



KOMISJA
EUROPEJSKA

Bruksela, dnia 8.4.2019 r.
COM(2019) 168 final

**KOMUNIKAT KOMISJI DO PARLAMENTU EUROPEJSKIEGO, RADY,
EUROPEJSKIEGO KOMITETU EKONOMICZNO-SPOŁECZNEGO I KOMITETU
REGIONÓW**

Budowanie zaufania do sztucznej inteligencji ukierunkowanej na człowieka

1. WPROWADZENIE – EUROPEJSKA STRATEGIA NA RZECZ SZTUCZNEJ INTELIGENCJI

Sztuczna inteligencja (SI) może zmienić nasz świat na lepszy: może przyczynić się do poprawy opieki zdrowotnej, zmniejszenia zużycia energii, zwiększenia bezpieczeństwa samochodów oraz bardziej wydajnego użytkowania wody i zasobów naturalnych przez rolników. Sztuczną inteligencję można wykorzystać do przewidywania zmiany środowiska i klimatu, poprawy zarządzania ryzykiem finansowym oraz zapewnienia narzędzi do wytwarzania produktów dostosowanych do naszych potrzeb, generujących jak najmniejszą ilość odpadów. Sztuczna inteligencja może również pomóc w wykrywaniu nadużyć finansowych i zagrożeń dla cyberbezpieczeństwa oraz umożliwić organom ścigania skuteczniejsze zwalczanie przestępczości.

Sztuczna inteligencja może przynieść korzyści całemu społeczeństwu i gospodarce. Jest to technologia strategiczna, którą obecnie rozwija się i jest wykorzystywana w szybkim tempie na całym świecie. Sztuczna inteligencja niesie ze sobą nowe wyzwania dla przyszłości pracy i rodzi wątpliwości natury prawnej i etycznej.

Aby przezwyciężyć te problemy i w jak największym stopniu wykorzystać możliwości stwarzane przez SI, w kwietniu 2018 r. Komisja opublikowała strategię europejską.¹ Strategia w centrum rozwoju SI stawia człowieka – **SI ukierunkowana na człowieka**. Jest to trzytorowe podejście do zwiększenia zdolności technologicznych i przemysłowych UE oraz wykorzystania SI w całej gospodarce, przygotowania się do zmian społeczno-gospodarczych oraz zapewnienia odpowiednich ram etycznych i prawnych.

W ramach działań dotyczących strategii na rzecz SI, **Komisja wypracowała wraz z państwami członkowskimi skoordynowany plan w dziedzinie SI²**, przedstawiony w grudniu 2018 r. w celu stworzenia synergii i zebrania danych – surowca dla wielu zastosowań SI – oraz zwiększenia wspólnych inwestycji. Celem planu jest wzmocnienie współpracy transgranicznej oraz mobilizacja wszystkich zainteresowanych podmiotów w celu zwiększenia inwestycji publicznych i prywatnych do poziomu **co najmniej 20 mld EUR rocznie** w kolejnej dekadzie³. Komisja podwoiła swoje inwestycje w SI w ramach programu „Horyzont 2020” i planuje zainwestować 1 mld EUR rocznie z programu „Horyzont Europa” i programu „Cyfrowa Europa”, wspierając w szczególności wspólne przestrzenie danych w dziedzinie zdrowia, transportu i produkcji oraz duże obiekty eksperymentalne, takie jak inteligentne szpitale i infrastrukturę dla pojazdów zautomatyzowanych, a także program badań strategicznych.

Aby zrealizować ten wspólny program badań strategicznych, innowacji i wdrożenia, Komisja zintensyfikowała **dialog ze wszystkimi zainteresowanymi stronami** reprezentującymi przemysł, instytuty badawcze i władze publiczne. Nowy program „Cyfrowa Europa” będzie również miał kluczowe znaczenie dla umożliwienia dostępu do SI małym i średnim przedsiębiorstwom we wszystkich państwach członkowskich za pośrednictwem ośrodków innowacji cyfrowych, ulepszonych obiektów eksperymentalnych i doświadczalnych, przestrzeni danych i programów szkoleniowych.

Europejskie etyczne podejście do SI, oparte na renomie bezpiecznych i wysokiej jakości produktów, wzmacnia zaufanie obywateli do postępu cyfrowego i ma na celu stworzenie

¹

² COM(2018) 795.

³ W tym celu Komisja zaproponowała, aby w następnym okresie programowania 2021–2027 Unia przeznaczyła co najmniej 1 mld EUR rocznie na finansowanie inwestycji w SI z programów „Horyzont Europa” i „Cyfrowa Europa”.

przewagi konkurencyjnej europejskich przedsiębiorstw działających w obszarze SI. Celem niniejszego komunikatu jest uruchomienie kompleksowego etapu pilotażowego z udziałem jak najszerszego kręgu zainteresowanych stron w celu przetestowania praktycznego wdrażania wytycznych w zakresie etyki w odniesieniu do rozwoju i wykorzystania SI.

2. BUDOWANIE ZAUFANIA DO SI UKIERUNKOWANEJ NA CZŁOWIEKA

Europejska strategia w zakresie SI oraz skoordynowany plan jasno wskazują, że zaufanie **jest warunkiem wstępnym dla zapewnienia podejścia do SI ukierunkowanego na człowieka**: Sztuczna inteligencja nie jest celem samym w sobie, ale narzędziem, które musi służyć ludziom, mającym na względzie zwiększenie dobrostanu ludzi. W tym celu należy zadbać o wiarygodność SI. Wartości, na których opierają się nasze społeczeństwa, powinny być w pełni zintegrowane z kierunkiem rozwoju SI.

Unia opiera się na **wartościach, takich jak poszanowanie godności osoby ludzkiej, wolność, demokracja, równość, praworządność, jak również poszanowanie praw człowieka**, w tym praw osób należących do mniejszości⁴. Wartości te są wspólne dla społeczeństw wszystkich państw członkowskich, w których panuje pluralizm, niedyskryminacja, tolerancja, sprawiedliwość, solidarność i równość. Oprócz tego, **Karta praw podstawowych Unii Europejskiej** ujmuje w jednym tekście wszystkie prawa jednostki, prawa obywatelskie, polityczne, gospodarcze i społeczne, z których korzystają mieszkańcy Unii.

UE dysponuje **silnymi ramami regulacyjnymi**, które określają globalny standard w odniesieniu do SI ukierunkowanej na człowieka. Ogólne rozporządzenie o ochronie danych zapewnia wysoki standard ochrony danych osobowych i wymaga wdrożenia środków mających na celu zapewnienie ochrony danych już w fazie projektowania oraz domyślnie⁵. Rozporządzenie w sprawie swobodnego przepływu danych nieosobowych usuwa przeszkody w swobodnym przepływie danych nieosobowych i pozwala na przetwarzanie wszystkich kategorii danych w dowolnym miejscu w Europie. Przyjęty niedawno akt ws. cyberbezpieczeństwa przyczyni się do wzmocnienia zaufania do internetu, a proponowane rozporządzenie w sprawie prywatności i łączności elektronicznej⁶ również zmierza do osiągnięcia tego celu.

Sztuczna inteligencja niesie ze sobą nowe wyzwania, ponieważ umożliwia „uczenie się” przez maszyny oraz podejmowanie i wdrażanie decyzji bez udziału człowieka. Już niedługo ten rodzaj funkcjonalności stanie się standardem w przypadku wielu rodzajów towarów i usług, od smartfonów po pojazdy zautomatyzowane, roboty i aplikacje internetowe. Jednak decyzje podejmowane za pomocą algorytmów mogą się opierać na niekompletnych danych, przez co nie są wiarygodne, mogą być zafałszowane w wyniku cyberataków, mogą też być nieobiektywne lub po prostu błędne. Bezrefleksyjne poleganie na technologii może prowadzić

⁴ UE jest również stroną Konwencji ONZ o prawach osób niepełnosprawnych.

⁵ Rozporządzenie (UE) 2016/679. Ogólne rozporządzenie o ochronie danych gwarantuje swobodny przepływ danych osobowych w Unii. Zawiera także przepisy dotyczące podejmowania decyzji opartych wyłącznie na zautomatyzowanym przetwarzaniu danych, w tym profilowaniu. Osoby, których to dotyczy, mają prawo do uzyskania informacji o przypadku automatycznego podejmowania decyzji oraz do otrzymywania istotnych informacji na temat logiki leżącej u podstaw automatycznego podejmowania decyzji, jak również znaczenia i przewidywanych konsekwencji przetwarzania danych dla tych osób. W takich przypadkach mają one również prawo do interwencji człowieka, wyrażenia swojego stanowiska i zaskarżenia decyzji.

⁶ COM(2017) 10.

do otrzymania problematycznych wyników, a także może zniechęcić obywateli do ich uznania lub wykorzystania.

W związku z tym technologia SI powinna być rozwijana w sposób, który stawia człowieka w centrum i dzięki temu jest godna zaufania publicznego. Oznacza to, że zastosowania SI powinny być zgodne nie tylko z prawem, ale również z zasadami etycznymi i gwarantować, by ich zastosowania nie powodowały niezamierzonych szkód. Na każdym etapie rozwoju SI należy zapewnić różnorodność pod względem płci, pochodzenia rasowego lub etnicznego, religii lub przekonań, niepełnosprawności i wieku. Zastosowania SI powinny wzmocnić pozycję obywateli i zapewnić poszanowanie ich praw podstawowych. Ich celem powinno być zwiększenie ludzkich możliwości, a nie ich zastąpienie, a także umożliwienie dostępu osobom z niepełnosprawnościami.

W związku z tym istnieje potrzeba opracowania **wytycznych w zakresie etyki**, które opierałyby się na istniejących ramach regulacyjnych i które powinny być stosowane przez projektantów, dostawców i użytkowników SI na rynku wewnętrznym, przez ustanowienie jednakowych zasad etycznych we wszystkich państwach członkowskich. Dlatego też Komisja powołała **grupę ekspertów wysokiego szczebla ds. SI**⁷, reprezentującą szeroki krąg zainteresowanych stron, która opracowała wytyczne w zakresie etyki dotyczące SI oraz przygotowała zestaw zaleceń dotyczących szerszej polityki w zakresie sztucznej inteligencji. Jednocześnie utworzono **Europejski sojusz na rzecz SI**⁸, otwartą platformę z udziałem ponad 2 700 członków, w celu zapewnienia szerszego wkładu w prace grupy ekspertów wysokiego szczebla ds. SI.

Grupa ekspertów wysokiego szczebla ds. SI opublikowała pierwszy projekt wytycznych w zakresie etyki w grudniu 2018 r. W oparciu o **konsultacje z zainteresowanymi stronami**⁹ oraz **spotkania z przedstawicielami państw członkowskich**¹⁰ grupa ekspertów wysokiego szczebla ds. SI przedstawiła Komisji poprawiony dokument w marcu 2019 r. W dotychczasowych informacjach zwrotnych zainteresowane strony zasadniczo przyjęły z zadowoleniem praktyczny charakter wytycznych i konkretne wytyczne oferujące projektantom, dostawcom i użytkownikom SI wskazówki na temat tego, jak zapewnić wiarygodność.

2.1. Wytyczne w zakresie etyki dotyczące wiarygodnej sztucznej inteligencji przygotowane przez grupę ekspertów wysokiego szczebla ds. SI

Wytyczne opracowane przez grupę ekspertów wysokiego szczebla ds. SI, do których odnosi się niniejszy komunikat¹¹, opierają się w szczególności na pracach Europejskiej Grupy do

⁷ <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

⁸ <https://ec.europa.eu/digital-single-market/en/european-ai-alliance>

⁹ W wyniku tych konsultacji otrzymano uwagi od 511 organizacji, stowarzyszeń, przedsiębiorstw, instytutów badawczych, osób fizycznych i innych podmiotów. Podsumowanie otrzymanych informacji zwrotnych jest dostępne na stronie internetowej: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

¹⁰ Prace grupy ekspertów zostały pozytywnie ocenione przez państwa członkowskie, zaś Rada przyjęła w dniu 18 lutego 2019 r. konkluzję, uwzględniając między innymi zbliżającą się publikację wytycznych w zakresie etyki oraz wspierając wysiłki Komisji mające na celu przeniesienia unijnego etycznego podejścia na poziom globalny: <https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/en/pdf>

¹¹ <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

spraw Etyki w Nauce i Nowych Technologiach (EGE) oraz Agencji Praw Podstawowych Unii Europejskiej (FRA).

W wytycznych stwierdzono, że aby osiągnąć „wiarygodną sztuczną inteligencję” niezbędne są trzy elementy: (1) powinna ona być zgodna z przepisami prawa (2), powinna być rozwijana z poszanowaniem zasad etycznych i (3) powinna być solidna.

W oparciu o te trzy składniki i wartości europejskie, o których mowa w sekcji 2, w wytycznych określono siedem kluczowych wymogów, które należy spełnić, aby zastosowania SI można było uznać za wiarygodne. Wytyczne zawierają również listę kontrolną pomocną w sprawdzaniu, czy wymogi te zostały spełnione.

Siedem kluczowych wymogów to:

- przewodnia i nadzorczą rolę człowieka;
- techniczna solidność i bezpieczeństwo;
- ochrona prywatności i danych;
- przejrzystość;
- różnorodność, niedyskryminacja i sprawiedliwość;
- dobrostan społeczny i środowiskowy;
- odpowiedzialność.

Podczas gdy wymogi te mają mieć zastosowanie do wszystkich systemów SI w różnych środowiskach i branżach, należy wziąć pod uwagę szczególnie kontekst, w jakim są one stosowane, z myślą o ich konkretnym i proporcjonalnym wdrożeniu z zastosowaniem podejścia opartego na oddziaływaniu. Dla przykładu, zastosowanie SI sugerujące niewłaściwą lekturę jest o wiele mniej groźne niż błędne diagnozowanie nowotworów i w związku z tym może podlegać mniej rygorystycznemu nadzorowi.

Wytyczne opracowane przez grupę ekspertów wysokiego szczebla ds. SI nie mają charakteru wiążącego i jako takie nie tworzą żadnych nowych zobowiązań prawnych. Jednak wiele istniejących przepisów prawa unijnego (dotyczących zastosowań lub obszarów) odzwierciedla już jeden lub kilka kluczowych wymogów, na przykład przepisy dotyczące bezpieczeństwa, ochrony danych osobowych, prywatności lub ochrony środowiska.

Komisja z zadowoleniem przyjmuje prace grupy ekspertów wysokiego szczebla ds. SI i uznaje jej cenny wkład w kształtowanie polityki.

2.2. Kluczowe wymogi dotyczące wiarygodnej SI

Komisja popiera następujące kluczowe wymogi dotyczące wiarygodnej SI oparte na wartościach europejskich. Zachęca zainteresowane strony do stosowania wymogów i do przetestowania listy kontrolnej służącej stworzeniu atmosfery zaufania niezbędnej do skutecznego rozwoju i stosowania SI. Komisja z zadowoleniem przyjmie informację zwrotną od zainteresowanych stron, aby ocenić, czy lista kontrolna zawarta w wytycznych wymaga dalszego dostosowania.

I. Przewodnia i nadzorczą rolę człowieka

Systemy SI powinny wspierać ludzi w dokonywaniu lepszych i bardziej świadomych wyborów zgodnych z ich celami. Powinny one pełnić rolę czynników wspierających rozwój pręźnie się rozwijającego i sprawiedliwego społeczeństwa poprzez wzmacnianie przewodniej roli człowieka i **praw podstawowych**, a nie zmniejszać, ograniczać czy wypaczać ludzką autonomię. **Ogólny dobrostan użytkownika** powinien stanowić zasadniczy sens funkcjonowania systemu.

Nadzorcza rola człowieka pomaga zagwarantować, że system SI nie podważa niezależności człowieka ani nie powoduje innych negatywnych skutków. W zależności od konkretnego systemu opartego na SI i obszaru jego stosowania należy zapewnić odpowiedni stopień **środków kontroli**, w tym zdolność dostosowania, dokładność i przejrzystość działania systemów opartych na SI¹². **Nadzór** można sprawować, stosując mechanizmy zarządzania, takie jak zapewnienie zasady udziału czynnika ludzkiego, ludzkiej interwencji lub ludzkiej kontroli¹³. Należy zapewnić, by organy publiczne miały możliwość wykonywania swoich uprawnień nadzorczych zgodnie z ich mandatami. Gdy pozostałe elementy systemów pozostają takie same, im mniejszy nadzór ma być sprawowany przez człowieka nad systemem SI, tym bardziej niezbędne są szczegółowe testy i rygorystyczne zarządzanie.

II. Techniczna solidność i bezpieczeństwo

Aby sztuczna inteligencja była wiarygodna, algorytmy muszą być bezpieczne, niezawodne i wystarczająco solidne, by poradzić sobie z błędami lub niespójnościami na wszystkich etapach cyklu życia systemu SI, a także z błędnymi wynikami końcowymi. Systemy SI muszą być **niezawodne**, wystarczająco bezpieczne czyli **odporne** zarówno na jawne ataki, jak i na bardziej subtelne próby manipulacji danymi lub algorytmami, a także muszą zapewnić **plan awaryjny** w przypadku problemów. Decyzje systemów muszą być **precyzyjne** lub przynajmniej prawidłowo odzwierciedlać swój poziom precyzji, a ich wyniki powinny być **odtworzalne**.

Ponadto systemy SI powinny być bezpieczne z założenia, a bezpieczeństwo należy uwzględniać na etapie ich projektowania, aby zagwarantować, że są one **bezpieczne w sposób możliwy do sprawdzenia** na każdym etapie, z uwzględnieniem bezpieczeństwa fizycznego i psychicznego wszystkich zainteresowanych. Obejmuje to minimalizację i, w miarę możliwości, odwracalność niezamierzonych konsekwencji lub błędów w działaniu systemu. Należy wprowadzić procedury mające na celu wyjaśnienie i ocenę potencjalnych zagrożeń związanych ze stosowaniem systemów SI w różnych obszarach zastosowań.

¹² Ogólne rozporządzenie o ochronie danych daje osobom fizycznym prawo do tego, by nie podlegać decyzji, która opiera się wyłącznie na zautomatyzowanym przetwarzaniu, i wywołuje wobec tej osoby skutki prawne lub w podobny sposób istotnie na nią wpływa (art. 22 ogólnego rozporządzenia o ochronie danych).

¹³ Zasada udziału czynnika ludzkiego (ang. *human-in-the loop*) odnosi się do ludzkiego udziału w każdym cyklu decyzyjnym systemu, który w wielu przypadkach nie jest ani możliwy, ani pożądany. Zasada ludzkiej interwencji (ang. *human-on-the-loop*) odnosi się do możliwości interwencji człowieka podczas cyklu projektowania systemu oraz monitorowania działania systemu. Zasada ludzkiej kontroli (ang. *human-in-command*) odwołuje się do możliwości nadzorowania ogólnego działania systemu SI (w tym szerszych skutków gospodarczych, społecznych, prawnych i etycznych jego zastosowania) oraz do możliwości decydowania o tym, kiedy i w jaki sposób stosować system w określonej sytuacji, aby określić różne poziomy swobody podejmowania decyzji przez człowieka podczas korzystania z systemu lub zapewnić możliwość uchylenia decyzji podjętej przez system.

III. Ochrona prywatności i danych

Należy zagwarantować **ochronę** prywatności i **danych** na **wszystkich etapach** cyklu życia systemu SI. Cyfrowe rejestry zachowań ludzi mogą umożliwić systemom SI wnioskowanie nie tylko o preferencjach poszczególnych osób, ich wieku i płci, ale także o ich orientacji seksualnej, poglądach religijnych lub politycznych. By ludzie mieli większe zaufanie do przetwarzania danych, należy zagwarantować ich pełną kontrolę nad własnymi danymi oraz że dane dotyczące tych osób nie będą wykorzystywane na ich niekorzyść lub w sposób dyskryminujący.

Oprócz ochrony prywatności i danych osobowych, konieczne jest spełnienie wymogów w celu zapewnienia wysokiej jakości systemów SI. Jakość stosowanych zbiorów danych ma zasadnicze znaczenie dla skuteczności działania systemów SI. Podczas gromadzenia danych mogą pojawić się społecznie skonstruowane uprzedzenia lub nieścisłości, błędy i pomyłki. Kwestie te należy rozwiązać przed rozpoczęciem trenowania systemu SI w oparciu o pewien zbiór danych. Ponadto należy zapewnić **integralność** danych. Stosowane procesy i zbiory danych należy testować i dokumentować na każdym etapie, czyli podczas planowania, trenowania, testowania i wdrażania. Powinno to mieć zastosowanie również do systemów SI, które nie zostały opracowane w miejscu ich zastosowania, lecz zostały nabyte poza nim. Wreszcie, **dostęp** do danych należy odpowiednio uregulować i kontrolować.

IV. Przejrzystość

Należy zapewnić **identyfikowalność** systemów SI – należy rejestrować i dokumentować zarówno decyzje podejmowane przez systemy, jak i cały proces (w tym opis gromadzenia i etykietowania danych oraz opis zastosowanego algorytmu), na podstawie którego wydano decyzje. W związku z tym należy zapewnić najbardziej szczegółową **wytłumaczalność** procesu podejmowania decyzji w oparciu o algorytm, dostosowaną do potrzeb zainteresowanych osób. Należy kontynuować badania nad opracowaniem mechanizmów wytłumaczalności. Ponadto należy wyjaśnić stopień, w jakim system SI ma wpływ na organizacyjny proces podejmowania decyzji, uzasadnić dobór projektu systemu, a także przyczyny jego zastosowania (zapewniając w ten sposób nie tylko przejrzystość danych i systemu, ale również przejrzystość modelu biznesowego).

Istotne jest również odpowiednie **informowanie** różnych zainteresowanych stron o możliwościach i ograniczeniach systemu SI w sposób właściwy dla rozpatrywanej sprawy. Ponadto systemy SI powinny być rozpoznawalne jako takie, zapewniając, że użytkownicy będą świadomi tego, że wchodzą w interakcje z systemem SI i będą wiedzieli, które osoby są za niego odpowiedzialne.

V. Różnorodność, niedyskryminacja i sprawiedliwość

Zestawy danych wykorzystywane przez systemy SI (zarówno podczas trenowania, jak i działania) mogą zostać wypaczone z powodu uwzględnienia nieumyślnych uprzedzeń historycznych, niekompletności danych i złych modeli zarządzania. Kontynuacja tego rodzaju stronniczości może prowadzić do (bez)pośredniej dyskryminacji. Szkoda może również wynikać z zamierzonego wykorzystania uprzedzeń (konsumentów) lub z nieuczciwej konkurencji. Ponadto sposób, w jaki systemy SI są projektowane (np. sposób, w jaki pisany jest kod programowania algorytmu), może również być nacechowany stronniczością. Takie problemy należy rozwiązywać na samym początku

projektowania systemu.

Tworzenie **zespołów projektowych z zagwarantowaniem różnorodności** oraz zapewnienie mechanizmów **uczestnictwa**, w szczególności uczestnictwa obywateli, może zapobiegać tego rodzaju zagrożeniom. W całym cyklu życia systemu wskazane są konsultacje z zainteresowanymi stronami, na które system może mieć bezpośredni lub pośredni wpływ. Systemy SI powinny uwzględniać cały zakres ludzkich zdolności, umiejętności i wymogów oraz zapewniać dostępność dzięki podejściu opartemu na projektowaniu uniwersalnym, dążąc do równego dostępu osób z niepełnosprawnościami.

VI. Dobrostan społeczny i środowiskowy

Aby sztuczną inteligencję uznać za godną zaufania, należy uwzględnić jej wpływ na **środowisko i inne istoty zdolne do odczuwania**. Najlepiej byłoby, gdyby wszyscy ludzie, w tym przyszłe pokolenia, mogli czerpać korzyści z różnorodności biologicznej i środowiska nadającego się do życia. W związku z tym należy wspierać zrównoważony rozwój i **odpowiedzialność ekologiczną** systemów SI. To samo dotyczy rozwiązań z zakresu SI dotyczących obszarów o globalnym znaczeniu, takich jak np. cele ONZ w zakresie zrównoważonego rozwoju.

Ponadto należy brać pod uwagę wpływ systemów SI nie tylko z perspektywy indywidualnej, lecz również z perspektywy **społeczeństwa jako całości**. Zastosowanie systemów SI powinno być dokładnie przemyślane, szczególnie w sytuacjach związanych z procesem demokratycznym, w tym w procesie opiniotwórczym, podczas podejmowania decyzji politycznych lub w kontekście wyborów. Ponadto należy rozważyć **skutki społeczne** zastosowań SI. Chociaż systemy SI mogą być wykorzystywane do poprawy umiejętności społecznych, mogą się one również przyczynić do ich pogorszenia.

VII. Odpowiedzialność

Należy wprowadzić mechanizmy zapewniające odpowiedzialność za systemy SI oraz ich wyniki, zarówno przed, jak i po ich zastosowaniu. **Możliwość kontroli** systemów SI ma w tym względzie kluczowe znaczenie, ponieważ ocena systemów SI przez audytorów wewnętrznych i zewnętrznych oraz dostępność sprawozdań z takiej oceny w dużym stopniu przyczyniają się do wiarygodności tej technologii. Należy zapewnić możliwość kontroli zewnętrznej w przypadku zastosowań mających wpływ na prawa podstawowe, w tym zastosowań o istotnym znaczeniu dla bezpieczeństwa.

Należy identyfikować, oceniać, dokumentować i ograniczać **ewentualne negatywne skutki** stosowania systemów SI. Zastosowanie ocen skutków wspomaga ten proces. Oceny te powinny być proporcjonalne do ryzyka, jakie stwarzają systemy SI. **Kompromisy** pomiędzy wymogami – których często nie można uniknąć – należy rozwiązać w sposób racjonalny i metodyczny oraz rejestrować. Wreszcie, w przypadku wystąpienia niesprawiedliwych negatywnych skutków należy przewidzieć dostępne mechanizmy zapewniające **odpowiednie środki dochodzenia roszczeń**.

2.3. Dalsze działania: faza pilotażowa z udziałem jak najszerszego kręgu zainteresowanych stron

Osiągnięcie porozumienia w sprawie tych kluczowych wymogów wobec systemów SI jest pierwszym ważnym krokiem w kierunku wypracowania wytycznych w zakresie etycznej SI.

Na następnym etapie Komisja zadba o to, by wytyczne te zostały przetestowane i wdrożone w praktyce.

W tym celu Komisja rozpocznie teraz ukierunkowany etap pilotażowy mający na celu uzyskanie uporządkowanych informacji zwrotnych od zainteresowanych stron. Działanie to będzie się koncentrować w szczególności na liście kontrolnej sporządzonej przez grupę ekspertów wysokiego szczebla dla każdego z kluczowych wymogów.

Prace te będą toczyły się dwutorowo: (i) etap pilotażowy wytycznych z udziałem zainteresowanych stron, które projektują lub wykorzystują SI, w tym organów administracji publicznej, oraz (ii) kontynuacja konsultacji z zainteresowanymi stronami i proces podnoszenia świadomości w poszczególnych państwach członkowskich i wśród różnych grup zainteresowanych stron, w tym w sektorze przemysłu i usług:

- (i) począwszy od czerwca 2019 r., wszystkie zainteresowane strony i osoby prywatne zostaną zaproszone do przetestowania listy kontrolnej i podzielenia się swoimi uwagami na temat sposobów jej poprawy. Ponadto grupa ekspertów wysokiego szczebla ds. SI opracuje szczegółową analizę wraz z zainteresowanymi stronami z sektora prywatnego i publicznego w celu zgromadzenia bardziej szczegółowych informacji zwrotnych na temat tego, w jaki sposób wytyczne można wdrażać w wielu różnych dziedzinach zastosowań. Wszystkie informacje zwrotne dotyczące funkcjonalności i wykonalności wytycznych zostaną ocenione do końca 2019 r.
- (ii) Równolegle Komisja zorganizuje dalsze działania informacyjne, umożliwiając przedstawicielom grupy ekspertów wysokiego szczebla ds. SI przedstawienie wytycznych odpowiednim zainteresowanym stronom w państwach członkowskich, w tym w sektorze przemysłu i usług, oraz zapewni tym zainteresowanym stronom dodatkową możliwość przedstawienia uwag dotyczących wytycznych w zakresie sztucznej inteligencji i wniesienia wkładu do ich treści.

Komisja weźmie pod uwagę prace grupy ekspertów ds. etyki w zakresie opartej na sieci i zautomatyzowanej mobilności¹⁴ i będzie współpracować z finansowanymi przez UE projektami badawczymi w zakresie SI oraz z odpowiednimi partnerstwami publiczno-prywatnymi w celu wdrożenia kluczowych wymogów¹⁵. Przykładowo Komisja będzie wspierać, we współpracy z państwami członkowskimi, opracowanie wspólnej bazy danych obrazów medycznych początkowo obejmującej najpowszechniejsze formy nowotworów, tak aby można było wytrenować algorytmy do diagnozowania objawów z bardzo wysoką dokładnością. Podobnie współpraca Komisji i państw członkowskich umożliwia zwiększenie liczby korytarzy transgranicznych służących testowaniu pojazdów opartych na sieci i pojazdów zautomatyzowanych. Wytyczne powinny być stosowane w tych projektach i testowane, a wyniki testów zostaną uwzględnione w procesie oceny.

Faza pilotażowa i konsultacje z zainteresowanymi stronami skorzystają z prac unijnego sojuszu na rzecz SI oraz działalności AI4EU – platformy sztucznej inteligencji na żądanie. Projekt AI4EU¹⁶, który rozpoczął się w styczniu 2019 r., gromadzi algorytmy, narzędzia, zbiory danych i usługi, wspierając organizacje, w szczególności małe i średnie przedsiębiorstwa, we wdrażaniu rozwiązań w zakresie SI. Unijny sojusz na rzecz SI, wraz z

¹⁴ Zob. komunikat Komisji w sprawie opartej na sieci i zautomatyzowanej mobilności, COM (2018) 283.

¹⁵ W ramach Europejskiego Funduszu Obronnego Komisja opracuje również szczegółowe wytyczne w zakresie etyki dotyczące oceny wniosków projektowych dotyczących zastosowań SI w dziedzinie obronności.

¹⁶ <https://ec.europa.eu/digital-single-market/en/news/artificial-intelligence-ai4eu-project-launches-1-january-2019>

AI4EU, będzie nadal aktywizować środowiska działające w obszarze SI w całej Europie, również w celu pilotowania wytycznych w zakresie etyki dotyczących SI i promowania sztucznej inteligencji ukierunkowanej na człowieka.

Na początku 2020 r. w oparciu o ocenę informacji zwrotnych otrzymanych na etapie pilotażowym **grupa ekspertów wysokiego szczebla ds. SI dokona przeglądu i aktualizacji wytycznych.** W oparciu o ten przegląd i zdobyte doświadczenia **Komisja oceni wyniki i zaproponuje dalsze działania.**

Etyczna sztuczna inteligencja jest korzystna dla wszystkich. Zagwarantowanie poszanowania podstawowych wartości i praw nie tylko jest niezbędne samo w sobie, ale także ułatwia akceptację ze strony opinii publicznej i zwiększa przewagę konkurencyjną europejskich przedsiębiorstw działających w obszarze SI poprzez stworzenie marki ukierunkowanej na człowieka wiarygodnej SI słynącej z etycznych i bezpiecznych produktów. Opinia ta opiera się na ogólnej renomie przedsiębiorstw europejskich słynących z bezpiecznych produktów wysokiej jakości. Faza pilotażowa pomoże zagwarantować, że produkty SI będą potwierdzać tę opinię.

2.4. W kierunku międzynarodowych wytycznych w zakresie etyki dotyczących SI

Międzynarodowe dyskusje na temat etyki SI stały się bardziej intensywne po tym, jak prezydencja Japonii na szczycie G-7 umieściła ten temat w porządku obrad w 2016 r. Biorąc pod uwagę międzynarodowe powiązania rozwoju SI w zakresie obiegu danych, rozwoju algorytmów i inwestycji w badania, **Komisja będzie kontynuować swoje wysiłki na rzecz przeniesienia podejścia Unii na poziom globalny i wypracowania konsensusu w sprawie SI ukierunkowanej na człowieka**¹⁷.

Prace grupy ekspertów wysokiego szczebla ds. SI, a w szczególności wykaz wymogów i proces zaangażowania zainteresowanych stron, stanowią dodatkowy wartościowy wkład w międzynarodowe dyskusje prowadzone przez Komisję. Unia Europejska może odgrywać wiodącą rolę w opracowywaniu międzynarodowych wytycznych w zakresie SI oraz, w miarę możliwości, powiązanych mechanizmów oceny.

W związku z tym Komisja zamierza:

wzmocnić współpracę z partnerami o podobnych poglądach:

- zbada, w jakim stopniu można osiągnąć zbieżność z projektami wytycznych w zakresie etyki z państw trzecich (np. z Japonii, Kanady, Singapuru) i w oparciu o tę grupę państw o podobnych poglądach, przygotuje szerszą dyskusję, wspieraną przez działania służące wdrożeniu Instrumentu Partnerstwa na rzecz Współpracy z Państwami Trzecimi¹⁸ oraz

¹⁷ Wysoki Przedstawiciel Unii do Spraw Zagranicznych i Polityki Bezpieczeństwa zamierza, przy wsparciu Komisji, wyciągnąć wnioski z konsultacji prowadzonych w ramach Organizacji Narodów Zjednoczonych, Global Tech Panel i innych wielostronnych forów, i w szczególności koordynować propozycje mające na celu sprostanie tym złożonym wyzwaniom w obszarze bezpieczeństwa.

¹⁸ Rozporządzenie Parlamentu Europejskiego i Rady (UE) nr 234/2014 z dnia 11 marca 2014 r. ustanawiające Instrument Partnerstwa na rzecz Współpracy z Państwami Trzecimi (Dz.U. L 77 z 15.3.2014, s. 77). Na przykład planowany projekt „międzynarodowego sojuszu na rzecz ukierunkowanego na człowieka podejścia do sztucznej inteligencji” ułatwi wspólne inicjatywy z partnerami o podobnych poglądach w celu promowania wytycznych w zakresie etyki oraz przyjęcia wspólnych zasad i wniosków operacyjnych. Umożliwi to UE i krajom o podobnych poglądach omawianie wniosków operacyjnych wynikających z wytycznych w zakresie etyki dotyczących SI zaproponowanych

- zbada, w jaki sposób przedsiębiorstwa z państw spoza UE i organizacje międzynarodowe mogą wnieść wkład w fazę pilotażową wytycznych poprzez testowanie i zatwierdzanie.

nadal odgrywać aktywną rolę w międzynarodowych dyskusjach i inicjatywach:

- uczestniczyć w forach wielostronnych, takich jak G-7 i G-20;
- angażować się w dialog z krajami spoza UE oraz organizować dwustronne i wielostronne spotkania w celu osiągnięcia konsensusu w sprawie SI ukierunkowanej na człowieka;
- przyczyniać się do odpowiedniej działalności normalizacyjnej w organizacjach międzynarodowych zajmujących się opracowywaniem norm w celu promowania tej wizji oraz
- wspierać gromadzenie i rozpowszechnianie informacji na temat polityki publicznej we współpracy z odpowiednimi organizacjami międzynarodowymi.

3. WNIOSKI

UE opiera się na szeregu podstawowych wartości i zbudowała na tej podstawie solidne i zrównoważone ramy regulacyjne. Konieczne jest opracowanie wytycznych w zakresie etyki dotyczących rozwoju i stosowania SI w oparciu o istniejące ramy regulacyjne z uwagi na nowatorski charakter SI oraz szczególne wyzwania, jakie niesie ze sobą ta technologia. Jedynie w sytuacji, gdy SI tworzy się i stosuje w sposób zapewniający poszanowanie dla szeroko podzielanych wartości etycznych, można ją uznać za wiarygodną.

Mając powyższe na względzie, Komisja z zadowoleniem przyjmuje wynik prac grupy ekspertów wysokiego szczebla ds. SI. W oparciu o kluczowe wymogi, jakie SI musi spełniać, by można ją uznać za wiarygodną, Komisja rozpocznie obecnie ukierunkowany etap pilotażowy w celu zapewnienia, aby wypracowane dzięki niemu wytyczne w zakresie etyki dotyczące rozwoju i stosowania SI można było wdrożyć w praktyce. Komisja będzie również dążyć do wypracowania szerokiego konsensusu społecznego w sprawie SI ukierunkowanej człowieka, w tym ze wszystkimi zainteresowanymi stronami i partnerami międzynarodowymi.

Wymiar etyczny SI nie jest luksusem czy dodatkiem: musi stanowić integralną część rozwoju SI. Dążąc do stworzenia ukierunkowanej na człowieka SI opartej na zaufaniu, zabezpieczamy poszanowanie naszych podstawowych wartości społecznych i tworzymy charakterystyczną markę Europy i jej przemysłu jako lidera w zakresie najnowocześniejszej SI, której cały świat może zaufać.

Aby zapewnić etyczny rozwój SI w Europie w szerszym kontekście, Komisja stosuje kompleksowe podejście obejmujące w szczególności następujące kierunki działań, które mają zostać wdrożone do trzeciego kwartału 2019 r.:

- W ramach programu „Horyzont 2020” Komisja rozpocznie tworzenie **sieci ośrodków doskonałości w zakresie badań naukowych nad SI**. Wybierze maksymalnie cztery sieci, skupiając się na naukowych lub technologicznych wyzwaniach, takich jak

przez grupę ekspertów w celu wypracowania wspólnego podejścia. Ponadto zapewni on monitorowanie wykorzystania technologii SI w skali globalnej. Wreszcie projekt ma na celu organizowanie działań w zakresie dyplomacji publicznej towarzyszących wydarzeniom międzynarodowym, np. przez G-7, G-20 oraz Organizację Współpracy Gospodarczej i Rozwoju.

wytłumaczalność oraz zaawansowana interakcja człowiek-maszyna, które stanowią kluczowe elementy wiarygodnej SI.

- Rozpocznie tworzenie **sieci ośrodków innowacji cyfrowych**¹⁹ skupiających się na wykorzystaniu SI przy produkcji oraz na dużych zbiorach danych.
- Wraz z państwami członkowskimi i zainteresowanymi stronami Komisja rozpocznie dyskusje przygotowawcze w celu opracowania i wdrożenia **modelu wymiany danych i najlepszego wykorzystania wspólnych przestrzeni danych**, ze szczególnym naciskiem na transport, opiekę zdrowotną i produkcję przemysłową²⁰.

Ponadto Komisja pracuje nad sprawozdaniem w sprawie wyzwań, jakie SI stwarza dla przepisów w zakresie bezpieczeństwa i odpowiedzialności oraz nad dokumentem zawierającym wytyczne dotyczące wdrażania dyrektywy w sprawie odpowiedzialności za produkty²¹. Jednocześnie Wspólne Przedsięwzięcie w dziedzinie Europejskich Obliczeń Wielkiej Skali (EuroHPC)²² stworzy kolejną generację superkomputerów, ponieważ duże zdolności obliczeniowe są niezbędne do przetwarzania danych i trenowania SI, a Europa musi opanować pełen cyfrowy łańcuch wartości. Prowadzona z państwami członkowskimi i sektorem przemysłu współpraca w obszarze mikroelektronicznych podzespołów i układów (ECSEL)²³ oraz europejska inicjatywa dotycząca procesorów²⁴ przyczynią się do rozwoju technologii procesorów o niskim zużyciu energii na potrzeby obliczeń wielkiej skali.

Podobnie jak prace nad wytycznymi w zakresie etyki dotyczącymi SI, wszystkie te inicjatywy opierają się na **ściślejszej współpracy wszystkich zainteresowanych stron**, państw członkowskich, przemysłu, podmiotów społecznych i obywateli. Podsumowując, europejskie podejście do sztucznej inteligencji pokazuje, w jaki sposób konkurencyjność gospodarcza i zaufanie społeczne muszą opierać się na tych samych podstawowych wartościach i wzajemnie się umacniać.

¹⁹ <http://s3platform.jrc.ec.europa.eu/digital-innovation-hubs>

²⁰ Niezbędne środki zostaną uruchomione z programu „Horyzont 2020” (w ramach którego na lata 2018–2020 przeznaczono blisko 1,5 mld EUR na rzecz SI) oraz jego planowanego następcy – programu „Horyzont Europa”, części cyfrowej instrumentu „Łącząc Europę”, a w szczególności z przyszłego programu „Cyfrowa Europa”. Projekty będą również czerpać z zasobów sektora prywatnego i programów państw członkowskich.

²¹ Zob. komunikat Komisji „Sztuczna inteligencja dla Europy”, COM(2018) 237.

²² <https://eurohpc-ju.europa.eu>

²³ www.ecsel.eu

²⁴ www.european-processor-initiative.eu