



# Zalecenie UNESCO w sprawie etyki sztucznej inteligencji

przyjęte 23 listopada 2021 roku

Tłumaczenie opracowane  
przez Stały Sekretariat Polskiego Komitetu  
do spraw UNESCO

# Słowo wstępne od Polskiego Komitetu do spraw UNESCO

Szybki rozwój sztucznej inteligencji oraz jej gwałtownie rosnący wpływ na życie indywidualne i zbiorowe sprawiają, że jej etyczne i różnorodne społeczne aspekty stają się centralnymi zagadnieniami z punktu widzenia zrównoważonego rozwoju, bezpieczeństwa, równości społecznej, utrzymywania pokoju, a także przyszłości praw człowieka, demokracji i wolności osobistej, jak i dzisiejszej sytuacji w tych dziedzinach.

W 2021 roku, na mocy rezolucji 41. sesji Konferencji Generalnej<sup>1</sup>, UNESCO przyjęło Zalecenie w sprawie etyki sztucznej inteligencji. Jest to pierwszy i dotąd jedyny instrument prawny o globalnym zasięgu poświęcony tej problematyce. Jego stosowanie zwiększa możliwości zarówno rozwoju polityk krajowych, jak i międzynarodowej współpracy i dialogu.

Przyjęcie Zalecenia, w następstwie prac Grupy Ekspertów Ad Hoc, szerokich, angażujących różnorodnych interesariuszy konsultacji, a w końcu negocjacji międzyrządowych, zostało uznane za wielki sukces<sup>2</sup>. Polscy eksperci i przedstawiciele Polski aktywnie uczestniczyli w pracach nad Zaleceniem na wszystkich etapach jego powstawiania.

Zalecenie UNESCO dla państw członkowskich, zwane również rekomendacją UNESCO, jest instrumentem normatywnym o charakterze niewiążącym, stanowiącym zbiór wskazań, które państwa członkowskie wdrażają w sposób dostosowany do własnych uwarunkowań. W regularnych terminach określonych przez Konferencję Generalną UNESCO, państwa członkowskie mają obowiązek składać sprawozdania na temat implementacji Zalecenia. Na poziomie międzynarodowym Zalecenie UNESCO jest dokumentem ukierunkowującym wymianę doświadczeń, współpracę i dialog między krajami i różnymi zainteresowanymi podmiotami.

---

<sup>1</sup> 41 C/Resolution 34: <https://unesdoc.unesco.org/ark:/48223/pf0000380399> [dostęp: 10.05.2023]

<sup>2</sup> Por. strony Ethics of Artificial Intelligence w portalu UNESCO: <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics> [dostęp: 10.05.2023]

# Spis treści

Preambuła .....	5
I. Zakres stosowania .....	10
II. Cele i zadania .....	13
III. Wartości i zasady .....	15
III.1. Wartości .....	16
III.2. Zasady .....	19
IV. Obszary działania dotyczące polityk .....	26
Obszar polityki 1: Ocena oddziaływania pod względem etycznym .....	27
Obszar polityki 2: Zarządzanie zgodne z etyką .....	28
Obszar polityki 3: Polityka dotycząca danych .....	33
Obszar polityki 4: Rozwój i współpraca międzynarodowa .....	35
Obszar polityki 5: Środowisko i ekosystemy .....	36
Obszar polityki 6: Równość płci .....	38
Obszar polityki 7: Kultura .....	39
Obszar polityki 8: Edukacja i badania naukowe .....	41
Obszar polityki 9: Komunikacja i informacja .....	44
Obszar polityki 10: Gospodarka i praca .....	45
Obszar polityki 11: Zdrowie i dobrostan społeczny .....	47
V. Monitoring i ewaluacja .....	50
VI. Wykorzystanie niniejszego Zalecenia .....	52
VII. Promowanie niniejszego Zalecenia .....	53
VIII. Postanowienia końcowe .....	54



## Preambuła

Konferencja Generalna Organizacji Narodów Zjednoczonych dla Wychowania, Nauki i Kultury (UNESCO), obradująca w Paryżu na 41. sesji w dniach 9–24 listopada 2021 r.,

**Uznając** głęboki i dynamiczny, pozytywny i negatywny wpływ, jaki sztuczna inteligencja (AI) wywiera na społeczeństwa, środowisko, ekosystemy i życie ludzkie, w tym umysł człowieka, spowodowany częściowo nowymi sposobami oddziaływania jej zastosowań na myślenie, interakcje i podejmowanie decyzji oraz następstwami zastosowań sztucznej inteligencji w dziedzinie edukacji, nauk humanistycznych, społecznych i przyrodniczych, kultury, komunikacji i informacji,

**Przypominając**, że UNESCO, w myśl Aktu konstytucyjnego, dąży do budowania pokoju i bezpieczeństwa, promując współpracę między narodami poprzez edukację, naukę, kulturę, komunikację i informację w celu umacniania powszechnego poszanowania sprawiedliwości, praworządności oraz praw człowieka i podstawowych wolności, które są uznawane przez narody świata,

**W przekonaniu**, że niniejsze Zalecenie, jako instrument służący ustanawianiu standardów, reprezentujący podejście globalne, opierający się na prawie międzynarodowym i stawiający w centrum uwagi ludzką godność i prawa człowieka, a także równość płci, sprawiedliwość społeczną i ekonomiczną oraz rozwój, dobrostan fizyczny i psychiczny człowieka, różnorodność, wzajemne powiązania, włączenie społeczne oraz ochronę środowiska i ekosystemów, może ukierunkować rozwój technologii AI na taki, który będzie oparty na poczuciu odpowiedzialności,

**Kierując się** celami i zasadami zawartymi w Karcie Narodów Zjednoczonych,

**Uważając**, że technologie sztucznej inteligencji mogą się dobrze przysłużyć ludzkości i mogą być wykorzystywane we wszystkich krajach, ale jednocześnie budzą zasadnicze obawy natury etycznej, na przykład w kwestii uprzedzeń, które mogą się utrzymywać i pogłębiać, prowadząc z kolei do dyskryminacji, nierówności, powiększania się podziału cyfrowego, wykluczenia i zagrożeń dla różnorodności kulturowej, społecznej i biologicznej, do podziałów na tle społecznym lub ekonomicznym, a także w kwestii transparentności, rozumienia zasad działania algorytmów i wykorzystywania służących im danych oraz potencjalnego oddziaływania tych technologii między innymi na godność ludzką, prawa człowieka i podstawowe wolności, równość płci, demokrację, procesy

społeczne, gospodarcze, polityczne i kulturowe, praktyki naukowe i inżynierskie, dobrostan zwierząt oraz środowisko i ekosystemy,

**Uważając również**, że technologie AI mogą pogłębiać występujące na świecie podziały i nierówności w poszczególnych krajach i pomiędzy nimi, oraz że należy dbać o sprawiedliwość, zaufanie i rzetelność, tak aby żaden kraj ani żaden człowiek nie pozostawał w tyle, czy to dzięki sprawiedliwemu dostępowi i współdzieleniu korzyści z technologii AI, czy dzięki ochronie przed ich negatywnymi skutkami, przy założeniu, że w poszczególnych krajach występują różne uwarunkowania i że należy szanować wolę tych osób, które nie są zainteresowane korzystaniem ze wszystkich osiągnięć technologicznych,

**Świadoma faktu**, że wszystkie kraje stoją w obliczu przyspieszonego tempa korzystania z technologii informacyjno-komunikacyjnych i technologii sztucznej inteligencji oraz rosnącego zapotrzebowania na umiejętności korzystania z mediów i kompetencje informacyjne, a także faktu, że gospodarka cyfrowa wiąże się z poważnymi wyzwaniami społecznymi, gospodarczymi i środowiskowymi oraz możliwościami dzielenia korzyści, istotnymi zwłaszcza dla krajów o niskich i średnich dochodach, w tym między innymi krajów najslabiej rozwiniętych, rozwijających się krajów śródlądowych i małych rozwijających się państw wyspiarskich<sup>3</sup>, co wymaga poszanowania, ochrony i promowania kultur, wartości i wiedzy ludności rdzennej w celu rozwijania zrównoważonych gospodarek cyfrowych,

**Przyznając** również, że technologie AI potencjalnie mogą wywierać korzystny wpływ na środowisko i ekosystemy, ale w celu urzeczywistnienia tych korzyści należy brać pod uwagę ewentualne szkody i negatywne oddziaływanie na środowisko i ekosystemy oraz stosować odpowiednie środki zaradcze,

**Stwierdzając**, że uwzględnianie zagrożeń i problemów natury etycznej nie powinno hamować innowacyjności i rozwoju, lecz raczej stwarzać nowe możliwości oraz pobudzać taki rozwój badań i innowacyjności, który kierowałby się etyką, stawiając u podstaw technologii AI prawa człowieka i podstawowe wolności, wartości i zasady oraz refleksję moralną i etyczną,

**Przypominając również**, że w listopadzie 2019 roku Konferencja Generalna UNESCO, na swej 40. sesji, przyjęła rezolucję 40 C/37, w której upoważniła Dyrektora Generalną do „opracowania międzynarodowego instrumentu normatywnego w zakresie etyki sztucznej inteligencji (AI) w postaci zalecenia”, który miał być przedłożony Konferencji Generalnej na 41. sesji w 2021 roku,

---

<sup>3</sup> ang.: low and middle-income countries (LMICs), least developed countries (LDCs), land-locked developing countries (LLDCs), small island developing States (SIDS).

**Uznając**, że rozwój technologii opartych na sztucznej inteligencji wymaga odpowiadającej mu edukacji w zakresie korzystania z danych, mediów i informacji, a także dostępu do niezależnych, pluralistycznych i zaufanych źródeł informacji, w tym w związku z usiłowaniami ograniczania ryzyka związanego z dezinformacją, powielaniem nieprawdziwych informacji i mową nienawiści, a także szkód, które może pociągać za sobą niewłaściwe wykorzystywanie danych osobowych,

**Zauważając**, że ramy normatywne dla technologii opartych na sztucznej inteligencji i ich implikacji społecznych opierają się na międzynarodowych i krajowych ramach prawnych, prawach człowieka i podstawowych wolnościach, etyce, potrzebie dostępu do danych, informacji i wiedzy, wolności badań naukowych i innowacyjności, dobrostanie człowieka, środowiska i ekosystemu, a także, że tworzą one pomost łączący wartości i zasady etyczne z wyzwaniem i możliwościami związanymi z technologiami opartymi na sztucznej inteligencji, budowany na fundamencie porozumienia i wspólnoty celów,

**Uznając ponadto**, że wartości i zasady etyczne mogą wspierać tworzenie i wdrażanie środków politycznych i norm prawnych opartych na prawach, dostarczając wskazówek potrzebnych w związku z szybkim tempem rozwoju technologicznego,

**Wyrażając również przekonanie**, że powszechnie akceptowane standardy etyczne opracowane dla technologii AI, przy pełnym poszanowaniu prawa międzynarodowego, a w szczególności prawa dotyczącego praw człowieka, mogą odgrywać kluczową rolę w rozwoju norm związanych z AI na całym świecie,

**Mając na uwadze** Powszechną deklarację praw człowieka (1948), instrumenty międzynarodowe dotyczące praw człowieka, w tym konwencję dotyczącą statusu uchodźców (1951), konwencję dotyczącą dyskryminacji w zakresie zatrudnienia i wykonywania zawodu (1958), międzynarodową konwencję w sprawie likwidacji wszelkich form dyskryminacji rasowej (1965), międzynarodowy pakt praw obywatelskich i politycznych (1966), międzynarodowy pakt praw gospodarczych, społecznych i kulturalnych (1966), konwencję w sprawie likwidacji wszelkich form dyskryminacji kobiet (1979), konwencję o prawach dziecka (1989), konwencję o prawach osób niepełnosprawnych (2006), konwencję w sprawie zwalczania dyskryminacji w oświacie (1960), konwencję w sprawie ochrony i promowania różnorodności form wyrazu kulturowego (2005), jak również wszelkie inne właściwe w tym przypadku międzynarodowe instrumenty, zalecenia i deklaracje,

**Powołując się również na** Deklarację Narodów Zjednoczonych o prawie do rozwoju (1986), Deklarację o odpowiedzialności obecnych pokoleń wobec przyszłych pokoleń (1997), Powszechną deklarację w sprawie bioetyki i praw człowieka (2005), Deklarację

Narodów Zjednoczonych o prawach ludów tubylczych (2007), rezolucję Zgromadzenia Ogólnego Narodów Zjednoczonych w sprawie przeglądu Światowego Szczytu Społeczeństwa Informacyjnego (A/RES/70/125) (2015), rezolucję Zgromadzenia Ogólnego Narodów Zjednoczonych „Przekształcamy nasz świat: Agenda 2030 na rzecz zrównoważonego rozwoju” (A/RES/70/1) (2015), Zalecenie w sprawie zachowania i dostępu do dziedzictwa dokumentacyjnego, w tym dziedzictwa cyfrowego (2015), Deklarację zasad etycznych w odniesieniu do zmian klimatu (2017), Zalecenie dotyczące nauki i naukowców (2017), Wskaźniki uniwersalności Internetu (zatwierdzone przez Międzynarodowy Program Rozwoju Komunikacji UNESCO w 2018 r.), w tym zasady ROAM – Rights, Openness, Access and Multi-stakeholder approach (prawa, otwartość, powszechny dostęp i podejście z udziałem wielu interesariuszy) zatwierdzone przez Konferencję Generalną UNESCO w 2015 r., rezolucję Rady Praw Człowieka „Prawo do prywatności w erze cyfrowej” (A/HRC/RES/42/15) (2019) oraz rezolucję Rady Praw Człowieka „Nowe i powstające technologie cyfrowe a prawa człowieka” (A/HRC/RES/41/11) (2019),

**Podkreślając**, że szczególną uwagę należy zwrócić na kraje o niskich i średnich dochodach, w tym, ale nie tylko, na kraje najslabiej rozwinięte, rozwijające się kraje śródlądowe i małe rozwijające się państwa wyspiarskie, ponieważ dysponują one własnym potencjałem, ale nie są dostatecznie reprezentowane w debacie na temat etyki sztucznej inteligencji, co może powodować niedostateczne uwzględnianie lokalnej wiedzy, pluralizmu kulturowego, systemów wartości i wymogu powszechnej sprawiedliwości przy ocenianiu pozytywnych i negatywnych skutków zastosowań technologii sztucznej inteligencji,

**Świadoma również** istnienia wielości polityk krajowych, innych ram regulacyjnych i inicjatyw podejmowanych przez odpowiednie podmioty systemu Narodów Zjednoczonych, organizacje międzyrządowe, w tym organizacje regionalne, a także przez sektor prywatny, organizacje zawodowe, organizacje pozarządowe i środowisko naukowe, które dotyczą etyki i regulacji w zakresie technologii AI,

**Wyrażając ponadto przekonanie**, że technologie sztucznej inteligencji mogą przynieść istotne korzyści, ale ich wdrażanie może potęgować napięcia związane z innowacyjnością, asymetrią dostępu do wiedzy i technologii, w tym z deficytem umiejętności cyfrowych i obywatelskich, ograniczającym zdolność społeczeństwa do angażowania się w tematykę AI, a także ograniczać dostęp do informacji i powodować nierówności w zakresie potencjału ludzkiego i instytucjonalnego, ograniczenia w dostępie do innowacji technologicznych oraz braku w zakresie odpowiedniej infrastruktury fizycznej i cyfrowej i ram legislacyjnych, w tym prawa dotyczącego danych, oraz że z tymi wszystkimi problemami należy się zmierzyć,



**Podkreślając**, że zacieśnianie współpracy i umacnianie solidarności na poziomie globalnym, w tym poprzez multilateralizm, jest konieczne w celu zapewnienia sprawiedliwego dostępu do technologii AI i podjęcia wyzwań, jakie niosą one ze sobą dla różnorodności i powiązań między kulturami i systemami etycznymi, a także w celu przeciwdziałania potencjalnym nadużyciom, pełnego wykorzystywania potencjału AI, zwłaszcza w dziedzinie rozwoju, oraz zagwarantowania, że krajowe strategie dotyczące AI będą uwzględniały zasady etyczne,

**Biorąc w pełni pod uwagę** fakt, że szybkie tempo rozwoju technologii sztucznej inteligencji stanowi utrudnienie przy wdrażaniu i zarządzaniu tymi technologiami w sposób zgodny z zasadami etycznymi, a także wyzwanie w zakresie poszanowania i ochrony różnorodności kulturowej, oraz fakt, że rozwój tych technologii może potencjalnie naruszać normy i wartości etyczne przyjęte na poziomie lokalnym lub w danym regionie,

1. **Przyjmuje** niniejsze Zalecenie w sprawie etyki sztucznej inteligencji w dniu dwudziestego trzeciego listopada 2021 roku;
2. **Zaleca** państwom członkowskim dobrowolne stosowanie postanowień niniejszego Zalecenia poprzez podejmowanie stosownych kroków, w tym wszelkich środków prawnych lub innych, które miałyby zastosowanie w tym przypadku, zgodnie z praktyką konstytucyjną i systemem rządów w każdym państwie, w celu wprowadzania w życie, w ramach ustawodawstwa krajowego, zasad i norm zawartych w Zaleceniu, z poszanowaniem prawa międzynarodowego, w tym prawa międzynarodowego w dziedzinie praw człowieka;
3. **Zaleca również**, aby państwa członkowskie włączyły do wdrażania niniejszego Zalecenia wszystkich interesariuszy, w tym podmioty gospodarcze, w celu zapewnienia, że wszyscy oni będą odgrywali przypadającą im rolę oraz aby poleciły to Zalecenie uwadze władz, organów, organizacji naukowo-badawczych i akademickich oraz instytucji i organizacji należących do sektora publicznego, prywatnego i społeczeństwa obywatelskiego, których dotyczy sprawa technologii AI, tak aby na rozwój i zastosowania tych technologii miały wpływ zarówno rzetelne badania naukowe, jak i analiza i ocena z punktu widzenia etyki.

# I. Zakres stosowania

1. Niniejsze Zalecenie dotyczy kwestii etycznych związanych z dziedziną sztucznej inteligencji w zakresie, w jakim wiążą się one z mandatem UNESCO. Postrzega się w nim etykę AI jako systematycznie prowadzone rozważania nad standardami normatywnymi, w oparciu o holistyczne, wszechstronne, wielokulturowe i stale ewoluujące ramy wzajemnie ze sobą powiązanych wartości, zasad i działań, którymi powinny kierować się społeczeństwa, by w sposób odpowiedzialny zajmować się już znanymi i jeszcze niezbadanymi oddziaływaniami technologii AI na człowieka i społeczeństwo oraz na środowisko i ekosystemy, ramy, które dają podstawy do akceptowania lub odrzucania tych technologii. W Zaleceniu etyka jest rozumiana jako dynamiczna podstawa oceniania i ukierunkowywania technologii AI z punktu widzenia standardów normatywnych, mająca na celu poszanowanie godności ludzkiej i dobrostan człowieka oraz zapobieganie szkodom i oparta na etyce nauki i techniki.
2. Celem niniejszego Zalecenia nie jest przyjęcie jednej definicji AI, gdyż taka definicja w miarę postępu technologicznego musiałaby się stale zmieniać. Jego celem jest raczej wzięcie pod uwagę tych cech systemów AI, które mają zasadnicze znaczenie z punktu widzenia etyki. Dlatego też niniejsze Zalecenie traktuje systemy AI jako posiadające zdolność do przetwarzania danych i informacji w sposób, który przypomina inteligentne zachowanie i zazwyczaj obejmuje aspekty takie jak rozumowanie, uczenie się, postrzeganie, przewidywanie, planowanie czy kontrola. Przy takim podejściu najistotniejsze są trzy elementy:
  - a. Systemy oparte na sztucznej inteligencji (AI) są technologiami przetwarzania informacji, które integrują modele i algorytmy dające zdolność uczenia się i wykonywania zadań poznawczych prowadzących do takich wyników, jak prognozowanie i podejmowanie decyzji w środowisku fizycznym i wirtualnym. Systemy AI są przeznaczone do działań o różnym stopniu autonomii poprzez modelowanie i reprezentację wiedzy oraz poprzez wykorzystywanie danych i mierzenie współzmienności. Systemy AI mogą obejmować kilka metod, takich jak wskazane poniżej, ale nie tylko:
    - i. uczenie maszynowe, w tym uczenie głębokie i uczenie wzmacniające;
    - ii. rozumowanie maszynowe, w tym planowanie, harmonogramowanie, reprezentacja wiedzy i rozumowanie, wyszukiwanie i optymalizacja.

Systemy AI mogą być wykorzystywane w systemach cyber-fizycznych, w tym w Internecie rzeczy, systemach robotycznych, robotyce społecznej oraz w interfejsach człowiek-komputer, które obejmują funkcje sterowania, percepcji, przetwarzania danych zebranych przez czujniki oraz funkcjonowanie siłowników w środowisku związanym z systemami AI.

- b.** Kwestie etyczne dotyczące systemów AI odnoszą się do wszystkich etapów cyklu życia systemów AI, rozumianego w tym przypadku jako zakres rozpoczynający się od badań, poprzez projektowanie i rozwój po wdrażanie i używanie, w tym utrzymanie, eksploatację, komercjalizację, finansowanie, monitorowanie i ewaluację, walidację, wyłączenie z obrotu i utylizowanie. Podmioty działające w obszarze AI można zdefiniować jako wszelkie podmioty zaangażowane w co najmniej jeden etap cyklu życia systemu AI i mogą to być zarówno osoby fizyczne, jak i prawne, takie jak m.in. badacze, programiści, inżynierowie, naukowcy zajmujący się danymi, użytkownicy docelowi, przedsiębiorstwa, uczelnie wyższe oraz podmioty publiczne i prywatne.
- c.** Systemy AI powodują wyłonienie się zagadnień etycznych nowego typu, które odnoszą się między innymi, ale nie tylko, do ich wpływu na proces podejmowania decyzji, zatrudnienie i pracę, interakcje społeczne, opiekę zdrowotną, edukację, media, dostęp do informacji, podział cyfrowy, dane osobowe i ochronę konsumentów, środowisko, demokrację, rządy prawa, bezpieczeństwo i ochronę porządku publicznego, technologie podwójnego zastosowania oraz prawa człowieka i podstawowe wolności, w tym wolność wypowiedzi, prywatność i niedyskryminację. Nowe wyzwania etyczne wynikają również stąd, że algorytmy sztucznej inteligencji mogą utrwalać i pogłębiać istniejące uprzedzenia, a tym samym umacniać już istniejące formy dyskryminacji, uprzedzeń i stereotypów. Niektóre z tych kwestii są związane ze zdolnością systemów AI do wykonywania zadań, które wcześniej mogły wykonywać tylko istoty żywe, a w niektórych przypadkach wyłącznie ludzie. Cechy te nadają systemom AI istotną nową rolę w działalności człowieka i w społeczeństwie, a także w relacjach człowieka ze środowiskiem i ekosystemami, tworząc nowy kontekst, w którym dzieci i młodzież będą dorastać, dążyć do zrozumienia świata i samych siebie, krytycznie rozumieć media i informacje oraz uczyć się podejmowania decyzji. W długiej perspektywie systemy AI mogą podważać właściwe człowiekowi doświadczenie i poczucie sprawczości, wywołując dodatkowe obawy między innymi o możliwość rozumienia samego siebie oraz zachodzących interakcji społecznych, kulturowych i środowiskowych, a także obawy o zachowanie autonomii, sprawczości, poczucia wartości i godności.

- 3.** Niniejsze Zalecenie zwraca szczególną uwagę na szersze implikacje etyczne systemów AI w głównych obszarach wchodzących w zakres kompetencji UNESCO – edukacji, nauce, kulturze oraz komunikacji i informacji, co było przedmiotem wstępnego studium na temat sztucznej inteligencji, sporządzonego przez Światową Komisję UNESCO do spraw Etyki Wiedzy Naukowej i Technologii (COMEST):
- a.** w obszarze edukacji, ponieważ życie w społeczeństwach, w których cyfryzacja odgrywa coraz większą rolę, wymaga wdrożenia nowych praktyk edukacyjnych, refleksji etycznej, krytycznego myślenia, odpowiedzialności przy tworzeniu rozwiązań oraz nowych umiejętności, zważywszy na skutki dla rynku pracy, możliwości zatrudnienia i aktywności obywatelskiej.
  - b.** w obszarze nauki rozumianej w najszerszym znaczeniu i obejmującej wszystkie dziedziny akademickie od nauk przyrodniczych i medycznych po nauki społeczne i humanistyczne, ponieważ technologie AI wnoszą nowe możliwości i nowe podejście badawcze, mają wpływ na nasze koncepcje naukowego rozumienia i interpretacji oraz tworzą nowe podstawy do podejmowania decyzji.
  - c.** w obszarze tożsamości i różnorodności kulturowej, ponieważ technologie AI mogą wzbogacać rozwój przemysłów kultury i przemysłów kreatywnych, ale też mogą prowadzić do zwiększonej koncentracji podaży treści kulturowych, danych, rynków i dochodów w ręku zaledwie kilku podmiotów, co może mieć negatywne skutki dla różnorodności i pluralizmu w zakresie języków, mediów, form wyrazu kulturowego, uczestnictwa i równości.
  - d.** w obszarze komunikacji i informacji, ponieważ technologie oparte na sztucznej inteligencji odgrywają coraz większą rolę w przetwarzaniu, strukturyzowaniu i dostarczaniu informacji. Kwestie zautomatyzowanego dziennikarstwa i algorytmicznego dostarczania wiadomości oraz moderowania i przechowywania treści w mediach społecznościowych i wyszukiwarkach to tylko niektóre przykłady dotyczące zagadnień związanych między innymi z dostępem do informacji, dezinformacją, powielaniem nieprawdziwych informacji, mową nienawiści, powstawaniem nowych form narracji społecznej, dyskryminacją, wolnością słowa, prywatnością oraz edukacją medialną i informacyjną.
- 4.** Niniejsze Zalecenie jest adresowane do państw członkowskich zarówno jako do podmiotów działających w obszarze AI, jak i do podmiotów odpowiedzialnych za tworzenie ram prawnych i regulacyjnych dotyczących całego cyklu życia systemów AI oraz za promowanie odpowiedzialnego biznesu. Dostarcza ono również wytycznych dotyczących etyki wszystkim podmiotom działającym w obszarze AI, w tym sektorowi publicznemu i prywatnemu, dając podstawy do oceny etycznych aspektów oddziaływania systemów AI w całym cyklu ich życia.

## II. Cele i zadania

5. Niniejsze Zalecenie ma na celu stworzenie podstaw do takiego funkcjonowania systemów AI, które będzie służyło ludzkości, jednostkom, społeczeństwom, środowisku i ekosystemom oraz zapobiegało powstawaniu szkód. Jego celem jest również pobudzanie pokojowego wykorzystywania systemów AI.
6. Jako dopełnienie istniejących na świecie ram etycznych dotyczących AI, niniejsze Zalecenie ma stanowić powszechnie akceptowany instrument normatywny, który koncentruje się nie tylko na określeniu wartości i zasad, ale także na ich wprowadzaniu w życie poprzez konkretne rekomendacje dotyczące polityk, kładący duży nacisk na włączenie do tej problematyki kwestii równości płci oraz ochrony środowiska i ekosystemów.
7. Z uwagi na to, że złożoność kwestii etycznych związanych z AI wymaga współpracy wielu interesariuszy działających na różnych płaszczyznach i w różnych dziedzinach w ramach społeczności międzynarodowych, regionalnych i krajowych, niniejsze Zalecenie ma na celu umożliwienie zainteresowanym stronom przyjęcia na siebie współodpowiedzialności w oparciu o dialog prowadzony na poziomie globalnym z uwzględnieniem wymiaru międzykulturowego.
8. Cele niniejszego Zalecenia są następujące:
  - a. wyznaczenie uniwersalnego kanonu wartości, zasad i działań, którym państwa powinny się kierować przy tworzeniu prawa, polityk lub innych instrumentów dotyczących AI, w sposób zgodny z prawem międzynarodowym;
  - b. wskazywanie kierunku działania osobom, grupom, społecznościom, instytucjom i firmom prywatnym w celu uwzględnienia kwestii etycznych na wszystkich etapach cyklu życia systemów AI;
  - c. ochrona, promowanie i poszanowanie praw człowieka i podstawowych wolności, godności ludzkiej i równości, w tym równości płci; ochrona interesów obecnych i przyszłych pokoleń; ochrona środowiska, różnorodności biologicznej i ekosystemów oraz poszanowanie różnorodności kulturowej na wszystkich etapach cyklu życia systemu AI;
  - d. wspieranie otwartego dialogu z udziałem wielu interesariuszy i uwzględniającego wiele dyscyplin oraz dążenie do osiągnięcia konsensusu w kwestiach etycznych związanych z systemami AI;

- e. promowanie sprawiedliwego dostępu do osiągnięć i wiedzy w dziedzinie AI oraz współdzielenia korzyści, ze szczególnym uwzględnieniem potrzeb i wkładu krajów o niskich i średnich dochodach, w tym krajów najslabiej rozwiniętych, rozwijających się krajów śródlądowych oraz małych rozwijających się państw wyspiarskich.

### III. Wartości i zasady

9. Przedstawione poniżej wartości i zasady powinny być przestrzegane w pierwszym rzędzie przez wszystkie podmioty uczestniczące w cyklu życia systemów AI, a w stosownych przypadkach i zależnie od potrzeb, promowane poprzez tworzenie nowych przepisów prawnych, regulacji i wytycznych biznesowych lub uaktualnianie już istniejących dokumentów. Należy się przy tym kierować zgodnością z prawem międzynarodowym, w tym z Kartą Narodów Zjednoczonych i zobowiązaniami państw członkowskich w zakresie praw człowieka oraz uzgodnionymi na szczeblu międzynarodowym celami dotyczącymi zrównoważoności społecznej, politycznej, środowiskowej, edukacyjnej, naukowej i gospodarczej, takimi jak Cele Zrównoważonego Rozwoju Narodów Zjednoczonych.
10. Wartości odgrywają istotną rolę jako ideały motywujące tworzenie środków działania i norm prawnych. Podczas gdy całość przedstawionych poniżej wartości jest źródłem inspiracji dla właściwego postępowania i stanowi fundament zasad, zasady jako takie wyrażają te wartości w sposób bardziej konkretny poprzez deklaracje i działania związane z prowadzoną polityką.
11. Wszystkie wyszczególnione poniżej wartości i zasady są właściwe same w sobie i odnoszą się do każdego konkretnego kontekstu, niemniej niekiedy mogą między nimi pojawiać się napięcia. Każda sytuacja tego typu wymaga przeprowadzenia oceny, w której zostanie uwzględniony dany kontekst w celu poradzenia sobie z potencjalnymi napięciami. W przyjętych rozstrzygnięciach należy kierować się zasadami proporcjonalności oraz zgodności z prawami człowieka i podstawowymi wolnościami. W każdym przypadku jakiegokolwiek ograniczenia praw człowieka i podstawowych wolności muszą mieć podstawę prawną, muszą być uzasadnione, konieczne i proporcjonalne, a także zgodne z zobowiązaniami państw, wynikającymi z prawa międzynarodowego. Dokonywanie racjonalnych rozstrzygnięć będzie zazwyczaj wymagało zaangażowania szerokiego grona właściwych interesariuszy, dialogu społecznego, a także refleksji etycznej, należytej staranności i przeprowadzenia oceny oddziaływania.
12. Wiarygodność i integralność cyklu życia systemów AI są kluczowe dla funkcjonowania tych technologii dla dobra ludzkości, jednostek, społeczeństw, środowiska i ekosystemów oraz dla urzeczywistniania przez nie wartości i zasad określonych w niniejszym Zaleceniu. Ludzie powinni mieć uzasadnione powody, by nie wątpić,

że systemy AI mogą przynosić korzyść jednostkom i społeczeństwu, jeśli zostaną zastosowane odpowiednie środki ograniczające czynniki ryzyka. Zasadniczym warunkiem wiarygodności jest to, by systemy AI przez cały cykl życia podlegały starannemu monitorowaniu przez właściwych interesariuszy, stosownie do sytuacji. Z uwagi na to, że wiarygodność wynika ze stosowania zasad zawartych w niniejszym dokumencie, proponowane działania w zakresie poszczególnych polityk są ukierunkowane na promowanie wiarygodności na wszystkich etapach cyklu życia systemów AI.

## III.1. Wartości

### Poszanowanie, ochrona i promowanie praw człowieka i podstawowych wolności oraz godności ludzkiej

- 13.** Nienaruszalna i przyrodzona godność każdego człowieka jest podstawą powszechnego, niepodzielnego, niezbywalnego, współzależnego i wzajemnie powiązanego systemu praw człowieka i podstawowych wolności. Dlatego też poszanowanie, ochrona i promowanie godności ludzkiej oraz praw ustanowionych przez prawo międzynarodowe, w tym międzynarodowe prawo dotyczące praw człowieka, mają zasadnicze znaczenie w całym cyklu życia systemów AI. Pojęcie godności ludzkiej odnosi się do uznania niezbywalnej i równej wartości każdej istoty ludzkiej, bez względu na rasę, kolor skóry, urodzenie, płeć, wiek, język, religię, poglądy polityczne, kraj pochodzenia, pochodzenie etniczne, pochodzenie społeczne, ekonomiczne czy społeczne uwarunkowania dotyczące urodzenia lub niepełnosprawność i wszelkie inne przesłanki.
- 14.** Żadna istota ludzka, ani żadna społeczność ludzka nie mogą doznawać krzywdy lub być podporządkowywane, pod względem fizycznym, ekonomicznym, społecznym, politycznym, kulturowym czy psychicznym, na jakimkolwiek etapie cyklu życia systemów AI. W całym cyklu życia systemów AI celem powinno być podnoszenie jakości życia istot ludzkich, przy czym o definiowaniu «jakości życia» powinny decydować dane osoby lub grupy, o ile nie dochodzi do pogwałcenia praw człowieka i podstawowych wolności czy naruszenia godności ludzkiej w rozumieniu powyższej definicji.
- 15.** Jednostki mogą wchodzić w interakcje z systemami AI w całym ich cyklu życia i uzyskiwać od nich pomoc, taką jak opieka nad osobami słabszymi lub znajdującymi się w trudnej sytuacji, w tym między innymi opieka nad dziećmi, osobami



starszymi, osobami z niepełnosprawnościami lub chorymi. W trakcie takich interakcji osoby te nigdy nie mogą być traktowane przedmiotowo, nie może też dochodzić do naruszania ich godności osobistej w jakikolwiek inny sposób, a prawa człowieka i podstawowe wolności nie mogą być pogwałcone czy naruszane.

- 16.** Prawa człowieka i podstawowe wolności muszą być respektowane, chronione i promowane w całym cyklu życia systemów AI. Rządy, sektor prywatny, społeczeństwo obywatelskie, organizacje międzynarodowe, społeczność techniczna i środowisko akademickie muszą przestrzegać postanowień instrumentów i ram prawnych dotyczących praw człowieka, gdy uczestniczą w procesach związanych z cyklem życia systemów AI. Nowe technologie powinny dostarczać nowych sposobów promowania, obrony i egzekwowania praw człowieka, a nie ich naruszania.

### Dobro środowiska i ekosystemów

- 17.** Dobro środowiska i ekosystemów powinno być uznawane, chronione i promowane w całym cyklu życia systemów AI. Co więcej, środowisko i ekosystemy stanowią dla człowieka i innych istot żywych warunek egzystencji i bez nich czerpanie korzyści z osiągnięć AI nie byłoby możliwe.
- 18.** Wszystkie podmioty zaangażowane w cykl życia systemów AI muszą przestrzegać obowiązującego prawa międzynarodowego i ustawodawstwa krajowego, norm i praktyk, takich jak zasada przezroczności w odniesieniu do ochrony i odnawialności zasobów środowiska i ekosystemów oraz zrównoważonego rozwoju. Powinny one redukować oddziaływanie systemów AI na środowisko, w tym między innymi ślad węglowy, w celu ograniczenia do minimum czynników ryzyka związanych ze zmianami klimatycznymi i środowiskowymi oraz zapobiegając eksploatacji niezgodnej z zasadami zrównoważonego rozwoju, użytkowaniu i przekształcaniu zasobów naturalnych w sposób przyczyniający się do pogorszenia stanu środowiska i degradacji ekosystemów.

### Zapewnienie różnorodności i włączenia społecznego

- 19.** Przez cały cykl życia systemów AI należy zapewnić poszanowanie, ochronę i promowanie różnorodności i włączenia społecznego, zgodnie z prawem międzynarodowym, w tym z prawem w dziedzinie praw człowieka. Można tego dokonać poprzez promowanie aktywnego uczestnictwa wszystkich ludzi czy grup bez względu na rasę, kolor skóry, urodzenie, płeć, wiek, język, religię, poglądy polityczne, kraj pochodzenia, pochodzenie etniczne, pochodzenie społeczne, ekonomiczne

ne czy społeczne uwarunkowania dotyczące urodzenia lub niepełnosprawność i wszelkie inne przesłanki.

20. Zakres wyborów dotyczących stylu życia, przekonań, opinii, wypowiedzi lub osobistych doświadczeń, obejmujący fakultatywny charakter korzystania z systemów AI i przyczynianie się do ich tworzenia, nie może być zawężany w żadnej fazie cyklu życia systemów AI.
21. Ponadto należy dokładać starań, między innymi poprzez współpracę międzynarodową, by przewyciężyć dotkliwe dla społeczności lokalnych braki występujące zarówno w zakresie infrastruktury technicznej, edukacji i umiejętności, jak i regulacji prawnych, szczególnie w krajach o niskich i średnich dochodach, krajach najuboższych, rozwijających się krajach śródlądowych i rozwijających się małych państwach wyspiarskich i w żadnym razie nie można ich wykorzystywać.

### Życie w społeczeństwach pokojowych, sprawiedliwych i spójnych wewnętrznie

22. Podmioty działające w obszarze AI powinny odgrywać aktywną rolę w kształtowaniu się pokojowych i sprawiedliwych społeczeństw, w oparciu o przyszłość budowaną na wzajemnych powiązaniach i kierowaniu się dobrem wszystkich, w poszanowaniu praw człowieka i podstawowych wolności. Wartość życia w pokojowych i sprawiedliwych społeczeństwach wskazuje na potencjalny wkład, jaki systemy AI przez cały cykl ich życia mogą wносить w tworzenie powiązań między wszystkimi istotami żywymi i ich relacji ze środowiskiem naturalnym.
23. Pojęcie wzajemnych powiązań między ludźmi opiera się na tym, że każdy człowiek należy do większej całości, która rozwija się wtedy, gdy wszystkie jej części składowe mają możliwość rozwoju. Życie w pokojowych, sprawiedliwych i spójnych wewnętrznie społeczeństwach wymaga organicznej, bezpośredniej i bezinteresownej więzi solidarności, którą cechuje stałe poszukiwanie relacji pokojowych oraz kierowanie się troską o innych i o środowisko naturalne w najszerszym rozumieniu tego słowa.
24. Poszanowanie powyższej wartości wymaga promowania – przez cały cykl życia systemów AI – pokoju, włączenia społecznego, sprawiedliwości, równości i wzajemnych powiązań w tym sensie, że procesy związane z cyklem życia systemów AI nie mogą segregować ludzi i społeczności, uprzedmiotawiać ich lub zagrażać ich wolności, autonomii w podejmowaniu decyzji i bezpieczeństwu, ani dzielić i nastawiać jednostek i grup przeciwko sobie czy też stwarzać zagrożenia dla współistnienia ludzi, innych istot żywych i środowiska naturalnego.

## III.2. Zasady

### Proporcjonalność i zasada „nie szkodzić”

25. Należy przyjąć, że technologie AI same z siebie niekoniecznie muszą służyć dobru człowieka, środowiska i ekosystemów. Poza tym procesy związane z cyklem życia systemów AI nie powinny wykraczać poza to, co niezbędne z punktu widzenia uzasadnionych celów lub zadań i muszą być dostosowane do danej sytuacji. Gdyby się okazało, że mogą powstać jakiegokolwiek szkody, w wyniku których mogłyby ucierpieć istoty ludzkie, mogłoby dojść do naruszenia praw człowieka i podstawowych wolności czy strat, które mogłyby ponieść społeczności, społeczeństwa jako całość lub środowisko i ekosystemy, należy im przeciwdziałać poprzez zapewnienie wdrożenia procedur dotyczących oceny ryzyka i podjęcie działań prewencyjnych.
26. Decyzja o zastosowaniu systemów AI oraz o doborze metody AI powinna być uzasadniona w następujący sposób: (a) wybrana metoda powinna być właściwa i proporcjonalna z punktu widzenia przyjętego uprawnionego celu; (b) wybrana metoda nie może naruszać podstawowych wartości określonych w niniejszym dokumencie, a w szczególności jej zastosowanie nie może powodować łamania lub naruszania praw człowieka; oraz (c) metoda powinna być dobrana zależnie od kontekstu i oparta na rzetelnych podstawach naukowych. W sytuacjach, w których decyzje miałyby nieodwracalne lub trudne do odwrócenia skutki albo dotyczyłyby życia i śmierci, ostateczną decyzję musi podejmować człowiek. Systemy AI nie mogą być wykorzystywane w szczególności do takich celów, jak stosowanie punktacji społecznej lub sprawowanie masowej inwigilacji.

### Bezpieczeństwo i ochrona

27. Należy unikać niepożądanych szkód (ryzyka związanego z bezpieczeństwem) i podatności na ataki (ryzyka związanego z ochroną) oraz brać je pod uwagę, zapobiegać im i eliminować je w całym cyklu życia systemów AI w celu zapewnienia bezpieczeństwa i ochrony ludziom, środowisku i ekosystemom. Warunkiem bezpieczeństwa i ochrony w przypadku sztucznej inteligencji jest opracowanie trwałych, chroniących prywatność ram dostępu do danych, które będą sprzyjać poprawie jakości trenowania i walidacji modeli AI, wykorzystujących dane dobrej jakości.

## Bezstronność i niedyskryminacja

- 28.** Podmioty działające w obszarze AI powinny promować sprawiedliwość społeczną oraz stać na straży uczciwości i braku wszelkiej dyskryminacji, przestrzegając prawa międzynarodowego. Zakłada to podejście włączające, służące zapewnieniu, że korzyści płynące z technologii AI będą dostępne i osiągalne dla wszystkich, z uwzględnieniem szczególnych potrzeb różnych grup wiekowych, systemów kulturowych, różnych grup językowych, osób z niepełnosprawnościami, dziewcząt i kobiet oraz osób znajdujących się w niekorzystnej sytuacji, wykluczonych i podatnych na zagrożenia czy osób znajdujących się w trudnym położeniu. Państwa członkowskie powinny działać na rzecz promowania włączenia społecznego w dostępie dla wszystkich, w tym dla społeczności lokalnych, do systemów AI dostarczających treści i usług odpowiadających potrzebom lokalnym, z uwzględnieniem wielojęzyczności i różnorodności kulturowej. Państwa członkowskie powinny dokładać starań, by zająć się problemem nierówności cyfrowych i zapewnić włączenie społeczne w dostępie do AI oraz w uczestnictwie w rozwijaniu tej technologii. Na szczeblu krajowym państwa członkowskie powinny promować równość w zakresie dostępu i uczestnictwa w cyklu życia systemów AI między obszarami wiejskimi i miejskimi oraz między wszystkimi ludźmi, bez względu na rasę, kolor skóry, urodzenie, płeć, wiek, język, religię, poglądy polityczne, pochodzenie narodowe, etniczne, społeczne, ekonomiczne czy społeczne uwarunkowania dotyczące pochodzenia lub niepełnosprawność i wszelkie inne przesłanki. Na szczeblu międzynarodowym państwa najbardziej zaawansowane technologicznie są obowiązane okazywać solidarność państwom najmniej zaawansowanym w celu zapewnienia takiego podziału korzyści płynących z technologii AI, by dostęp i uczestnictwo w cyklu życia systemów AI tych ostatnich przyczyniały się do ustanowienia bardziej sprawiedliwego porządku międzynarodowego w dziedzinie informacji, komunikacji, kultury, edukacji i badań naukowych oraz stabilności społeczno-gospodarczej i politycznej.
- 29.** Podmioty działające w obszarze AI powinny w miarę możliwości dokładać wszelkich starań, by ograniczać do minimum wspieranie lub utrwalanie dyskryminacyjnych czy wypaczonych zastosowań i wyników w całym cyklu życia systemów AI oraz zapobiegać temu w celu zapewnienia bezstronności takich systemów. Powinny być dostępne skuteczne środki zaradcze przeciwdziałające dyskryminacji i nieprawidłowym decyzjom algorytmów.
- 30.** Ponadto w całym cyklu życia systemów AI należy rozwiązywać problemy dotyczące nierówności cyfrowych i nierówności w dostępie do wiedzy w poszczególnych krajach oraz pomiędzy nimi, w tym w zakresie dostępu i jakości dostępu do tech-

nologii i danych, w myśl odpowiednich ram prawnych przyjętych na szczeblu krajowym, regionalnym i międzynarodowym, jak również w zakresie konektywności, wiedzy i umiejętności oraz efektywnej partycypacji społecznej, tak by każda osoba była równo traktowana.

## Zrównoważoność

- 31.** Rozwój zrównoważonych społeczeństw opiera się na urzeczywistnianiu złożonego zbioru celów obejmujących człowieka, społeczeństwo, kulturę, gospodarkę i środowisko. Technologie AI mogą sprzyjać realizacji celów zrównoważonego rozwoju bądź ją utrudniać, w zależności od sposobu ich zastosowania w krajach o zróżnicowanym poziomie rozwoju. Dlatego też ciągle poddawanie ocenie wpływu technologii AI na człowieka, społeczeństwo, kulturę, gospodarkę i środowisko powinno w pełni uwzględniać ich implikacje dla zrównoważoności jako zbioru stale ewoluujących, wieloaspektowych celów, takich jak te, które obecnie są określone w Celach Zrównoważonego Rozwoju Narodów Zjednoczonych.

## Prawo do prywatności i ochrona danych

- 32.** Prywatność jako niezbędny warunek ochrony godności człowieka, jego autonomii i zdolności do działania musi być respektowana, chroniona i promowana w całym cyklu życia systemów AI. Jest istotne, by dane dla systemów AI były gromadzone, wykorzystywane, udostępniane, archiwizowane i usuwane w sposób zgodny z prawem międzynarodowym oraz odpowiadający wartościom i zasadom określonym w niniejszym Zaleceniu, przy jednoczesnym przestrzeganiu właściwych ram prawnych przyjętych na szczeblu krajowym, regionalnym i międzynarodowym.
- 33.** Odpowiednie ramy ochrony danych i mechanizmy zarządzania powinny być tworzone w oparciu o podejście uwzględniające udział wielu interesariuszy na szczeblu krajowym lub międzynarodowym, broniące przez systemy sądownicze i stosowane przez cały cykl życia systemów AI. Ramy ochrony danych i wszelkie związane z tym mechanizmy powinny opierać się na międzynarodowych zasadach i normach ochrony danych odnoszących się do gromadzenia, wykorzystywania i ujawniania danych osobowych oraz korzystania z przysługujących im praw przez osoby, których te dane dotyczą, przy zapewnieniu zgodności celu z prawem i ważnej podstawy prawnej do przetwarzania danych osobowych, w tym świadomej zgody.
- 34.** Systemy algorytmiczne wymagają przeprowadzania dostosowanej do nich oceny oddziaływania na prywatność, obejmującej również aspekty społeczne i etyczne, dotyczące ich stosowania oraz innowacyjnego podejścia, polegającego na posza-

nowaniu prywatności od momentu projektowania systemu. Podmioty działające w obszarze AI muszą zagwarantować, że odpowiadają za projektowanie i wdrażanie systemów AI w sposób zapewniający ochronę danych osobowych w całym cyklu życia systemów AI.

## Nadzór i decyzje w gestii człowieka

35. Państwa członkowskie powinny zapewnić, by zawsze można było przypisać osobom fizycznym lub istniejącym podmiotom prawnym odpowiedzialność etyczną i prawną za każdy etap cyklu życia systemów AI, w tym w przypadku wyrównania szkód związanych z systemami AI. Nadzór ze strony człowieka odnosi się zatem nie tylko do nadzoru indywidualnego, ale również, w określonych przypadkach, do całościowego nadzoru publicznego.
36. Może się niekiedy zdarzyć, że ludzie postanowią zdać się na systemy AI z racji ich wydajności, ale decyzja o zrzeczeniu się kontroli w jakimś ograniczonym zakresie pozostanie w gestii człowieka, gdyż ludzie mogą korzystać z systemów AI przy podejmowaniu decyzji i działań, ale system AI nigdy nie może zastąpić człowieka w ponoszeniu ostatecznej odpowiedzialności i wypełnianiu obowiązku rozliczenia się. Co do zasady, decyzje dotyczące życia i śmierci nie mogą być powierzane systemom AI.

## Transparentność i wyjaśnialność

37. Transparentność i wyjaśnialność sposobu działania systemów AI to często niezbędne warunki wstępne dla zapewnienia poszanowania, ochrony i promowania praw człowieka, podstawowych wolności i zasad etycznych. Transparentność jest niezbędna, by odpowiednie krajowe i międzynarodowe regulacje dotyczące odpowiedzialności mogły skutecznie funkcjonować. Brak transparentności może również ograniczać możliwości skutecznego kwestionowania decyzji opartych na wynikach wygenerowanych przez systemy AI, a tym samym może naruszać prawo do rzetelnego procesu sądowego i odpowiedniego zadośćuczynienia, a także zawężać obszary, w których te systemy mogą być legalnie wykorzystywane.
38. Mimo że należy dokładać starań, by zwiększać transparentność i wyjaśnialność sposobu działania systemów AI, w tym systemów o oddziaływaniu eksterytorialnym, w całym cyklu ich życia z uwagi na umacnianie demokratycznego systemu zarządzania, stopień transparentności i wyjaśnialności powinien być zawsze dostosowany do kontekstu i zakresu ich oddziaływania, gdyż może zaistnieć konieczność znalezienia równowagi między zasadami transparentności i wyjaśnialności

a innymi zasadami, takimi jak prywatność i bezpieczeństwo. Osoby fizyczne powinny być w pełni poinformowane, gdy decyzja jest podejmowana na podstawie algorytmów albo przez algorytmy, zwłaszcza wtedy, gdy ma to wpływ na ich bezpieczeństwo lub na przestrzeganie praw człowieka. W takiej sytuacji powinny one mieć możliwość zwrócenia się do odpowiedniego podmiotu działającego w obszarze AI lub do instytucji publicznej o wyjaśnienie sposobu, w jaki system AI generuje określone wyniki. Ponadto osoby fizyczne powinny mieć dostęp do uzasadnienia decyzji mającej wpływ na ich prawa i wolności, a także powinny mieć możliwość przedstawienia swoich uwag wyznaczonemu pracownikowi firmy prywatnej lub instytucji publicznej, który będzie w stanie dokonać weryfikacji i korekty decyzji. Podmioty działające w obszarze AI powinny w odpowiedni sposób i we właściwym terminie informować użytkowników, gdy produkt lub usługa są dostarczane bezpośrednio lub z pomocą systemów AI.

- 39.** Z socjotechnicznego punktu widzenia, większa transparentność przyczynia się do powstawania bardziej pokojowych, sprawiedliwych, demokratycznych i włączających społeczeństw. Umożliwia ona kontrolę publiczną, która może ograniczać korupcję i dyskryminację, a także wspierać wykrywanie i zapobieganie naruszeniom praw człowieka. Transparentność służy dostarczaniu właściwych informacji odpowiednim adresatom w celu lepszego zrozumienia i budowania zaufania. W odniesieniu do systemów AI, transparentność może umożliwiać ludziom zrozumienie, w jaki sposób wprowadzane są poszczególne etapy danego systemu AI, zależnie od kontekstu i stopnia delikatności kwestii, jakiej dotyczy. Może ona również obejmować wgląd w czynniki, które mają wpływ na konkretne prognozy lub decyzje, oraz weryfikację, czy odpowiednie gwarancje dotyczące na przykład bezpieczeństwa czy równości zostały przez system AI należycie wprowadzone. W przypadku poważnego zagrożenia w zakresie poszanowania praw człowieka, zasada transparentności może również wymagać udostępnienia kodu lub zbiorów danych.
- 40.** Wyjaśnialność (wytłumaczalność) odnosi się do możliwości zrozumienia i uzyskania wglądu w wyniki wygenerowane przez systemy AI. Zasada wyjaśnialności sztucznej inteligencji obejmuje również możliwość zrozumienia danych wejściowych, wyjściowych i sposobu funkcjonowania każdego bloku w schemacie algorytmu oraz sposobu, w jaki przyczynia się on do uzyskiwanych wyników. Wyjaśnialność jest ściśle związana z transparentnością, gdyż wyniki i procesy cząstkowe prowadzące do wyników powinny być zrozumiałe i możliwe do prześledzenia, odpowiednio do kontekstu. Podmioty działające w obszarze AI powinny zobowiązywać się do zapewnienia możliwości wytłumaczenia zasad działania danego algorytmu. W przypadku aplikacji AI oddziaływających na użytkownika docelowego

w sposób, który nie jest tymczasowy, łatwo odwracalny lub w inny sposób związany z niewielkim ryzykiem, należy zapewnić, aby każda decyzja, która doprowadziła do podjęcia określonego działania była opatrzona wyczerpującym wyjaśnieniem, tak aby wynik można było uznać za transparentny.

41. Transparentność i wyjaśnialność są ściśle związane z odpowiednimi środkami dotyczącymi odpowiedzialności i rozliczalności, jak również z wiarygodnością systemów AI.

## Odpowiedzialność i rozliczalność

42. Podmioty działające w obszarze AI i państwa członkowskie powinny szanować, chronić i promować prawa człowieka i podstawowe wolności, a także promować ochronę środowiska i ekosystemów, przyjmując na siebie w stosownym zakresie odpowiedzialność etyczną i prawną, przestrzegając prawa krajowego i międzynarodowego, w szczególności zobowiązań państw członkowskich w zakresie praw człowieka oraz wytycznych w zakresie etyki, dotyczących całego cyklu życia systemów AI, w tym w stosunku do podmiotów działających w obszarze AI na ich rzeczywistym terytorium i pod ich nadzorem. Zawsze powinna istnieć możliwość ostatecznego przypisania odpowiedzialności etycznej i odpowiedzialności za decyzje i działania oparte w jakikolwiek sposób na systemie AI podmiotom działającym w obszarze AI, stosownie do ich roli w cyklu życia danego systemu AI.
43. Należy opracować odpowiednie mechanizmy nadzoru, oceny oddziaływania, audytu i należytej staranności, w tym mechanizmy dotyczące ochrony informatorów, aby zapewnić rozliczalność systemów AI i ich oddziaływania na wszystkich etapach cyklu ich życia. Zarówno rozwiązania techniczne, jak i instytucjonalne powinny zapewniać możliwość kontroli i śledzenia (funkcjonowania) systemów AI, w szczególności w celu przewyżczenia jakichkolwiek niezgodności z normami i standardami w zakresie praw człowieka oraz zapobiegania zagrożeniom dla środowiska i ekosystemów.

## Świadomość i edukacja

44. Świadomość społeczna oraz rozumienie technologii AI i wartości, jaką przedstawiają dane, powinny być promowane poprzez otwartą i dostępną edukację, aktywność obywatelską, szkolenia z zakresu umiejętności cyfrowych i etyki AI, edukację medialną i informacyjną oraz szkolenia organizowane wspólnie przez rządy, organizacje międzyrządowe, społeczeństwo obywatelskie, środowiska akademickie, media, liderów społeczności i sektor prywatny, z uwzględnieniem róż-



norodności językowej, społecznej i kulturowej, mające na celu zapewnienie efektywnej partycypacji społecznej, tak aby wszyscy członkowie społeczeństwa mogli świadomie podejmować decyzje w sprawie korzystania z systemów AI i podlegali ochronie przed ich nadmiernym wpływem.

45. Uczenie się o oddziaływaniu systemów AI powinno obejmować naukę o prawach człowieka i podstawowych wolnościach, ich wprowadzaniu w życie i działaniu na ich rzecz, co oznacza, że podejście do systemów AI i ich rozumienie powinno się opierać na ich oddziaływaniu na prawa człowieka i na dostępie do korzystania z tych praw, a także na ich wpływie na środowisko i ekosystemy.

### Adaptacyjne zarządzanie i współpraca, z udziałem wielu podmiotów

46. Przy wykorzystywaniu danych należy przestrzegać prawa międzynarodowego i suwerenności narodowej. Oznacza to, że państwa, przestrzegając prawa międzynarodowego, mogą regulować dane generowane na ich terytorium lub przesyłane tranzytem przez ich terytorium oraz podejmować kroki zmierzające do objęcia ich skutecznymi regulacjami, w tym w zakresie ochrony danych, w oparciu o poszanowanie prawa do prywatności, zgodnie z prawem międzynarodowym oraz innymi normami i standardami dotyczącymi praw człowieka.
47. Udział różnych interesariuszy w całym cyklu życia systemów AI jest niezbędny przy włączającym podejściu do zarządzania AI, umożliwiającym powszechne współdzielenie korzyści i przyczyniającym się do zrównoważonego rozwoju. Do interesariuszy należą między innymi rządy, organizacje międzyrządowe, społeczność techniczna, społeczeństwo obywatelskie, naukowcy, środowisko akademickie, media, sektor edukacji, decydenci, podmioty gospodarcze należące do sektora prywatnego, instytucje praw człowieka i organy do spraw równości, organy monitorujące zwalczanie dyskryminacji oraz grupy działające na rzecz młodzieży i dzieci. Należy przyjąć otwarte standardy i zapewniać interoperacyjność w celu ułatwiania współpracy. Powinny być podejmowane działania, które biorą pod uwagę zachodzące przemiany technologiczne i pojawianie się nowych grup interesariuszy oraz pozwalają na znaczący udział zmarginalizowanych grup, społeczności i osób, a także, w stosownych przypadkach, zapewniają poszanowanie autonomii ludności rdzennej w zarządzaniu danymi, które jej dotyczą.

## IV. Obszary działania dotyczące polityk

48. Działania opisane w wymienionych poniżej obszarach poszczególnych polityk urzeczywistniają wartości i zasady określone w niniejszym Zaleceniu. Działalność państw członkowskich polega przede wszystkim na wprowadzaniu skutecznych środków, w tym na przykład ram lub mechanizmów strategicznych oraz na zapewnianiu ich stosowania przez innych interesariuszy, takich jak podmioty gospodarcze należące do sektora prywatnego, instytucje akademickie i badawcze oraz społeczeństwo obywatelskie, między innymi poprzez zachęcanie wszystkich zainteresowanych stron do tworzenia narzędzi oceny wpływu na prawa człowieka, rządy prawa, demokrację i etykę oraz narzędzi należytej staranności, zgodnie z istniejącymi wytycznymi, w tym Wytycznymi Narodów Zjednoczonych dotyczącymi biznesu i praw człowieka. Proces opracowywania takich polityk lub mechanizmów powinien obejmować wszystkich interesariuszy oraz uwzględniać uwarunkowania i priorytety każdego państwa członkowskiego. UNESCO może być partnerem i wspierać państwa członkowskie w tworzeniu, a także w monitorowaniu i ocenie funkcjonowania mechanizmów strategicznych.
49. UNESCO uznaje, że państwa członkowskie będą gotowe do wdrażania niniejszego Zalecenia w niejednakowym stopniu z punktu widzenia nauki, techniki, gospodarki, edukacji, prawa, regulacji, infrastruktury, uwarunkowań społecznych i innych. Warto zauważyć, że «gotowość» jest tu stanem dynamicznym. W celu umożliwienia skutecznego wdrażania niniejszego Zalecenia, UNESCO: (1) opracuje metodologię oceny stanu gotowości, aby pomóc zainteresowanym państwom członkowskim w określeniu ich statusu na poszczególnych etapach drogi obejmującej złożony zbiór celów; i (2) zapewni wsparcie zainteresowanym państwom członkowskim poprzez opracowanie metodologii UNESCO dotyczącej oceny oddziaływania technologii AI pod względem etycznym (*Ethical Impact Assessment – EIA*), a także informowanie o najlepszych praktykach, wytycznych dotyczących przeprowadzania oceny oddziaływania i innych mechanizmach oraz wynikach prac analitycznych.

## Obszar polityki 1: Ocena oddziaływania pod względem etycznym

- 50.** Państwa członkowskie powinny przyjmować ramy regulacyjne dla ocen oddziaływania, takich jak ocena oddziaływania pod względem etycznym, w celu określenia i oszacowania korzyści, problemów i zagrożeń związanych z systemami AI, jak również wprowadzać odpowiednie środki mające na celu zapobieganie, łagodzenie i monitorowanie zagrożeń, w tym mechanizmy certyfikacji. Takie oceny oddziaływania powinny określać wpływ na prawa człowieka i podstawowe wolności, w szczególności, choć nie wyłącznie, na prawa osób wykluczonych i podatnych na zagrożenia lub znajdujących się w trudnym położeniu, na prawa pracownicze, środowisko i ekosystemy oraz ujawniać implikacje etyczne i społeczne, a także ułatwiać partycypację społeczną zgodnie z wartościami i zasadami określonymi w niniejszym Zaleceniu.
- 51.** Państwa członkowskie i firmy prywatne powinny opracować mechanizmy należytej staranności i nadzoru w celu określenia zagrożeń, zapobiegania im, łagodzenia ich skutków oraz rozliczania się ze sposobu, w jaki rozwiązują problemy związane z oddziaływaniem systemów AI na poszanowanie praw człowieka, prawa i włączenie społeczne. Państwa członkowskie powinny również móc oceniać wpływ społeczno-gospodarczy systemów AI na poziom ubóstwa i zapewniać, by przepaść między osobami żyjącymi w dobrobycie, a tymi, które żyją w ubóstwie, a także nierówności cyfrowe między państwami i wewnątrz państw, nie pogłębiały się wraz ze stosowaniem technologii opartych na sztucznej inteligencji na masową skalę, obecnie i w przyszłości. W tym celu należy w szczególności wdrożyć egzekwowalne protokoły transparentności, odpowiadające dostępowi do informacji, w tym informacji leżących w interesie publicznym, które znajdują się w posiadaniu podmiotów prywatnych. Państwa członkowskie, podmioty gospodarcze w sektorze prywatnym i podmioty społeczeństwa obywatelskiego powinny badać socjologiczne i psychologiczne skutki, jakie stosowanie rekomendacji wygenerowanych w oparciu o AI może mieć dla ludzi z punktu widzenia ich autonomii decyzyjnej. Systemy AI zidentyfikowane po przeprowadzeniu oceny oddziaływania pod względem etycznym jako potencjalnie zagrażające przestrzeganiu praw człowieka, przed wprowadzeniem ich na rynek powinny być szeroko testowane przez podmioty działające w obszarze AI, w tym w razie potrzeby w warunkach rzeczywistych.
- 52.** W ramach oceny oddziaływania pod względem etycznym państwa członkowskie i przedsiębiorstwa należące do sektora usług biznesowych powinny wdrożyć od-

powiednie środki w celu monitorowania wszystkich etapów cyklu życia systemu AI, w tym funkcjonowania algorytmów wykorzystywanych do podejmowania decyzji, danych, a także podmiotów zaangażowanych w ten proces, zwłaszcza w dziedzinie usług publicznych i w przypadku konieczności bezpośredniej interakcji z użytkownikiem docelowym. Zobowiązania państw członkowskich wynikające z regulacji dotyczących praw człowieka powinny być traktowane jako jeden z aspektów etycznych przy ocenie systemów AI.

- 53.** Rządy powinny przyjąć ramy regulacyjne określające procedurę, przeznaczoną w szczególności dla organów publicznych, dotyczącą przeprowadzania oceny systemów AI pod względem etycznym w celu przewidywania konsekwencji, ograniczania zagrożeń, unikania szkodliwych skutków, ułatwiania partycypacji obywatelskiej i podejmowania wyzwań społecznych. W ramach takiej oceny należy również ustanowić odpowiednie mechanizmy nadzoru, w tym zasady sprawdzalności, identyfikowalności i wyjaśnialności, które umożliwiają ocenianie algorytmów, danych i procesów projektowych, a także dokonywanie zewnętrznego przeglądu systemów AI. Oceny skutków pod względem etycznym powinny być transparentne i w stosownych przypadkach udostępniane opinii publicznej. Takie oceny powinny być wielodyscyplinarne, otwarte i włączające, powinny angażować wiele podmiotów i obejmować różne kultury. Organy publiczne powinny być zobowiązane do monitorowania systemów AI wdrożonych i/lub tworzonych przez te organy poprzez wprowadzenie odpowiednich mechanizmów i narzędzi.

## Obszar polityki 2: Zarządzanie zgodne z etyką

- 54.** Państwa członkowskie powinny zapewnić, by mechanizmy zarządzania AI miały charakter włączający, były przejrzyste, wielodyscyplinarne, wielostronne (co stwarza możliwość ograniczania szkód i dochodzenia roszczeń również za granicą) i angażowały wiele podmiotów. Co istotne, zarządzanie powinno obejmować aspekty związane z prognozowaniem i skuteczną ochroną, monitorowaniem skutków, egzekwowaniem prawa i praktyczną realizacją zadośćuczynienia.
- 55.** Państwa członkowskie powinny zapewnić badanie i naprawianie szkód spowodowanych przez systemy oparte na sztucznej inteligencji poprzez wprowadzenie solidnych mechanizmów wykonawczych i działań naprawczych, aby zagwarantować przestrzeganie praw człowieka i podstawowych wolności oraz poszanowanie praworządności w świecie cyfrowym i fizycznym. Takie mechanizmy i działania powinny obejmować mechanizmy naprawcze pozostające w dyspozycji podmio-

tów gospodarczych działających w sektorze prywatnym i publicznym. W tym celu należy promować weryfikowalność i identyfikowalność systemów AI. Ponadto państwa członkowskie powinny wzmocnić swoje możliwości instytucjonalne, by wywiązywać się z tych zobowiązań oraz powinny współpracować z naukowcami i innymi interesariuszami w celu badania, zapobiegania i łagodzenia skutków wszelkich potencjalnie złośliwych zastosowań systemów AI.

- 56.** Zachęca się państwa członkowskie do tworzenia krajowych i regionalnych strategii dotyczących AI oraz do rozważenia rozwiązań z zakresu miękkiego zarządzania, takich jak mechanizm certyfikacji systemów AI i wzajemna uznawalność certyfikacji, zależnie od stopnia delikatności kwestii związanych z daną dziedziną i ich potencjalnego oddziaływania na prawa człowieka, środowisko i ekosystemy lub innych względów etycznych określonych w niniejszym Zaleceniu. Mechanizm taki mógłby obejmować różne poziomy audytu systemów, danych oraz przestrzegania wytycznych etycznych i wymogów proceduralnych z uwagi na aspekty etyczne. Jednocześnie nie powinien on hamować rozwoju innowacyjności, stawiać w niekorzystnej sytuacji małych i średnich przedsiębiorstw i start-upów, społeczeństwa obywatelskiego czy też organizacji badawczych i naukowych z powodu nadmiernego obciążenia obowiązkami administracyjnymi. W celu zapewnienia solidności i niezmiennej rzetelności systemu oraz zgodności z wytycznymi dotyczącymi etyki przez cały cykl życia systemu AI, mechanizmy te powinny obejmować również komponent systematycznego monitoringu, który uwzględniałby w razie potrzeby wymóg ponownej certyfikacji.
- 57.** Państwa członkowskie i organy publiczne powinny dokonywać transparentnej autoewaluacji istniejących i proponowanych systemów AI, która powinna obejmować w szczególności przeprowadzenie analizy, czy przyjęcie AI w danym przypadku jest właściwe, a jeśli tak, jaka powinna być zastosowana metoda i czy przyjęcie tej metody nie pociągałoby za sobą pogwałcenia lub naruszenia zobowiązań państw członkowskich w obszarze praw człowieka, co w takim przypadku prowadziło do zakazu jej zastosowania.
- 58.** Państwa członkowskie powinny zachęcać podmioty publiczne, podmioty gospodarcze należące do sektora prywatnego i organizacje społeczeństwa obywatelskiego do włączania różnych interesariuszy do procesu zarządzania AI oraz do rozważenia możliwości powołania niezależnego inspektora do spraw etyki AI lub ustanowienia innego mechanizmu dotyczącego nadzoru nad przeprowadzaniem oceny skutków pod względem etycznym, audytów i stałego monitoringu, a także opracowywania wytycznych w sprawie etyki systemów AI. Zachęca się państwa członkowskie, podmioty gospodarcze należące do sektora prywatnego i organiza-

cje społeczeństwa obywatelskiego, by przy wsparciu UNESCO tworzyły sieć niezależnych inspektorów do spraw etyki AI, którzy wspieraliby ten proces na poziomie krajowym, regionalnym i międzynarodowym.

- 59.** Państwa członkowskie powinny wspierać tworzenie ekosystemu cyfrowego i ułatwiać dostęp do niego w celu rozwijania na poziomie krajowym systemów AI respektujących zasady etyczne i sprzyjających włączeniu społecznemu, między innymi poprzez eliminowanie luk w dostępie do cyklu życia systemu AI, jednocześnie wnosząc wkład do współpracy międzynarodowej. Taki ekosystem obejmuje w szczególności technologie i infrastrukturę cyfrową oraz mechanizmy dzielenia się wiedzą na temat AI w zależności od potrzeb.
- 60.** Państwa członkowskie we współpracy z organizacjami międzynarodowymi, korporacjami transnarodowymi, instytucjami akademickimi i społeczeństwem obywatelskim powinny ustanowić mechanizmy zapewniające aktywny udział wszystkich państw członkowskich, w szczególności krajów o niskich i średnich dochodach, w tym krajów najsłabiej rozwiniętych, rozwijających się krajów śródlądowych i małych rozwijających się państw wyspiarskich, w międzynarodowych dyskusjach dotyczących zarządzania AI. Ten cel może być realizowany poprzez dostarczanie środków finansowych, zapewnianie równej partycypacji regionów lub za pomocą jakichkolwiek innych mechanizmów. Ponadto, dla zapewnienia powszechnego udziału w debatach dotyczących AI, państwa członkowskie powinny wspierać wyjazdy zagraniczne przedstawicieli podmiotów zajmujących się AI na terytorium danego kraju i poza nim, zwłaszcza z krajów o niskich i średnich dochodach, w tym w szczególności z krajów najsłabiej rozwiniętych, rozwijających się krajów śródlądowych i małych rozwijających się państw wyspiarskich, by mogli oni uczestniczyć w tych debatach.
- 61.** Zmiany wprowadzane w istniejących przepisach krajowych dotyczących systemów AI lub tworzone na nowo przepisy muszą być zgodne z zobowiązaniami państw członkowskich w zakresie prawa dotyczącego praw człowieka oraz promować prawa człowieka i podstawowe wolności w całym cyklu życia systemu AI. Promowanie praw człowieka i podstawowych wolności powinno również przybierać formę inicjatyw w zakresie zarządzania, modeli dobrych praktyk dotyczących współpracy w zakresie systemów AI oraz krajowych i międzynarodowych wytycznych technicznych i metodologicznych tworzonych w celu adaptacji do rozwoju technologii AI. Różne sektory, w tym sektor prywatny, w działaniach dotyczących systemów AI muszą przestrzegać, chronić i promować prawa człowieka i podstawowe wolności, wykorzystując już istniejące i nowe instrumenty w połączeniu z niniejszym Zaleceniem.

- 62.** Państwa członkowskie, które wprowadzają systemy AI przeznaczone do zastosowań wrażliwych z punktu widzenia praw człowieka, takich jak egzekwowanie prawa, opieka socjalna, zatrudnienie, media i informacja, opieka zdrowotna i system niezależnego sądownictwa, powinny ustanawiać mechanizmy monitorowania oddziaływań społecznych i gospodarczych takich systemów przez odpowiednie organy kontroli, w tym przez niezależne organy ochrony danych, nadzór branżowy i publiczne organy nadzorcze.
- 63.** Państwa członkowskie powinny wzmocnić zdolność sądownictwa do podejmowania decyzji dotyczących działania systemów AI zgodnie z zasadą praworządności, prawem i standardami międzynarodowymi, w tym w zakresie posługiwania się systemami AI w prowadzonych rozważaniach sędziowskich, przy jednoczesnym przestrzeganiu zasady dotyczącej nadzoru ze strony człowieka. W przypadku wykorzystywania systemów AI przez sądownictwo konieczne są odpowiednie zabezpieczenia, aby zagwarantować między innymi ochronę podstawowych praw człowieka, rządu prawa, niezawisłość sędziowską oraz przestrzeganie zasady nadzoru ze strony człowieka, a także, aby zapewnić rozwój i stosowanie systemów AI w sądownictwie w sposób oparty na zaufaniu, interesie publicznym i prymacie człowieka.
- 64.** Państwa członkowskie powinny zapewnić, aby rządy i organizacje wielostronne odgrywały wiodącą rolę w zapewnianiu bezpieczeństwa i ochrony systemów AI, przy udziale wielu interesariuszy. W szczególności państwa członkowskie, organizacje międzynarodowe i inne właściwe organy powinny wypracować międzynarodowe standardy, które definiują wymierne, sprawdzalne poziomy bezpieczeństwa i transparentności, tak aby można było dokonywać obiektywnej oceny systemów i określać poziomy zgodności. Ponadto państwa członkowskie i podmioty gospodarcze powinny stale wspierać badania strategiczne nad potencjalnymi zagrożeniami bezpieczeństwa ze strony technologii AI oraz zachęcać do prowadzenia badań nad transparentnością i wyjaśnialnością, włączeniem społecznym i opanowaniem tych technologii, przeznaczając dodatkowe środki na te cele w różnych dziedzinach i na różnych poziomach, na przykład w zakresie języka technicznego i języka naturalnego.
- 65.** Państwa członkowskie powinny wdrażać polityki, mające na celu zapewnienie zgodności działań podmiotów zajmujących się AI z międzynarodowymi przepisami, standardami i zasadami dotyczącymi praw człowieka w całym cyklu życia systemów AI, które jednocześnie w pełni uwzględniają istniejącą różnorodność kulturową i społeczną, w tym miejscowe zwyczaje i tradycje religijne, przy należytym poszanowaniu nadrzędności i uniwersalności praw człowieka.

- 66.** Państwa członkowskie powinny wprowadzić mechanizmy zobowiązujące podmioty działające w obszarze AI do ujawniania i zwalczania wszelkiego rodzaju stereotypów w wynikach generowanych przez systemy AI i w danych, powstałych czy to w wyniku zamierzonego działania, czy zaniedbania, oraz do zagwarantowania, by zestawy danych szkoleniowych dla systemów AI nie sprzyjały powstawaniu nierówności pod względem kulturowym, ekonomicznym czy społecznym, ugruntowywaniu się uprzedzeń, szerzeniu dezinformacji i powielaniu nieprawdziwych informacji, ani ograniczaniu wolności wypowiedzi i dostępu do informacji. Szczególną uwagę należy zwrócić na te regiony, w których nie ma wystarczającej liczby danych.
- 67.** Państwa członkowskie powinny wdrożyć politykę promowania i rozszerzania zakresu różnorodności i włączenia społecznego, który odzwierciedlałby przekrój społeczny ich populacji w zespołach zajmujących się rozwojem AI i w zestawach danych szkoleniowych, a także powinny zapewniać równy dostęp do technologii AI i płynących z nich korzyści, w szczególności grupom zmarginalizowanym, zarówno ze środowisk wiejskich, jak i miejskich.
- 68.** Państwa członkowskie powinny tworzyć, zweryfikować i ewentualnie dostosować ramy regulacyjne w celu zapewnienia rozliczalności i odpowiedzialności za treść i wyniki generowane przez systemy AI na różnych etapach cyklu ich życia. Powinny one, stosownie do potrzeb, ustalać zakres odpowiedzialności lub dostarczać jasnej interpretacji istniejących przepisów w celu przypisania odpowiedzialności za wyniki i funkcjonowanie systemów AI. Ponadto, przy tworzeniu ram regulacyjnych, państwa członkowskie powinny mieć na uwadze w szczególności fakt, iż ostatecznie odpowiedzialność i obowiązek rozliczalności musi zawsze spoczywać na osobach fizycznych lub prawnych, a systemy AI nie mogą posiadać osobowości prawnej. W związku z tym ramy regulacyjne powinny być spójne z zasadą nadzoru ze strony człowieka i opierać się na podejściu kompleksowym, skupiającym się na podmiotach AI i na procesach technologicznych zachodzących na różnych etapach cyklu życia systemów AI.
- 69.** W celu ustanowienia norm tam, gdzie one nie istnieją, lub dostosowania już istniejących ram prawnych, państwa członkowskie powinny włączyć do prowadzonych prac wszystkie podmioty działające w obszarze AI (w tym między innymi badaczy, przedstawicieli społeczeństwa obywatelskiego i organów ścigania, ubezpieczycieli, inwestorów, producentów, inżynierów, prawników i użytkowników). Normy te mogą przerodzić się w dobre praktyki oraz przepisy ustawowe i wykonawcze. Zachęca się ponadto państwa członkowskie do korzystania z mechanizmów, takich jak wzorce polityk i piaskownice regulacyjne, aby przyspieszyć tworzenie



przepisów ustawowych i wykonawczych oraz polityk, w tym dokonywanie ich regularnego przeglądu w celu nadążania za szybkim rozwojem nowych technologii, a także zapewnić możliwość przetestowania przepisów ustawowych i wykonawczych w bezpiecznym środowisku przed ich oficjalnym przyjęciem. Państwa członkowskie powinny wspierać samorządy lokalne w tworzeniu lokalnych polityk, regulacji i przepisów prawnych, zgodnych z krajowym i międzynarodowym porządkiem prawnym.

- 70.** Państwa członkowskie powinny ustanowić ścisłe wymogi dotyczące transparentności i wyjaśnialności systemów AI, by zagwarantować wiarygodność systemów AI w całym cyklu ich życia. Wymogi te powinny obejmować opracowanie i wdrożenie mechanizmów dotyczących ocen oddziaływania, które uwzględniają charakter danej dziedziny, planowane wykorzystanie i docelowych odbiorców oraz wykonalność każdego konkretnego systemu AI.

### Obszar polityki 3: Polityka dotycząca danych

- 71.** Państwa członkowskie powinny dążyć do opracowania strategii zarządzania danymi, które zapewnią ciągłą ocenę jakości danych szkoleniowych dla systemów AI, w tym ocenę adekwatności procesów gromadzenia i selekcji danych, właściwych środków bezpieczeństwa i ochrony danych, a także mechanizmów informacji zwrotnej, umożliwiających uczenie się na błędach i dzielenie się najlepszymi praktykami przez wszystkie podmioty działające w obszarze AI.
- 72.** Państwa członkowskie powinny wprowadzić odpowiednie gwarancje w celu ochrony prawa do prywatności zgodnie z prawem międzynarodowym, w tym zająć się takimi kwestiami, jak nadzór. Powinny one między innymi przyjąć lub umocnić ramy prawne zapewniające odpowiednią ochronę, zgodną z prawem międzynarodowym. Powinny usilnie zachęcać wszystkie podmioty działające w obszarze AI, w tym podmioty gospodarcze, do przestrzegania istniejących norm międzynarodowych, a w szczególności do przeprowadzania właściwej w danej sytuacji oceny oddziaływania na prywatność, w ramach ocen oddziaływania pod względem etycznym, które uwzględniają szerszy wpływ społeczno-gospodarczy planowanego przetwarzania danych, oraz do stosowania od samego początku zasady poszanowania prywatności w swoich systemach. Prywatność powinna być szanowana, chroniona i promowana w całym cyklu życia systemów AI.
- 73.** Państwa członkowskie powinny zapewnić, aby osoby fizyczne zachowały prawo do swoich danych osobowych i podlegały ochronie na podstawie ram regulacyj-

nych, które zakładają w szczególności: transparentność, odpowiednie zabezpieczenia w zakresie przetwarzania danych wrażliwych, odpowiedni poziom ochrony danych, skuteczne i spójne schematy i mechanizmy odnoszące się do kwestii odpowiedzialności, pełne korzystanie z praw przez osoby, których te dane dotyczą, oraz możliwość dostępu do własnych danych osobowych w systemach AI i do ich usunięcia, z wyjątkiem pewnych okoliczności uznawanych w prawie międzynarodowym, odpowiedni poziom ochrony w pełni odpowiadający przepisom o ochronie danych, w przypadku, gdy dane są wykorzystywane do celów komercyjnych, takich jak umożliwienie stosowania mikrotargetowania w reklamie lub przekazywane za granicę oraz skuteczny niezależny nadzór w ramach mechanizmu zarządzania danymi, zapewniający osobom fizycznym kontrolę nad swoimi danymi osobowymi, który wzmocnia korzyści płynące ze swobodnego przepływu informacji w skali międzynarodowej, w tym z dostępu do danych.

- 74.** Państwa członkowskie powinny ustanowić własne polityki dotyczące danych lub odpowiadające im ramy regulacyjne czy też wzmocnić istniejące polityki, tak by zapewnić pełne bezpieczeństwo danych osobowych i danych wrażliwych, których ujawnienie mogłoby narazić konkretne osoby na wyjątkową szkodę, cierpienie czy krzywdę. Chodzi na przykład o dane dotyczące przestępstw, postępowań karnych i wyroków skazujących oraz związanych z nimi środków bezpieczeństwa, o dane biometryczne, genetyczne i zdrowotne oraz o dane osobowe, takie jak dane dotyczące rasy, koloru skóry, urodzenia, płci, wieku, języka, religii, poglądów politycznych, pochodzenia narodowego, pochodzenia etnicznego, pochodzenia społecznego, uwarunkowań ekonomicznych czy społecznych dotyczących urodzenia lub niepełnosprawności i wszelkich innych przesłanek.
- 75.** Państwa członkowskie powinny promować otwarte dane. W tym zakresie państwa członkowskie powinny rozważyć dokonanie przeglądu swoich polityk i ram regulacyjnych, w tym w zakresie dostępu do informacji i otwartego rządu, w celu uwzględnienia wymogów związanych z AI oraz promowania mechanizmów, takich jak otwarte repozytoria danych finansowane ze środków publicznych lub znajdujących się w zasobach publicznych oraz kody źródłowe i wirtualne składnice danych, w celu wspierania między innymi bezpiecznej, sprawiedliwej, zgodnej z prawem i etyką wymiany danych.
- 76.** Państwa członkowskie powinny promować i ułatwiać korzystanie z solidnych zasobów danych wysokiej jakości na potrzeby tworzenia, rozwijania i wykorzystywania systemów opartych na sztucznej inteligencji, a także przywiązywać najwyższą wagę do nadzorowania ich gromadzenia i wykorzystywania. Mogłoby to, w miarę możliwości, obejmować inwestowanie w tworzenie wzorcowych zestawów

danych referencyjnych, które byłyby zróżnicowane i skonstruowane w oparciu o ważną podstawę prawną, w tym zgodę osób, których te dane dotyczą, gdy jest ona wymagana przez prawo. Należy zachęcać do stosowania norm dotyczących opisywania zbiorów danych, w tym rozgrupowywania danych w podziale według płci i innych kryteriów, tak by można było łatwo określić, w jaki sposób dany zbiór danych został zgromadzony i jakie ma właściwości.

- 77.** Jak wskazuje również raport Grupy Wysokiego Szczebla powołanej przez Sekretarza Generalnego ONZ na temat współpracy cyfrowej, państwa członkowskie powinny, przy wsparciu Organizacji Narodów Zjednoczonych i UNESCO, przyjąć w stosownych przypadkach podejście do danych oparte na idei cyfrowych dóbr wspólnych, zwiększyć interoperacyjność narzędzi i zbiorów danych oraz interfejsów systemów przechowujących dane, a także zachęcać podmioty gospodarcze należące do sektora prywatnego do dzielenia się z wszystkimi interesariuszami, według właściwości, danymi gromadzonymi przez nie w celach dotyczących badań naukowych, innowacyjności lub świadczeń publicznych. Powinny one również promować publiczne i prywatne starania na rzecz tworzenia platform współpracy w celu dzielenia się wysokiej jakości danymi w zaufanych i zabezpieczonych przestrzeniach danych.

## Obszar polityki 4: Rozwój i współpraca międzynarodowa

- 78.** Państwa członkowskie i korporacje transnarodowe powinny nadać etyce AI priorytetowe znaczenie poprzez włączenie debaty na temat kwestii etycznych związanych z AI do prac właściwych forów międzynarodowych i międzyrządowych oraz forów gromadzących wielu interesariuszy.
- 79.** Państwa członkowskie powinny zapewnić, by wykorzystywanie AI w takich obszarach związanych z rozwojem, jak między innymi edukacja, nauka, kultura, komunikacja i informacja, opieka zdrowotna, rolnictwo i zaopatrzenie w żywność, środowisko, zarządzanie zasobami naturalnymi i infrastrukturą, planowanie i wzrost gospodarczy, odpowiadało wartościom i zasadom określonym w niniejszym Zaleceniu.
- 80.** Państwa członkowskie powinny działać za pośrednictwem organizacji międzynarodowych w celu tworzenia platform współpracy międzynarodowej w zakresie AI na rzecz rozwoju, między innymi poprzez wnoszenie wkładu w postaci wiedzy specjalistycznej i jej finansowania, danych i infrastruktury oraz poprzez wspieranie

współpracy wielostronnej w celu rozwiązywania trudnych problemów rozwojowych, zwłaszcza w przypadku krajów o średnich i niskich dochodach, w szczególności krajów najsłabiej rozwiniętych, rozwijających się krajów śródlądowych i małych rozwijających się państw wyspiarskich.

- 81.** Państwa członkowskie powinny dążyć do promowania współpracy międzynarodowej w zakresie badań i innowacyjności w dziedzinie sztucznej inteligencji, w tym ośrodków i sieci zajmujących się badaniami i innowacjami, promujących szerszy udział i wiodącą rolę naukowców z krajów o niskich i średnich dochodach i innych krajów, w tym krajów najsłabiej rozwiniętych, rozwijających się krajów śródlądowych i małych rozwijających się państw wyspiarskich.
- 82.** Państwa członkowskie powinny wspierać badania nad etyką sztucznej inteligencji poprzez zaangażowanie organizacji międzynarodowych i instytucji badawczych, a także korporacji transnarodowych, których prace mogą tworzyć podstawy do wykorzystywania w sposób zgodny z zasadami etycznymi systemów opartych na sztucznej inteligencji przez podmioty publiczne i prywatne, w tym między innymi badania nad możliwością zastosowania określonych ram etycznych w danej kulturze i w danym kontekście oraz badania nad możliwościami opracowania wykonalnych technicznie rozwiązań zgodnych z ramami etycznymi.
- 83.** Państwa członkowskie powinny zachęcać do międzynarodowej współpracy i współdziałania w dziedzinie sztucznej inteligencji w celu eliminowania podziałów geotechnologicznych. Wymiana i konsultacje dotyczące zagadnień technologicznych powinny odbywać się między państwami członkowskimi i ich społeczeństwami, między sektorem publicznym i prywatnym oraz między krajami najbardziej i najmniej zaawansowanymi pod względem technologicznym, przy pełnym poszanowaniu prawa międzynarodowego.

## Obszar polityki 5: Środowisko i ekosystemy

- 84.** Państwa członkowskie i przedsiębiorstwa powinny badać bezpośredni i pośredni wpływ na środowisko w całym cyklu życia danego systemu AI, w tym między innymi jego ślad węglowy, zużycie energii oraz oddziaływanie związane z wydobyciem surowców do produkcji technologii AI, a także powinny ograniczać wpływ systemów AI i infrastruktur informatycznych na środowisko. Państwa członkowskie powinny zapewnić przestrzeganie prawa, polityk i praktyk w zakresie ochrony środowiska przez wszystkie podmioty działające w obszarze AI.

- 85.** Państwa członkowskie powinny wprowadzać, w zależności od potrzeb, zachęty w celu zapewnienia opracowania i przyjęcia rozwiązań opartych na poszanowaniu praw i na etyce AI, które sprzyjałyby odporności w obliczu katastrof, monitorowaniu, ochronie i odnowie środowiska i ekosystemów oraz ochronie całej planety. Takie systemy AI powinny uwzględniać udział społeczności lokalnych i ludności rdzennej w czasie całego cyklu życia systemów AI i wspierać podejścia w rodzaju przyjętych w gospodarce o obiegu zamkniętym oraz promować zrównoważone wzorce konsumpcji i produkcji. Systemy AI mogą służyć w przypadku, gdy jest to celowe i właściwe, na przykład do:
- a.** wspierania ochrony, monitorowania i zarządzania zasobami naturalnymi;
  - b.** wspierania przewidywania, zapobiegania, monitorowania i łagodzenia skutków problemów związanych z klimatem;
  - c.** wspierania bardziej wydajnego i zrównoważonego ekosystemu żywnościowego;
  - d.** wspierania przyspieszenia dostępu do zrównoważonej energii i przyjmowania jej na masową skalę;
  - e.** umożliwiania i promowania włączania zrównoważonej infrastruktury, zrównoważonych modeli biznesowych i zrównoważonego finansowania do głównego nurtu polityk na rzecz zrównoważonego rozwoju;
  - f.** wykrywania zanieczyszczeń lub prognozowania poziomu zanieczyszczeń, a tym samym udzielania pomocy właściwym interesariuszom w określaniu, planowaniu i wdrażaniu odpowiednich działań, mających na celu zapobieganie zanieczyszczeniom i narażeniu na ich skutki oraz ich ograniczanie.
- 86.** Przy doborze metod AI, zważywszy na fakt, że niektóre z nich potencjalnie mogą wymagać wielu danych lub być zasobożerne oraz wywierać wpływ na środowisko, państwa członkowskie powinny zapewnić, by stosujące je podmioty, zgodnie z zasadą proporcjonalności, preferowały metody oszczędne pod względem zużycia danych, energii i zasobów. Należy ustalić warunki, które zagwarantują uzyskanie wystarczających dowodów na to, że zastosowanie AI przyniesie zamierzony skutek, lub sprawią, że zabezpieczenia przyjęte przy zastosowaniu AI potwierdzą zasadność jej wykorzystania. Jeśli nie jest to możliwe, należy kierować się zasadą przezorności, a w przypadku, gdy mogą wystąpić negatywne skutki dla środowiska niewspółmierne do osiągniętych korzyści, sztuczna inteligencja nie powinna być stosowana.

## Obszar polityki 6: Równość płci

- 87.** Państwa członkowskie powinny zapewnić, by potencjał technologii cyfrowych i sztucznej inteligencji w zakresie zapewnienia równości płci został w pełni wykorzystany, a także, by prawa człowieka i podstawowe wolności dziewcząt i kobiet oraz ich bezpieczeństwo i integralność osobista nie były naruszane w żadnej fazie cyklu życia systemu opartego na sztucznej inteligencji. Ponadto ocena oddziaływania pod względem etycznym (EIA) powinna uwzględniać kwestie związane z płcią w ujęciu przekrojowym.
- 88.** Państwa członkowskie powinny przyznawać środki z budżetu publicznego przeznaczone na finansowanie programów dotyczących problematyki płci, zapewnić, by krajowa polityka cyfrowa obejmowała plan działania na rzecz równouprawnienia oraz kreować odpowiednie polityki, na przykład w zakresie edukacji pracowniczej, ukierunkowane na wspieranie dziewcząt i kobiet, w celu zagwarantowania, że w gospodarce cyfrowej napędzanej przez sztuczną inteligencję nie zostaną one pominięte. Należy rozważyć i wdrożyć specjalne inwestycje zmierzające do realizacji programów celowych i wprowadzenia języka dostosowanego do płci, aby zwiększyć możliwości kształcenia się dziewcząt i kobiet w zakresie nauk ścisłych, technologii, inżynierii i nauk matematycznych (STEM), w tym dyscyplin związanych z technologiami informacyjno-komunikacyjnymi (ICT), podnosić poziom ich przygotowania zawodowego i zwiększać szanse na rynku pracy oraz zapewniać równość dziewcząt i kobiet pod względem możliwości awansu i rozwoju zawodowego.
- 89.** Państwa członkowskie powinny zapewnić wykorzystanie możliwości systemów opartych na sztucznej inteligencji w dążeniu do osiągnięcia równości płci. Powinny one zapewnić, by technologie te nie pogłębiały już i tak znacznych różnic w traktowaniu płci, występujących w wielu dziedzinach w świecie analogowym, a przeciwnie – by te różnice eliminowały. Należą do nich: różnice w wynagrodzeniu kobiet i mężczyzn, nierówna reprezentacja w niektórych zawodach i rodzajach działalności, brak reprezentacji na najwyższych stanowiskach kierowniczych i w zarządach lub w zespołach badawczych zajmujących się sztuczną inteligencją, luka edukacyjna, luka w dostępie do technologii cyfrowych i sztucznej inteligencji oraz nierówności w ich przyjmowaniu, stosowaniu i przystępności cenowej, a także nierówny podział nieodpłatnej pracy i obowiązków opiekuńczych w naszych społeczeństwach.
- 90.** Państwa członkowskie powinny zapewnić, by stereotypy związane z płcią i dyskryminujące uprzedzenia nie były przenoszone do systemów sztucznej inteligencji, a przeciwnie, aby te systemy je wychwytywały i w aktywny sposób eliminowały.

Należy dołożyć starań, by uniknąć kumulowania się negatywnych skutków luk technologicznych w obszarze działań dotyczących równości płci i przeciwdziałania przemocy wobec dziewcząt, kobiet i grup niedoreprezentowanych, przybierającej takie formy jak nękanie, zastraszanie lub handel ludźmi, w tym w sferze cyfrowej.

- 91.** Państwa członkowskie powinny ułatwiać rozwój przedsiębiorczości, udział i zaangażowanie kobiet we wszystkich fazach cyklu życia systemów AI poprzez oferowanie i promowanie zachęt ekonomicznych i regulacyjnych wśród innych zachęt i systemów wsparcia, a także poprzez wdrażanie polityk mających na celu zrównoważony udział kobiet i mężczyzn w badaniach nad AI w środowisku akademickim, zapewnienie reprezentatywności pod względem płci na najwyższych stanowiskach kierowniczych w przedsiębiorstwach cyfrowych i przedsiębiorstwach zajmujących się sztuczną inteligencją, w zarządach i zespołach badawczych. Państwa członkowskie powinny zapewnić, by fundusze publiczne (przeznaczone na innowacyjność, badania i technologie) wspierały programy i przedsiębiorstwa sprzyjające włączeniu społecznemu, z widoczną reprezentatywnością pod względem płci, oraz zachęcać, by zasady pozytywnej dyskryminacji w podobny sposób były stosowane w przypadku funduszy prywatnych. Należy tworzyć i wdrażać polityki sprzyjające eliminowaniu wszelkich form nękania, jednocześnie zachęcając do stosowania najlepszych praktyk w zakresie promowania różnorodności w całym cyklu życia systemów AI.
- 92.** Państwa członkowskie powinny promować różnorodność płci w badaniach nad AI w środowisku akademickim i w przemyśle, oferując dziewczętom i kobietom zachęty do podjęcia pracy w tej dziedzinie, wprowadzając mechanizmy przeciwdziałania stereotypom związanym z płcią i molestowaniu w środowisku badaczy AI oraz zachęcając podmioty akademickie i prywatne do dzielenia się najlepszymi praktykami w zakresie wspierania różnorodności pod względem płci.
- 93.** UNESCO może pomóc w tworzeniu repozytorium najlepszych praktyk w zakresie zachęcania do udziału dziewcząt, kobiet i grup niedoreprezentowanych na wszystkich etapach cyklu życia systemów AI.

## Obszar polityki 7: Kultura

- 94.** Zachęca się państwa członkowskie do włączania, stosownie do sytuacji, systemów AI do działań na rzecz ochrony, wzbogacania, zrozumienia, promowania, zarządzania i udostępniania materialnego, dokumentacyjnego i niematerialnego dziedzictwa kulturowego, w tym języków zagrożonych, jak również języków i wiedzy

ludności rdzennej, na przykład poprzez wprowadzanie lub uaktualnianie programów edukacyjnych związanych z zastosowaniem systemów AI w tych obszarach, w których jest to właściwe oraz poprzez umacnianie podejścia partycypacyjnego w odniesieniu do instytucji i społeczeństwa.

95. Zachęca się państwa członkowskie do przeprowadzania analiz i zajęcia się kwestią wpływu kulturowego systemów sztucznej inteligencji, w szczególności aplikacji przetwarzających język naturalny, takich jak tłumaczenie automatyczne i asystenci głosowi, na niuanse językowe i formy wyrazu. Takie oceny oddziaływania powinny być brane pod uwagę przy planowaniu i wdrażaniu strategii mających na celu osiągnięcie maksymalnych korzyści ze stosowania tych systemów w zakresie redukcji nierówności na tle kulturowym i pogłębiania zrozumienia między ludźmi, a także przeciwdziałania negatywnym skutkom nierówności, takim jak ograniczenie stosowania, które może prowadzić do zaniku języków zagrożonych, lokalnych dialektów oraz odmian tonalnych i kulturowych związanych z językiem i środkami wyrazu.
96. Państwa członkowskie powinny promować edukację w zakresie AI oraz szkolenia dotyczące cyfryzacji dla artystów i osób zawodowo zajmujących się twórczością, by mogli oni ocenić, czy technologie AI nadają się do wykorzystania w ich zawodzie, oraz by mogli mieć swój udział w projektowaniu i wdrażaniu odpowiednich technologii AI, wykorzystywanych obecnie do tworzenia, produkcji, dystrybucji, rozpowszechniania i konsumpcji wielu dóbr i usług kulturalnych, mając na uwadze znaczenie zachowania dziedzictwa kulturowego, różnorodności kulturowej i wolności artystycznej.
97. Państwa członkowskie powinny upowszechniać wiedzę na temat narzędzi AI i zachęcać do ich testowania lokalne przemysły kulturalne oraz małe i średnie przedsiębiorstwa, działające w sferze kultury, w celu przeciwdziałania koncentracji na rynku kultury.
98. Państwa członkowskie powinny włączać firmy technologiczne i innych interesariuszy do promowania zróżnicowanej oferty i szerokiego dostępu do form wyrazu kulturowego, w szczególności w celu zapewnienia, by rekomendacje algorytmiczne zwiększały widoczność i możliwości odkrywania treści lokalnych.
99. Państwa członkowskie powinny wspierać nowe badania na styku dziedzin – sztucznej inteligencji i własności intelektualnej – na przykład w celu ustalenia, czy lub jak chronić na podstawie prawa własności intelektualnej utwory stworzone przy użyciu technologii AI. Państwa członkowskie powinny również oceniać, w jaki sposób technologie AI wpływają na prawa lub interesy właścicieli praw własności



intelektualnej, których dzieła są wykorzystywane do badań, rozwoju, trenowania lub wdrażania aplikacji AI.

- 100.** Państwa członkowskie powinny zachęcać krajowe muzea, galerie, biblioteki i archiwa do wykorzystywania systemów opartych na sztucznej inteligencji w eksponowaniu zbiorów oraz rozbudowywaniu bibliotek, baz danych i bazy wiedzy, zapewniając jednocześnie ich dostępność dla użytkowników.

## Obszar polityki 8: Edukacja i badania naukowe

- 101.** Państwa członkowskie powinny współpracować z organizacjami międzynarodowymi, instytucjami edukacyjnymi oraz podmiotami prywatnymi i pozarządowymi w celu zapewnienia społeczeństwu możliwości kształcenia się w zakresie AI na wszystkich poziomach, we wszystkich krajach, tak by wyposażyć ludzi w potrzebną wiedzę i umiejętności oraz zmniejszyć podziały i nierówności w dostępie do technologii cyfrowych, wynikające z szerokiego stosowania systemów AI.
- 102.** Państwa członkowskie powinny ułatwiać nabywanie „niezbędnych umiejętności” służących edukacji w zakresie AI, takich jak podstawowe umiejętności w zakresie czytania, pisania, liczenia, programowania i znajomości technologii cyfrowych oraz umiejętności korzystania z mediów i informacji, a także umiejętności krytycznego i kreatywnego myślenia, pracy zespołowej, komunikacji, kompetencji społeczno-emocjonalnych i etycznych dotyczących AI, zwłaszcza w tych krajach i regionach lub na takich obszarach w obrębie poszczególnych krajów, gdzie odczuwalne są znaczne braki w kształceniu w zakresie tych umiejętności.
- 103.** Państwa członkowskie powinny promować programy służące podnoszeniu ogólnego poziomu wiedzy na temat osiągnięć w zakresie AI, w tym na temat danych oraz możliwości i wyzwań związanych z technologiami AI, oddziaływania systemów AI na prawa człowieka i ich implikacji, między innymi w zakresie praw dziecka. Powinny to być programy dostępne zarówno dla osób nie mających wykształcenia technicznego, jak i dla techników i inżynierów.
- 104.** Państwa członkowskie powinny wspierać inicjatywy badawcze dotyczące odpowiedzialnego i etycznego wykorzystywania technologii sztucznej inteligencji między innymi w nauczaniu, szkoleniu nauczycieli i nauce zdalnej, aby zwiększyć korzyści i ograniczyć zagrożenia w tej dziedzinie. Inicjatywom tym powinna towarzyszyć przeprowadzana we właściwy sposób ocena jakości kształcenia

i ocena wpływu, jaki stosowanie technologii AI wywiera na uczniów i nauczycieli. Państwa członkowskie powinny również zapewnić, by technologie AI służyły upodmiotowieniu uczniów i nauczycieli oraz by wzbogacały ich doświadczenia. Powinny mieć przy tym na uwadze fakt, że aspekty relacyjne i społeczne oraz wartość tradycyjnych form kształcenia mają zasadnicze znaczenie zarówno w relacjach między nauczycielem i uczniem, jak i we wzajemnych relacjach między uczniami i powinny być uwzględniane przy rozpatrywaniu kwestii posługiwania się technologiami AI w edukacji. Systemy AI wykorzystywane w nauczaniu powinny spełniać ścisłe wymogi, jeśli chodzi o monitorowanie, ocenę umiejętności lub przewidywanie zachowania osób uczących się. Sztuczna inteligencja powinna wspierać proces uczenia się bez ograniczania zdolności poznawczych i bez pozyskiwania informacji wrażliwych, zgodnie z właściwymi standardami ochrony danych osobowych. Dane przekazane w związku ze zdobywaniem wiedzy, które zostały zgromadzone w wyniku interakcji między uczącym się a systemem AI, nie mogą być przedmiotem niewłaściwego zastosowania, sprzeniewierzenia lub wykorzystania w celach przestępczych, ani w celach komercyjnych.

- 105.** Państwa członkowskie powinny promować pełnienie roli przywódczej przez dziewczęta i kobiety oraz ich uczestnictwo, podobnie jak przedstawicieli różnych grup etnicznych i kulturowych, osób z niepełnosprawnościami, osób wykluczonych społecznie lub znajdujących się w trudnej sytuacji, należących do mniejszości oraz wszystkich osób, które nie czerpią pełnych korzyści z włączenia cyfrowego, w programach edukacyjnych na temat AI na wszystkich poziomach, a także monitorować najlepsze praktyki w tym zakresie i dzielić się nimi z innymi państwami członkowskimi.
- 106.** Państwa członkowskie powinny tworzyć, w sposób odpowiadający krajowym programom i tradycjom oświatowym, programy nauczania na temat etyki AI na wszystkich poziomach, a także promować powiązanie kształcenia w zakresie umiejętności technicznych AI z kształceniem na temat humanistycznych, etycznych i społecznych aspektów AI. Kursy online i zasoby cyfrowe dotyczące kształcenia w zakresie etyki AI powinny być opracowywane w językach lokalnych, w tym w językach rdzennych i uwzględniać różnorodność środowisk, zwłaszcza kwestię dostępności materiałów dla osób z niepełnosprawnościami.
- 107.** Państwa członkowskie powinny promować i wspierać badania nad AI, w szczególności badania poświęcone etyce AI, w tym na przykład poprzez inwestowanie w takie badania lub tworzenie zachęt do inwestowania w tej dziedzinie w sektorze publicznym i prywatnym, uznając, że badania naukowe w znacznym stopniu przyczyniają się do dalszego rozwoju i doskonalenia technologii AI i mając na

uwadze promowanie prawa międzynarodowego oraz wartości i zasad określonych w niniejszym Zaleceniu. Państwa członkowskie powinny również prezentować opinii publicznej przykłady najlepszych praktyk w działaniach naukowców i przedsiębiorstw rozwijających technologie AI w sposób zgodny z zasadami etycznymi oraz wspierać współpracę z nimi.

- 108.** Państwa członkowskie powinny zapewnić, by naukowcy zajmujący się sztuczną inteligencją byli przeszkoleni w zakresie etyki badań naukowych oraz wymagać od nich uwzględniania aspektów etycznych w koncepcjach, produktach i publikacjach, w szczególności w analizie wykorzystywanych przez nich zbiorów danych, w sposobie ich komentowania oraz w jakości i zakresie wyników, a także w możliwych aplikacjach.
- 109.** Państwa członkowskie powinny zachęcać podmioty gospodarcze należące do sektora prywatnego do ułatwiania społeczności naukowej dostępu do ich danych na potrzeby badań, zwłaszcza w krajach o niskich i średnich dochodach, w szczególności w krajach najslabiej rozwiniętych, rozwijających się krajach śródziemnomorskich i małych rozwijających się państwach wyspiarskich. Warunki dostępu powinny spełniać właściwe standardy w zakresie poszanowania prywatności i ochrony danych.
- 110.** W celu dokonywania krytycznej oceny badań nad AI i właściwego monitorowania potencjalnych nadużyć lub skutków niepożądanych, państwa członkowskie powinny zapewnić, by wszelkie przyszłe osiągnięcia dotyczące technologii AI opierały się na rzetelnych i niezależnych badaniach naukowych oraz promować interdyscyplinarne badania nad AI poprzez włączenie do nich dyscyplin innych niż nauki ścisłe, technika, inżynieria i nauki matematyczne (STEM), takich jak kulturoznawstwo, edukacja, etyka, stosunki międzynarodowe, prawo, lingwistyka, filozofia, nauki polityczne, socjologia i psychologia.
- 111.** Uznając, że technologie AI stwarzają ogromne możliwości dla postępu wiedzy i praktyki naukowej, zwłaszcza w dyscyplinach tradycyjnie posługujących się modelami, państwa członkowskie powinny zachęcać społeczności naukowe do brania pod uwagę korzyści, ograniczeń i zagrożeń wynikających z ich stosowania, w tym do podejmowania starań, by wnioski wyciągane na podstawie podejścia opartego na wykorzystywaniu danych, modeli i przetwarzaniu informacji były rzetelne i wiarygodne. Ponadto państwa członkowskie powinny ułatwiać udział środowiska naukowego w kreowaniu polityk i w prowadzeniu działalności informacyjnej, mającej na celu ukazywanie atutów i słabych punktów technologii AI.

## Obszar polityki 9: Komunikacja i informacja

- 112.** Państwa członkowskie powinny wykorzystywać systemy AI do zwiększania dostępu do informacji i wiedzy. Może to obejmować wsparcie dla naukowców, środowiska akademickiego, dziennikarzy, ogółu społeczeństwa i twórców oprogramowania, w celu poprawy sytuacji w zakresie wolności słowa, swobód akademickich i naukowych oraz dostępu do informacji, a także szybszego ogłaszania oficjalnych danych i informacji.
- 113.** Państwa członkowskie powinny zapewnić, by podmioty działające w obszarze AI brały pod uwagę poszanowanie i promowanie wolności wypowiedzi, a także ułatwianie dostępu do informacji w odniesieniu do automatycznego tworzenia, moderowania i ochrony treści. Odpowiednie ramy, w tym ramy regulacyjne, powinny sprzyjać przejrzystości w komunikacji online i w działaniach operatorów informacyjnych, zapewniać użytkownikom możliwość zapoznawania się z różnymi punktami widzenia oraz uwzględniać systemy szybkiego powiadamiania ich o przyczynach usunięcia informacji czy innego działania dotyczącego przetwarzania treści, a także mechanizmy odwoławcze umożliwiające dochodzenie zażądań.
- 114.** Państwa członkowskie powinny inwestować w rozwijanie umiejętności cyfrowych, medialnych i informacyjnych oraz promować te umiejętności w dążeniu do ugruntowywania krytycznego myślenia i upowszechniania kompetencji potrzebnych do rozumienia funkcjonowania i skutków działania systemów opartych na sztucznej inteligencji w celu przeciwdziałania dezinformacji, rozpowszechnianiu nieprawdziwych informacji i mowie nienawiści. Do spodziewanych efektów działań w tych dziedzinach należy lepsze rozumienie i poprawa zdolności oceniania zarówno pozytywnych, jak i potencjalnie szkodliwych skutków stosowania systemów rekomendacyjnych.
- 115.** Państwa członkowskie powinny tworzyć warunki korzystne dla mediów, tak by miały one uprawnienia i możliwości pozwalające na rzetelne informowanie o korzyściach i niedogodnościach związanych z systemami AI, a także powinny zachęcać media do wykorzystywania systemów AI w swej działalności w sposób zgodny z zasadami etycznymi.

## Obszar polityki 10: Gospodarka i praca

- 116.** Państwa członkowskie powinny oszacowywać wpływ systemów sztucznej inteligencji na rynki pracy i wynikające stąd potrzeby w zakresie edukacji, poszukując stosownych rozwiązań. Dotyczy to wszystkich krajów, ale zwłaszcza tych, których gospodarkę cechuje pracochłonność. Rozwiązania te mogą obejmować wprowadzenie szerszego zakresu „zasadniczych” i interdyscyplinarnych umiejętności na wszystkich poziomach kształcenia w celu zapewnienia obecnym pracownikom i ich kolejnym pokoleniom równych szans na zatrudnienie na szybko zmieniającym się rynku pracy oraz kształtowanie ich świadomości w sprawie etycznych aspektów systemów AI. Umiejętności dotyczące «uczenia się, jak się uczyć», zdolności komunikowania się, krytycznego myślenia, funkcjonowania w zespole, empatii i zdolności szybkiego przekwalifikowania się, powinny wchodzić w zakres nauczania zarówno w nauce specjalistycznych, technicznych umiejętności, jak i w przysposobieniu do wykonywania zadań wymagających niskich kwalifikacji. Kluczowe znaczenie ma transparentność w kwestii umiejętności poszukiwanych na rynku pracy i uaktualnianie programów nauczania pod tym kątem.
- 117.** Państwa członkowskie powinny wspierać zawieranie porozumień o współpracy między rządami, instytucjami akademickimi, instytucjami kształcenia i szkolenia zawodowego, firmami przemysłowymi, organizacjami pracowniczymi i organizacjami społeczeństwa obywatelskiego w celu wypełniania luki w zakresie potrzebnych kwalifikacji oraz dostosowywania programów i strategii szkoleniowych do uwarunkowań przyszłej pracy i potrzeb przemysłu, w tym małych i średnich przedsiębiorstw. Należy promować projektowe podejście w zakresie nauczania i uczenia się dotyczącego AI, które pozwala na nawiązywanie partnerstw między instytucjami publicznymi, podmiotami gospodarczymi należącymi do sektora prywatnego, uczelniami wyższymi i ośrodkami badawczymi.
- 118.** Państwa członkowskie powinny współpracować z podmiotami gospodarczymi należącymi do sektora prywatnego, organizacjami społeczeństwa obywatelskiego i innymi interesariuszami, w tym pracownikami i związkami zawodowymi, w celu zapewnienia warunków godnego przetrwania okresu przejściowego tym pracownikom, których miejsca pracy są zagrożone. Obejmuje to wprowadzenie programów służących podnoszeniu kwalifikacji i nabywaniu nowych kwalifikacji, znalezienie skutecznych mechanizmów zatrzymywania pracowników w okresach przejściowych oraz ewentualne wdrażanie programu socjalnego dla tych, którzy nie mogą się przekwalifikować. Państwa członkowskie powinny opracowywać

i wdrażać programy mające na celu zbadanie i podjęcie określonych wyzwań, które mogłyby obejmować m.in. podnoszenie kwalifikacji i zmianę kwalifikacji, wzmocnioną osłonę socjalną, proaktywne polityki i działania branżowe, ulgi podatkowe, nowe formy opodatkowania. Państwa członkowskie powinny zapewnić dostatecznie wysokie dofinansowanie publiczne na wspieranie takich programów. Odpowiednie regulacje, takie jak systemy podatkowe, powinny być dokładnie przeanalizowane i w razie potrzeby zmienione, aby przeciwdziałały skutkom bezrobocia spowodowanego automatyzacją, opierającą się na zastosowaniu sztucznej inteligencji.

- 119.** Państwa członkowskie powinny zachęcać naukowców do podejmowania badań nad wpływem systemów AI na lokalne środowisko pracy w celu prognozowania przyszłych tendencji i wyzwań. Badania te powinny mieć charakter interdyscyplinarny i dotyczyć wpływu systemów AI na gospodarkę, sektor społeczny i poszczególne obszary geograficzne, a także na interakcje człowiek-robot i na relacje międzyludzkie, w celu stworzenia systemu doradztwa upowszechniającego najlepsze praktyki w zakresie nabywania dodatkowych kwalifikacji i kierowania pracowników na inne stanowiska pracy.
- 120.** Państwa członkowskie powinny podjąć stosowne kroki w celu zapewnienia konkurencyjności rynków i ochrony konsumentów, rozważając możliwość podjęcia środków i wprowadzenia mechanizmów na szczeblu krajowym, regionalnym i międzynarodowym w celu zapobiegania nadużywaniu dominującej pozycji rynkowej, w tym monopolizacji, w odniesieniu do systemów AI w całym cyklu ich życia, niezależnie od tego, czy dotyczyłaby ona danych, badań, technologii czy rynku. Państwa członkowskie powinny zapobiegać wynikającym stąd nierównościom i oceniać poszczególne rynki, promując te spośród nich, które okażą się konkurencyjne. Należy zwrócić należytą uwagę na kraje o niskich i średnich dochodach, w szczególności kraje najsłabiej rozwinięte, rozwijające się kraje śródlądowe i małe rozwijające się państwa wyspiarskie, które są bardziej narażone i podatne na sytuacje nadużywania dominacji rynkowej z powodu między innymi braku infrastruktury, potencjału ludzkiego i odpowiednich regulacji. Podmioty z obszaru AI projektujące systemy AI w państwach, które ustanowiły lub przyjęły standardy etyczne dotyczące AI, powinny przestrzegać tych standardów przy wywozie takich produktów oraz przy projektowaniu lub wdrażaniu swoich systemów AI w krajach, w których takie standardy mogą nie być ustanowione, jednocześnie przestrzegając obowiązującego prawa międzynarodowego oraz krajowego ustawodawstwa, standardów i praktyk przyjętych w tych krajach.

## Obszar polityki 11: Zdrowie i dobrostan społeczny

- 121.** Państwa członkowskie powinny dążyć do stosowania skutecznych systemów sztucznej inteligencji w celu poprawy stanu zdrowia ludzi i ochrony prawa do życia, w tym do ograniczania epidemii, przy jednoczesnym budowaniu i umacnianiu solidarności międzynarodowej w celu przeciwdziałania globalnym zagrożeniom zdrowotnym i sytuacjom niepewności co do ryzyka kryzysów sanitarnych, a także powinny zapewniać, by wdrażanie przez nie systemów AI w opiece zdrowotnej było zgodne z prawem międzynarodowym i ich zobowiązaniami wynikającymi z postanowień dotyczących praw człowieka. Państwa członkowskie powinny zapewnić, aby podmioty uczestniczące w procedurach związanych z wykorzystaniem systemów opartych na sztucznej inteligencji w opiece zdrowotnej brały pod uwagę znaczenie relacji pacjenta z rodziną i z personelem medycznym.
- 122.** Państwa członkowskie powinny zapewnić, by rozwój i wdrażanie systemów AI związanych ze zdrowiem w ogóle, a w szczególności ze zdrowiem psychicznym, przy należyтым uwzględnieniu sytuacji dzieci i młodzieży, były regulowane w taki sposób, by te systemy były bezpieczne, skuteczne, efektywne, sprawdzone naukowo i medycznie oraz by umożliwiały innowacyjność i postęp w medycynie w oparciu o dowody naukowe. Ponadto w pokrewnym obszarze zdrowia cyfrowego usilnie zachęca się państwa członkowskie do aktywnego angażowania pacjentów i ich przedstawicieli na wszystkich ważnych etapach rozwoju systemu.
- 123.** Państwa członkowskie powinny zwrócić szczególną uwagę na regulacje w dziedzinie opieki zdrowotnej, dotyczące rozwiązań w zakresie prognozowania, diagnostyki i leczenia z zastosowaniem AI poprzez:
- a.** zapewnienie nadzoru w celu zminimalizowania i ograniczenia błędów
  - b.** zapewnienie, by specjalista, pacjent, opiekun lub użytkownik usług był włączony jako „ekspert w danej dziedzinie” do zespołu na wszystkich istotnych etapach tworzenia algorytmów;
  - c.** zwrócenie należytej uwagi na prywatność ze względu na potencjalną potrzebę monitorowania pacjenta pod względem medycznym oraz zapewnienie, by spełnione zostały wszystkie odpowiednie krajowe i międzynarodowe wymogi w zakresie ochrony danych;
  - d.** stworzenie skutecznych, ale nie utrudniających dostępu do opieki zdrowotnej mechanizmów, pozwalających na to, by osoby, których dane osobowe są ana-

lizowane, wiedziały o tym i wyrażały świadomą zgodę na wykorzystanie i analizę swoich danych;

- e. zapewnienie opieki medycznej przez człowieka i podejmowanie ostatecznych decyzji w sprawie diagnozy i leczenia zawsze przez ludzi, ale przy założeniu, że systemy AI mogą im pomagać w pracy;
  - f. w razie potrzeby, zapewnienie weryfikacji systemów AI przed zastosowaniem klinicznym przez komisję etyczną.
- 124.** Państwa członkowskie powinny prowadzić badania nad skutkami stosowania systemów AI dla zdrowia psychicznego i wprowadzić regulacje dotyczące ich potencjalnych negatywnych skutków, takich jak, między innymi, nasilenie się stanów depresyjnych i lękowych, izolacja społeczna, rozwój uzależnień, handel ludźmi, radykalizacja postaw i rozpowszechnianie błędnych informacji.
- 125.** Państwa członkowskie powinny opracować wytyczne dotyczące interakcji człowiek-robot i ich wpływu na relacje międzyludzkie, oparte na badaniach i uwzględniające dalszy rozwój robotyki, ze szczególnym uwzględnieniem aspektów dotyczących zdrowia psychicznego i fizycznego ludzi. Szczególną uwagę należy poświęcić wykorzystaniu robotów w opiece nad chorymi oraz w opiece nad osobami starszymi i niepełnosprawnymi, w edukacji, a także robotów przeznaczonych dla dzieci, robotów – zabawek, chatbotów i robotów towarzyszących dla dzieci i osób dorosłych. Ponadto technologie AI powinny być wykorzystywane w celu zwiększenia bezpieczeństwa i ergonomicznego przystosowania robotów, w tym w środowisku pracy człowiek-robot. Szczególną uwagę należy zwrócić na ryzyko posługiwania się sztuczną inteligencją w manipulowaniu ludźmi i wykorzystywaniu ludzkiej niewiedzy w nieuczciwych celach.
- 126.** Państwa członkowskie powinny zapewnić, by interakcja człowiek-robot była zgodna z wartościami i zasadami odnoszącymi się do wszelkich innych systemów AI, w tym z prawami człowieka i podstawowymi wolnościami, promowaniem różnorodności oraz ochroną osób szczególnie narażonych lub znajdujących się w trudnej sytuacji. Kwestie etyczne związane z systemami wykorzystywanymi w neurotechnologii oraz interfejsy mózg-komputer powinny być przedmiotem szczególnej uwagi w związku z potrzebą ochrony godności i autonomii człowieka.
- 127.** Państwa członkowskie powinny zapewnić, aby użytkownicy mogli łatwo rozpoznać, czy mają do czynienia z istotą żywą, czy z systemem AI naśladującym cechy ludzkie lub zwierzęce, oraz aby mieli oni możliwość skutecznej odmowy wchodzenia w taką interakcję i zażądania kontaktu z człowiekiem.



- 128.** Państwa członkowskie powinny wdrożyć politykę mającą na celu podniesienie stanu świadomości na temat antropomorfizacji technologii AI oraz technologii, które rozpoznają i naśladują ludzkie emocje, w tym w zakresie języka używanego do ich opisywania, a także powinny poddawać ocenie przejawy, implikacje etyczne i ewentualne ograniczenia takiej antropomorfizacji, w szczególności w kontekście interakcji robot-człowiek, zwłaszcza gdy dotyczy ona dzieci.
- 129.** Państwa członkowskie powinny zachęcać do podejmowania opartych na współpracy badań nad skutkami długotrwałej interakcji ludzi i systemów AI, zwracając szczególną uwagę na wpływ psychologiczny i poznawczy, jaki systemy te mogą wywierać na dzieci i młodzież, a także promować takie badania. W badaniach należy brać pod uwagę różnorakie normy, zasady, protokoły, podejście według dyscyplin oraz ocenę zmiany zachowania i nawyków, jak również wnikliwą ocenę dalekosiężnych skutków kulturowych i społecznych. Ponadto państwa członkowskie powinny zachęcać do podejmowania badań nad wpływem technologii opartych na sztucznej inteligencji na funkcjonowanie systemu opieki zdrowotnej i wyniki w leczeniu.
- 130.** Państwa członkowskie, jak również wszystkie zainteresowane podmioty, powinny wprowadzić mechanizmy umożliwiające szerokie włączenie dzieci i młodzieży do rozmów, debat i procesów decyzyjnych dotyczących wpływu systemów AI na ich życie i przyszłość.

## V. Monitoring i ewaluacja

- 131.** Państwa członkowskie powinny, zgodnie z własnymi uwarunkowaniami, systemem rządów i przepisami konstytucyjnymi, w sposób wiarygodny i transparentny monitorować i oceniać polityki, programy i mechanizmy dotyczące etyki AI, stosując połączenie podejścia ilościowego i jakościowego. UNESCO może je w tym wspierać poprzez:
- a.** opracowanie metodologii UNESCO dotyczącej oceny oddziaływania pod względem etycznym, opartej na rzetelnych badaniach naukowych i osadzonej w międzynarodowym systemie ochrony praw człowieka (EIA), wytycznych w sprawie wdrażania tej metodologii na wszystkich etapach cyklu życia systemów AI, a także materiałów szkoleniowych w celu wspierania działań państw członkowskich w zakresie szkolenia urzędników państwowych, decydentów i innych właściwych podmiotów z obszaru AI w zakresie EIA;
  - b.** opracowanie metodologii oceny stanu gotowości, w celu udzielania pomocy państwom członkowskim w określeniu stopnia ich przygotowania do wdrożenia Zalecenia z różnej perspektywy, na poszczególnych etapach;
  - c.** opracowanie metodologii UNESCO służącej dokonywaniu oceny ex ante i oceny ex post skuteczności i efektywności polityk dotyczących etyki i zachęt w zakresie AI w odniesieniu do określonych celów;
  - d.** przeprowadzenie pogłębionej analizy opartej na badaniach i dowodach oraz ustanowienie systemu sprawozdawczości w zakresie polityk dotyczących etyki AI;
  - e.** gromadzenie i upowszechnianie informacji o postępach i innowacjach, raportów badawczych, publikacji naukowych, danych i statystyk dotyczących polityk w zakresie etyki AI, w tym w ramach istniejących inicjatyw, w celu wspierania wymiany najlepszych praktyk i wzajemnego dzielenia się wiedzą dla przyspieszenia procesu wdrażania niniejszego Zalecenia.
- 132.** W procesach monitorowania i oceny powinien być zapewniony szeroki udział wszystkich zainteresowanych stron, w tym, ale nie wyłącznie, osób podatnych na zagrożenia lub znajdujących się w trudnej sytuacji. Należy przy tym zapewnić różnorodność pod względem społecznym i kulturowym oraz pod względem płci

w celu usprawnienia procesów uczenia się, a także umacniania powiązań między ustaleniami z badań, procesem decyzyjnym, transparentnością i rozliczaniem się z osiągniętych wyników.

- 133.** W celu promowania najlepszych polityk i praktyk związanych z etyką AI należy opracować odpowiednie narzędzia i wskaźniki służące do przeprowadzania oceny ich skuteczności i efektywności w odniesieniu do uzgodnionych standardów, priorytetów i celów, w tym konkretnych celów dotyczących osób należących do grup znajdujących się w niekorzystnej sytuacji, populacji wykluczonych społecznie oraz osób szczególnie narażonych lub znajdujących się w trudnej sytuacji, jak również wpływu systemów AI wywieranego na poziomie jednostki i społeczeństwa. Monitorowanie i ocena wpływu systemów AI oraz związanych z nimi polityk i praktyk w zakresie etyki AI powinny być prowadzone ustawicznie i w sposób systematyczny, proporcjonalnie do występujących zagrożeń. Powinno się to opierać na ramach uzgodnionych na szczeblu międzynarodowym i obejmować oceny instytucji prywatnych i publicznych, usługodawców i programów, w tym samooceny, a także badania ankietowe i opracowywanie zestawów wskaźników. Gromadzenie i przetwarzanie danych powinno odbywać się zgodnie z prawem międzynarodowym i ustawodawstwem krajowym w zakresie ochrony danych i prywatności danych oraz zgodnie z wartościami i zasadami przedstawionymi w niniejszym Zaleceniu.
- 134.** Państwa członkowskie mogą zechcieć w szczególności rozważyć ustanowienie mechanizmów monitorowania i oceny, takich jak komisja do spraw etyki, obserwatorium etyki AI, repozytorium poświęcone rozwojowi systemów AI zgodnemu z prawami człowieka i etyką lub też rozwijanie istniejących inicjatyw poprzez zajęcie się kwestią przestrzegania zasad etycznych w obszarach kompetencji UNESCO, ustanowienie mechanizmu wymiany doświadczeń, piaskownice regulacyjne AI oraz opracowanie przewodnika dla wszystkich podmiotów działających w obszarze AI na temat przeprowadzania oceny przestrzegania przez nie rekomendacji dotyczących poszczególnych polityk wymienionych w niniejszym dokumencie.

## VI. Wykorzystanie niniejszego Zalecenia

- 135.** Państwa członkowskie i wszystkie inne zainteresowane strony określone w niniejszym Zaleceniu powinny respektować, promować i chronić określone w nim wartości, zasady i normy etyczne dotyczące AI oraz podjąć wszelkie możliwe działania w celu wprowadzenia w życie zawartych w nim rekomendacji dotyczących poszczególnych polityk.
- 136.** Państwa członkowskie powinny dążyć do rozszerzenia i uzupełnienia swoich własnych działań dotyczących niniejszego Zalecenia, poprzez współpracę ze wszystkimi właściwymi krajowymi i międzynarodowymi organizacjami rządowymi i pozarządowymi, jak również z korporacjami transnarodowymi i organizacjami naukowymi, których działalność odpowiada zakresowi i celom niniejszego Zalecenia. Do ważnych narzędzi w tym zakresie należeć może opracowanie przez UNESCO metodologii oceny oddziaływania AI pod względem etycznym (EIA) i tworzenie krajowych komisji do spraw etyki AI.

## VII. Promowanie niniejszego Zalecenia

- 137.** UNESCO jest główną organizacją wyspecjalizowaną Narodów Zjednoczonych powołaną do promowania i upowszechniania niniejszego Zalecenia i w związku z tym będzie współpracować z innymi właściwymi jednostkami wchodzącymi w skład systemu Narodów Zjednoczonych, respektując ich mandat i unikając dublowania działań.
- 138.** UNESCO i niektóre z jej organów, takie jak Światowa Komisja Etyki Wiedzy Naukowej i Technologii (COMEST), Międzynarodowy Komitet Bioetyczny (IBC) i Międzypaństwowy Komitet Bioetyczny (IGBC), będą również współpracować z innymi międzynarodowymi, regionalnymi i subregionalnymi organizacjami rządowymi i pozarządowymi.
- 139.** Mimo że w ramach UNESCO promowanie i ochrona leżą w gestii rządów i organów międzyrządowych, społeczeństwo obywatelskie jest ważnym podmiotem, który stoi na straży interesów sektora publicznego. UNESCO musi w związku z tym uwzględniać i legitymizować jego rolę.

## VIII. Postanowienia końcowe

- 140.** Niniejsze Zalecenie należy traktować jako całość, a podstawowe wartości i zasady należy rozumieć jako komplementarne i wzajemnie ze sobą powiązane.
- 141.** Żadne z postanowień niniejszego Zalecenia nie może być interpretowane jako zastępujące, zmieniające lub w jakikolwiek sposób naruszające zobowiązania lub prawa państw wynikające z prawa międzynarodowego, ani jako przyzwolenie dla jakiegokolwiek państwa, innego podmiotu politycznego, gospodarczego lub społecznego, grupy lub osoby na zaangażowanie się w jakąkolwiek działalność lub popełnienie jakiegokolwiek czynu, które byłyby sprzeczne z prawami człowieka, podstawowymi wolnościami, godnością ludzką czy troską o środowisko i ekosystemy, zarówno żywe, jak i nieożywione.