środowiska dla rozwoju etycznej sztucznej inteligencji obejmuje również umożliwienie zgłaszania negatywnych skutków. Należy zapewnić zarówno możliwość zgłaszania działań lub decyzji, które przyczyniają się do nieetycznego działania sztucznej inteligencji, jak i reagowania na konsekwencje takiego działania. Identyfikacja, ocena, dokumentacja i minimalizacja potencjalnego negatywnego wpływu systemów sztucznej inteligencji jest szczególnie ważna dla osób bezpośrednio nimi dotkniętych. Autorzy dokumentu podkreślają, że odpowiednia ochrona musi być dostępna dla zgłaszających przypadki naruszenia zarówno osób, jak i organizacji pozarządowych, związków zawodowych lub innych podmiotów. Innym elementem tworzenia takiego środowiska jest zastosowanie ocen wpływu zarówno przed, jak i podczas opracowywania, wdrażania i stosowania systemów sztucznej inteligencji [s. 20].

Autorzy dokumentu zwracają się bezpośrednio do organizacji pracujących nad sztuczną inteligencją, w tym do firm. Rekomendują, że organizacje powinny ustanowić ramy i praktyki zarządzania, zarówno wewnętrzne, jak i zewnętrzne, zapewniające odpowiedzialność za etyczny wymiar decyzji związanych z rozwojem, wdrażaniem i wykorzystaniem systemów sztucznej inteligencji. Może to na przykład obejmować wyznaczenie osoby odpowiedzialnej za

kwestie etyczne dotyczące systemów sztucznej inteligencji lub wewnętrznego / zewnętrznego panelu lub rady ds. etyki. Wśród możliwych ról takiej osoby, panelu lub rady, należy wskazać zapewnienie nadzoru i doradztwa. Różne ośrodki certyfikacyjne mogą odgrywać w takich organach ważną rolę. Należy zapewnić kanały komunikacji z podmiotami nadzoru, dzielić się najlepszymi praktykami, omawiać dylematy lub zgłaszać pojawiające się problemy natury etycznej. Takie mechanizmy mogą uzupełniać nadzór prawny, ale nie mogą go zastąpić [s. 23].

W dokumencie określono siedem głównych wymagań, jakie powinna spełniać wiarygodna sztuczna inteligencja. Należą do nich: 1. Zapewnienie możliwości jednostek do (a) niezależnego działania i dokonywana wolnych wyborów (*human agency*) oraz (b) nadzoru nad działaniem produktów/usług. 2. Techniczne bezpieczeństwo i odporność na intencjonalne bądź nieintencjonalne wykorzystanie sztucznej inteligencji w celu wywołania negatywnych skutków. 3. Zapewnienie prywatności i bezpieczne zarządzanie danymi. 4. Zapewnienie transparentności, dzięki której użytkownicy będą świadomi, że mają do czynienia ze sztuczną inteligencją. 5. Dbałość o różnorodność, niedyskryminowanie i sprawiedliwe traktowanie wszystkich interesariuszy. 6. Troska o dobrobyt społeczny i środowisko naturalne. 7. Odpowiedzialność za podejmowane działania związane ze sztuczną inteligencją [s. 3].

Powyższe wskazówki mają na celu zapewnienie takiego środowiska tworzenia, rozwijania i wykorzystywania sztucznej inteligencji, która nie będzie powodowała szkody dla ludzi, ale też, o co postuluje m.in. Luciano Floridi [2018], dla świata zwierząt i środowiska naturalnego.

Skrótowo omówiony dokument stanowi jasne wytyczne wskazujące, jakie standardy etyczne i techniczne powinny być zapewniane przez firmy pracujące nad rozwojem sztucznej inteligencji. Jednak konieczna wydaje się kontrola etyczności rozwiązań bazujących na sztucznej inteligencji. Taką funkcję w obecnych warunkach prawnych mogą przyjąć na siebie organy administracji pań-

stwowej. Jednak ten wątek nie został wystarczająco podkreślony w dokumencie Komisji Europejskiej.

Zakończenie

W ostatnich kilku latach znacznemu przyśpieszeniu uległ proces tworzenia przez firmy produktów i usług bazujących na sztucznej inteligencji. Jest to możliwe nie tylko dzięki posiadaniu lepszej technologii, możliwości gromadzenia danych i zaawansowanych algorytmów do przetwarzania tych danych, ale przede wszystkim dzięki stale poszerzającemu się zakresowi ich pozyskiwania. Dostarczycielami danych pozwalających na rozwój sztucznej inteligencji jesteśmy my wszyscy, czy to bezpośrednio robiąc zakupy, czy korzystając z przeglądarek internetowych, czy pośrednio przekazując dane ośrodkom zdrowia lub urzędowi skarbowemu. Zamknięty cykl, w którym każdy z nas jest zarówno na początku, dostarczając dane, jak i na końcu, korzystając z usług i produktów powstałych na bazie tych danych, rodzi nowe, niespotykane wyzwania etyczne. Dlatego nadzór nad technologią sztucznej inteligencji jest niezbędny, zwłaszcza w jej aspekcie etycznym.

Transparentnych dyskusji nad aspektami etycznymi komercyjnego zastosowania sztucznej inteligencji, takich jak w przypadku samochodów autonomicznych, jest niewiele. Natomiast faktem jest, że większość prac nad rozwojem sztucznej inteligencji jest prowadzonych przez biznes. Dlatego istnieje potrzeba podjęcia takich działań, które będą spełniały trzy warunki: 1) dostępu do aktualnej wiedzy o realizowanych projektach, 2) udziału w debacie odpowiednich osób, 3) praktycznych i szybkich efektów działań.

Przedstawione w artykule badania wskazują, że firmy są świadome aspektu etycznego swoich prac nad sztuczną inteligencją. Inną sprawą jest, na ile traktują ten aspekt poważnie i uwzględniają przy projektowaniu produktów i usług, tak aby nie powodowały szkody. Opracowany przez Komisję Europejską dokument *Ethics Guidelines for Trustworthy Artificial Intelligence* może się stać ważnym wyznacz-

nikami standardów etycznych dla sztucznej inteligencji, pod warunkiem, że będzie stosowany przez firmy i nadzorowany przez administrację.

W niniejszym artykule, ze względu na jego określony cel, treść wytycznych została przedstawiona pobieżnie, ale warto podkreślić, że wiele z nich ma istotną wartość. Dokument zawiera również wątki niejasne, chociażby brak precyzyjnego opisu kategorii etycznych. Dlatego w dalszych badaniach nad etyką sztucznej inteligencji warto pochylić się krytycznie nad propozycją Komisji Europejskiej oraz przeanalizować, jak jest ona wykorzystywana przez organizacje rozwijające produkty i usługi oparte na sztucznej inteligencji.

BIBLIOGRAFIA

- A definition of AI: Main Capabilities and Scientific Disciplines*. 2019. High-Level Expert Group on Artificial Intelligence, Brussels.
- AI Momentum, Maturity and Models for Success*, SAS Institute. 2018. https://www.accenture.com/t20180919T202227Z_w_/us-en/_acnmedia/PDF-86/Accenture-AI-Momentum-Final.pdf [dostęp 2.04.2019].
- Bostrom N. 2016. *Superinteligencja. Scenariusze, strategie, zagrożenia*, Helion, Gliwice.
- Dawson D., et al. 2019. *Artificial Intelligence: Australia's Ethics Framework*, Data61 CSIRO, Australia.
- Draft Ethics Guidelines for Trustworthy AI*. 2018. European Commission, Brussels.
- Ethics Guidelines for Trustworthy Artificial Intelligence*. 2019. High-Level Expert Group on Artificial Intelligence, European Commission, Brussels.
- Floridi L. 2018. *AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*, „Minds and Machines” 28(4).
- Kuchta M. 2019. *Facebook przekaże 7,5 miliona dolarów na badania etyki sztucznej inteligencji*, Social Press 2019, <https://socialpress.pl/2019/01/facebook-przekaze-75-miliona-dolarow-na-badania-etyki-sztucznej-inteligencji> [dostęp 1.04.2019].
- Kwarciański T. 2018. *Moralność autonomiczna, czyli sztuczna inteligencja zamiast sumienia*, Obserwator Finansowy 2018, <https://www.obserwatorfinansowy.pl/forma/rotator/moralnosc-autonomiczna-czyli-sztuczna-inteligencja-zamiast-sumienia/> [dostęp 15.03.2019].
- Liderzy biznesu szkolą pracowników z etycznych aspektów wykorzystania sztucznej inteligencji*, www.sas.com/pl_pl/news/informacje-prasowe-pl/2018/liderzy-biznesu-szkola-pracownikow-z-etycznych-aspektow-wykorzystania-ai.html [dostęp 16.06.2019].
- Łupkowski P. 2005. *Rola etyki i antropologii w rozważaniach o sztucznej inteligencji*, „Ethos” 1–2 (69–70).
- Moral Machine*, <http://moralmachine.mit.edu/> [dostęp 15.03.2019].

O'Neil C. 2017. *Broń matematycznej zagłady*, przeł. M. Zieliński, Wydawnictwo Naukowe PWN.

Organizations Are Gearing Up for More Ethical and Responsible Use of Artificial Intelligence, <https://newsroom.accenture.com/news/organizations-are-gearing-up-for-more-ethical-and-responsible-use-of-artificial-intelligence-finds-study.htm> [dostęp 16.06.2019].

Podstawowe wartości etyczne dla sztucznej inteligencji, deutschland.de 2019, <https://www.deutschland.de/pl/news/podstawowe-wartosci-etyczne-dla-sztucznej-inteligencji> [dostęp 2.04.2019].

Przybysz P. 2017. *Czy należy obawiać się superinteligencji?*, „Studia Metodologiczne”, 38.

Założenia do strategii AI w Polsce. Plan działań Ministerstwa Cyfryzacji. 2018. Ministerstwo Cyfryzacji, Warszawa.

Summary

Models and conditions for the development of ethics of artificial intelligence by business

Business plays a dominant role in the development of products and services based on Artificial Intelligence (AI). The challenge is to ensure that products and services based on AI are ethical. The difficulty in dealing with this task stems from the facts that: a. the process of development is not transparent; b. companies do not always want to take part in discussions on the ethics of artificial intelligence and adhere to its results; c. the novelty of the issue causes that the results of these discussions do not arise quickly and are not practical. The aim of the article is to indicate the difficulty of conducting research on the ethics of AI developed by business. The article has been synthetically presented in what way business responds to ethical challenges brought about by the phenomenon of development of AI. The answer to these challenges was also given by the High-Level Working Group appointed by the European Commission to develop a code of artificial intelligence.

Key words: ethics of Artificial Intelligence, business ethics, AI, Artificial Intelligence