



Komisja Europejska

GRUPA EKSPERTÓW WYSOKIEGO SZCZEBLA DS. SZTUCZNEJ INTELIGENCJI



PROJEKT WYTYCZNYCH W ZAKRESIE ETYKI DOTYCZĄCYCH WIARYGODNEJ SZTUCZNEJ INTELIGENCJI

STRESZCZENIE

Dokument roboczy przeznaczony do konsultacji z
zainteresowanymi stronami

PROJEKT WYTYCZNYCH W ZAKRESIE ETYKI DOTYCZĄCYCH WIARYGODNEJ SZTUCZNEJ INTELIGENCJI



Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji
Projekt wytycznych w zakresie etyki dotyczących wiarygodnej sztucznej inteligencji

Komisja Europejska
Dyrekcja Generalna ds. Komunikacji Społecznej

Kontakt Nathalie Smuha – koordynator grupy ekspertów wysokiego szczebla ds. sztucznej inteligencji
Adres e-mail CNECT-HLG-AI@ec.europa.eu

European Commission
B-1049 Brussels

Dokument opublikowany w dniu 18 grudnia 2018 r. W języku angielskim.

Niniejszy dokument roboczy został opracowany przez grupę ekspertów wysokiego szczebla ds. sztucznej inteligencji bez uszczerbku dla indywidualnego stanowiska jej członków w konkretnych kwestiach i bez uszczerbku dla ostatecznej wersji dokumentu. Dokument ten będzie w późniejszym okresie zmieniany, a jego ostateczna wersja zostanie przedstawiona w marcu 2019 r., po konsultacjach z zainteresowanymi stronami za pośrednictwem europejskiego sojuszu na rzecz SI.

Ani Komisja Europejska, ani żadna osoba działająca w imieniu Komisji nie ponosi odpowiedzialności za sposób wykorzystania zamieszczonych poniżej informacji. Wyłącznie odpowiedzialność za treści zawarte w niniejszym dokumencie roboczym ponosi grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji. Personel służb Komisji służył pomocą w przygotowaniu wytycznych, niemniej jednak poglądy wyrażone w tym dokumencie odzwierciedlają opinię grupy ekspertów wysokiego szczebla ds. sztucznej inteligencji i w żadnym wypadku nie mogą być postrzegane jako oficjalne stanowisko Komisji Europejskiej. Dokument ten stanowi pierwszy rezultat prac grupy ekspertów wysokiego szczebla ds. sztucznej inteligencji. Jego ostateczna wersja zostanie przedstawiona Komisji w marcu 2019 r. Ostateczna wersja wyników realizacji drugiego zadania grupy – tj. przedstawienie zaleceń dotyczących polityki i inwestycyjnych w zakresie sztucznej inteligencji – nastąpi w połowie 2019 r.

Więcej informacji na temat grupy ekspertów wysokiego szczebla ds. sztucznej inteligencji można znaleźć w internecie (<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>).

Ponowne wykorzystanie dokumentów Komisji Europejskiej reguluje decyzja 2011/833/UE (Dz.U. L 330 z 14.12.2011, s. 39). Wykorzystywanie lub powielanie zdjęć lub innych materiałów, co do których UE nie przysługują prawa autorskie, wymaga bezpośredniej zgody właściciela praw autorskich.

STRESZCZENIE

Niniejszy dokument roboczy stanowi projekt wytycznych w zakresie etyki dotyczących sztucznej inteligencji i został opracowany przez działającą przy Komisji Europejskiej grupę ekspertów wysokiego szczebla ds. sztucznej inteligencji. Ostateczna wersja dokumentu ma zostać opublikowana w marcu 2019 r.

W dzisiejszych czasach sztuczna inteligencja (SI) jest jednym z głównych czynników przemian i bez wątpienia będzie miała istotny wpływ na tkankę społeczną. Stanowi ogromną szansę na zwiększenie dobrobytu i wzrostu gospodarczego, które Europa musi starać się osiągnąć. W ostatnim dziesięcioleciu dokonano znacznych postępów dzięki dostępności ogromnej ilości danych cyfrowych i wysoko wydajnych architektur obliczeniowych, jak również dzięki postępom w zakresie technik SI, takich jak uczenie się maszyn. Najważniejsze osiągnięcia związane z SI w zakresie pojazdów autonomicznych, opieki zdrowotnej, robotów domowych/usługowych, edukacji i cyberbezpieczeństwa codziennie poprawiają jakość naszego życia. SI ma ponadto kluczowe znaczenie dla sprostania wielu istotnym wyzwaniom, przed którymi stoi świat, takim jak zdrowie i dobrostan na świecie, zmiana klimatu, niezawodne systemy prawne i demokratyczne, jak również innym wyzwaniom wyrażonym w celach zrównoważonego rozwoju Organizacji Narodów Zjednoczonych.

SI może przynieść ogromne korzyści obywatelom i społeczeństwu, ale stwarza również pewne zagrożenia, którymi należy właściwie zarządzać. Biorąc pod uwagę fakt, że korzyści płynące z SI przewyższają związane z nią zagrożenia, musimy zadbać o wybranie takiej drogi, która **pozwole zmaksymalizować korzyści wynikające z SI, przy jednoczesnym zminimalizowaniu związanych z nią zagrożeń**. Aby to osiągnąć, **konieczne jest takie podejście do SI, w którego centrum jest człowiek** i dzięki któremu będziemy pamiętać, że rozwoju i wykorzystania SI nie należy postrzegać jako środka samego w sobie, lecz jako środka, którego celem jest zwiększenie dobrobytu człowieka. **Wiarygodna SI będzie naszym drogowskazem**, ponieważ człowiek będzie w stanie śmiało i w pełni czerpać korzyści z tej technologii jedynie wówczas, gdy będzie mógł jej zaufać.

Wiarygodna SI **obejmuje dwa elementy**: (1) powinna być stosowana z poszanowaniem praw podstawowych, obowiązujących przepisów oraz podstawowych zasad i wartości, gwarantując „**etyczny cel**”, oraz (2) powinna być **rzetelna technicznie** i niezawodna, ponieważ – nawet pomimo dobrych intencji – brak biegłości technologicznej może wyrządzić niezamierzone szkody.

W niniejszych wytycznych określono zatem **ramy wiarygodnej SI**:

- **Rozdział I** dotyczy **zapewniania etycznego celu SI** dzięki określeniu praw podstawowych, zasad i wartości, z którymi powinna być ona zgodna.
- Wychodząc od tych zasad, w **rozdziale II** przedstawiono **wytyczne w zakresie wdrażania wiarygodnej SI**, zarówno w odniesieniu do etycznego celu, jak i rzetelności technicznej. Zamieszczono w nim wykaz wymogów związanych z wiarygodną SI oraz przedstawiono przegląd metod technicznych i pozatechnicznych, które można wykorzystać do jej urzeczywistnienia.
- W **rozdziale III** dokonano **operacjonalizacji** wymogów poprzez przedstawienie konkretnej, lecz niewyczerpującej listy kontrolnej służącej do oceny wiarygodności SI. Listę tę należy następnie dostosować do konkretnych przypadków.

W przeciwieństwie do innych dokumentów poświęconych etycznej SI niniejsze wytyczne nie mają stanowić kolejnego wykazu kluczowych wartości i zasad dotyczących SI, lecz raczej zawierają wskazówki w zakresie ich konkretnego wdrożenia i operacjonalizacji w ramach systemów SI. Wytyczne uporządkowano według trzech poziomów abstrakcyjności, od poziomu najbardziej abstrakcyjnego w rozdziale I (prawa podstawowe, zasady i wartości) do poziomu najbardziej konkretnego w rozdziale III

(lista kontrolna oceny).

Wytyczne są skierowane do wszystkich **odpowiednich zainteresowanych stron tworzących, wdrażających lub wykorzystujących SI**, włączając w to przedsiębiorstwa, organizacje, naukowców, służby użyteczności publicznej, instytucje, osoby fizyczne lub inne podmioty. W ostatecznej wersji wytycznych przedstawiony zostanie mechanizm umożliwiający zainteresowanym stronom dobrowolne przyjęcie wytycznych.

Co ważne, celem niniejszych wytycznych nie jest zastąpienie jakiegokolwiek formy kształtowania polityki lub tworzenia uregulowań prawnych (kwestia ta zostanie rozpatrzona w ramach drugiego zadania realizowanego przez grupę ekspertów wysokiego szczebla ds. sztucznej inteligencji, tj. polityki i zaleceń inwestycyjnych, które mają zostać opublikowane w maju 2019 r.) ani powstrzymanie ich wprowadzania. Ponadto wytyczne należy postrzegać jako żywy dokument, który w miarę upływu czasu musi być regularnie aktualizowany w celu zapewnienia jego ciągłej adekwatności w miarę rozwoju technologii i naszej wiedzy na jej temat. Dokument ten powinien zatem stanowić punkt wyjścia do dyskusji na temat „**wiarygodnej SI wyprodukowanej w Europie**”.

Europa może promować swoje etyczne podejście do SI, jedynie utrzymując konkurencyjność na poziomie globalnym, zaś **etyczne podejście do SI ma kluczowe znaczenie dla zapewnienia odpowiedzialnej konkurencyjności**, ponieważ będzie sprzyjało budowaniu zaufania użytkowników i szerszemu stosowaniu SI. Niniejsze wytyczne nie mają na celu zahamowania innowacji w zakresie SI w Europie, lecz zakładają wykorzystanie etyki jako inspiracji do rozwoju unikalnej marki SI, która zapewni ochronę i korzyści zarówno obywatelom, jak i wspólnemu dobru. W ten sposób Europa może stać się liderem w zakresie najnowocześniejszych, bezpiecznych i etycznych systemów SI. Jedynie dzięki zapewnieniu wiarygodności obywatele Unii Europejskiej będą w pełni czerpać korzyści z SI.

Niniejsze wytyczne mają również **zachęcić do refleksji i dyskusji** na temat ram etycznych SI na **poziomie globalnym**, poza Europą.

STRESZCZENIE WYTYCZNYCH

W poszczególnych rozdziałach znajdują się wytyczne dotyczące realizacji wiarygodnej SI skierowane do wszystkich odpowiednich zainteresowanych stron tworzących, wdrażających lub wykorzystujących SI. Oto ich streszczenie:

Rozdział I: Kluczowe wytyczne dotyczące zapewnienia etycznego celu:

- Należy zapewnić, aby **człowiek znajdował się w centrum** SI: tworzeniu, wdrażaniu i wykorzystywaniu SI powinien przyświecać „**etyczny cel**”, wywodzący się z praw podstawowych, wartościach społecznych i zasadach etycznych, przynoszenie korzyści, nieszkodliwość, poszanowanie autonomii człowieka, sprawiedliwość i odpowiedzialność oraz je odzwierciedlający. Ma to kluczowe znaczenie dla pracy na rzecz tworzenia **wiarygodnej SI**.
- Dokonując prognozy możliwych skutków SI dla ludzi i dobra wspólnego, należy opierać się na prawach podstawowych, zasadach i wartościach etycznych. Należy zwracać **szczególną uwagę** na sytuacje dotyczące **słabszych grup**, takich jak dzieci, osoby z niepełnosprawnościami lub mniejszości, lub na sytuacje, w których istnieje **asymetria władzy lub dostępu do informacji**, np. między pracodawcami i pracownikami lub przedsiębiorstwami i konsumentami.
- Należy być świadomym tego, że SI, choć przynosi istotne korzyści obywatelom i społeczeństwu, może mieć również negatywne skutki. Należy zachować czujność w odniesieniu do obszarów o szczególnym znaczeniu.

Rozdział II: Kluczowe wytyczne w zakresie wdrażania wiarygodnej SI:

- Należy uwzględnić **wymogi związane z wiarygodną SI począwszy od najwcześniejszej fazy projektowania**: odpowiedzialność, zarządzanie danymi, projektowanie dla wszystkich, zarządzanie autonomią SI (nadzór ze strony człowieka), niedyskryminacja, poszanowanie autonomii ludzkiej, poszanowanie prywatności, rzetelność, bezpieczeństwo i przejrzystość.
- Należy uwzględnić metody techniczne i pozatechniczne w celu zapewnienia uwzględnienia tych wymogów w systemie SI. Wymogi te należy mieć również na uwadze podczas budowania zespołu pracującego nad systemem, jak również podczas tworzenia samego systemu, jego środowiska testowego i rozwijania potencjalnych zastosowań.
- Należy w sposób jasny i proaktywny przedstawić **zainteresowanym stronom** (klientom, pracownikom itp.) **informacje** na temat możliwości i ograniczeń systemu SI – umożliwi im to określenie realistycznych oczekiwań. Kluczowe znaczenie w tym zakresie ma zapewnienie **spójności** systemu SI.
- Należy zadbać o to, aby wiarygodna SI stała się **elementem kultury organizacji**, oraz przedstawić zainteresowanym stronom informacje na temat sposobu uwzględniania wiarygodnej SI przy projektowaniu i wykorzystywaniu systemów SI. Wiarygodną SI można również uwzględnić w deontologii lub kodeksach postępowania organizacji.
- Należy zapewnić udział i **włączenie zainteresowanych stron** w projektowanie i rozwijanie systemu SI. Ponadto należy zapewnić **różnorodność** przy tworzeniu zespołów rozwijających, wdrażających i testujących produkt.
- Należy dążyć do **ułatwienia możliwości kontroli** systemów SI, zwłaszcza w kontekstach lub sytuacjach o szczególnym znaczeniu. W miarę możliwości należy zaprojektować system w sposób umożliwiający śledzenie ścieżki wiodącej do poszczególnych decyzji od wprowadzonych uprzednio elementów, danych, wstępnie „przyuczonych” modeli itp. Ponadto należy określić **metody objaśniania** systemu SI.
- Należy zapewnić szczegółowy proces **zarządzania odpowiedzialnością**.
- Należy przewidzieć **szkolenia i edukację** oraz zadbać o to, by osoby zarządzające, opracowujące technologie, użytkownicy i pracodawcy byli świadomi i szkoleni w zakresie wiarygodnej SI.
- Należy pamiętać o tym, że mogą istnieć podstawowe rozbieżności między poszczególnymi celami (przejrzystość może otwierać drogę do nadużyć; identyfikowanie i zwalczanie uprzedzeń może kolidować z ochroną prywatności). Należy informować o związanych z tym kompromisach i dokumentować je.
- Należy wspierać badania i innowacje wspierające realizację wymogów związanych z wiarygodną SI.

Rozdział III: Kluczowe wytyczne w zakresie oceny wiarygodności SI:

- Przy rozwijaniu, wdrażaniu lub wykorzystywaniu SI należy przyjąć **listę kontrolną** służącą do oceny wiarygodności SI oraz dostosować ją do konkretnego przypadku zastosowania systemu SI
- Należy pamiętać, że lista kontrolna oceny **nigdy nie będzie wyczerpująca** i że zapewnienie wiarygodnej SI nie polega na mechanicznym zatwierdzaniu pozycji z listy, lecz na ciągłym procesie określania wymogów, oceny rozwiązań i zapewniania lepszych wyników w trakcie całego cyklu eksploatacji systemu SI.

Niniejsze wytyczne stanowią część wizji opartej na podejściu do SI, w którego centrum znajduje się człowiek i które pozwoli Europie stać się światowym innowatorem w zakresie etycznych, bezpiecznych i najnowocześniejszych systemów SI. Ich celem jest ułatwienie i umożliwienie rozwoju **„wiarygodnej SI wyprodukowanej w Europie”**, która przyczyni się do zwiększenia dobrobytu obywateli Unii Europejskiej.