



# AI-ERIT: AI Integration Framework for Higher Education: Ensuring HEIs' Readiness for Inescapable Transformation

## Wytyczne dotyczące audytu w zakresie sztucznej inteligencji



Opracowanie: BATMAN UNIVERSITESI Turcja



Fachhochschule  
des Mittelstands



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-

## Spis treści

<b>KRÓTKIE WPROWADZENIE DO PROJEKTU AI-ERIT</b> .....	<b>1</b>
<b>STRUKTURA DOKUMENTU</b> .....	<b>2</b>
<b>TERMINY I DEFINICJE</b> .....	<b>4</b>
<b>1. WSTĘP</b> .....	<b>5</b>
1.1. CEL I ZADANIA .....	5
1.2. ZNACZENIE AUDYTÓW AI W SZKOŁACH WYŻSZYCH .....	6
1.3. PRZEGLĄD KLUCZOWYCH KORZYŚCI .....	8
<b>2. KONCEPCJE PODSTAWOWE</b> .....	<b>8</b>
2.1. DEFINICJA SZTUCZNEJ INTELIGENCJI W SZKOLNICTWIE WYŻSZYM .....	8
2.2. WYZWANIA W ZARZĄDZANIU SZTUCZNĄ INTELIGENCJĄ I ROZWAŻANIA DOTYCZĄCE AUDYTU W INSTYTUCJACH SZKOLNICTWA WYŻSZEGO .....	9
2.3. ZASADY ETYCZNE WDRAŻANIA SZTUCZNEJ INTELIGENCJI W SZKOLNICTWIE WYŻSZYM .....	9
2.3.1. <i>Odpowiedzialność i odpowiedzialność</i> .....	9
2.3.2. <i>Stronniczość i uczciwość</i> .....	9
2.3.3. <i>Autonomia i sprawczość człowieka</i> .....	10
2.3.4. <i>Prywatność i ochrona danych</i> .....	10
2.3.5. <i>Bezpieczeństwo i ochrona</i> .....	10
2.3.6. <i>Inkluzywność</i> .....	10
2.3.7. <i>Przejrzystość i wyjaśnialność</i> .....	10
2.4. RAMY PRAWNE WDRAŻANIA SZTUCZNEJ INTELIGENCJI W SZKOLNICTWIE WYŻSZYM .....	11
2.5. ZASADY ODPOWIEDZIALNEGO KORZYSTANIA ZE SZTUCZNEJ INTELIGENCJI .....	14
2.6. PRZEKSZTAŁCANIE ZASAD ETYCZNYCH I RAM PRAWNYCH W PROCESY AUDYTOWE .....	15
<b>3. RAMY PROCESU AUDYTU AI</b> .....	<b>17</b>
<b>4. SZCZEGÓŁOWE INSTRUKCJE KROK PO KROKU</b> .....	<b>19</b>
<b>5. KROKI AUDYTU</b> .....	<b>25</b>
5.1. PRZYGOTOWANIE DO AUDYTU .....	25
5.1.1. <i>Określanie celów audytu</i> .....	25
5.1.2. <i>Zbierz zespoły audytorskie</i> .....	26
5.2. AUDYT ETYCZNEGO WYKORZYSTANIA SZTUCZNEJ INTELIGENCJI .....	27
5.2.1 <i>Audyt procesu badawczego wspomaganego sztuczną inteligencją</i> .....	28
5.2.2. <i>Audyt procesu edukacji wspomaganego sztuczną inteligencją</i> .....	36
5.2.3 <i>Audyt procesów administracyjnych wspomaganego przez sztuczną inteligencję</i> .....	44
5.2.4. <i>Uzasadnienie wartości odniesienia w wytycznych</i> .....	52
<b>6. OCENA OGÓLNEGO PROCESU AUDYTU</b> .....	<b>54</b>
<b>7. OGÓLNY RAPORT Z AUDYTU</b> .....	<b>56</b>
<b>ZAŁĄCZNIK</b> .....	<b>58</b>
AUDYT TECHNICZNY SYSTEMÓW/NARZĘDZI AI .....	58
OCENA TECHNICZNA SYSTEMÓW/NARZĘDZI AI .....	58
<b>REFERENCJE</b> .....	<b>66</b>



## Krótkie wprowadzenie do projektu AI-ERIT

Projekt AI-ERIT (2024-1-LT01-KA220-HED-000251565) ma na celu ułatwienie integracji sztucznej inteligencji (AI) w instytucjach szkolnictwa wyższego (HEI), ze szczególnym uwzględnieniem etyki, budowania potencjału i adaptacji technologicznej. Inicjatywa ta, finansowana ze środków programu Erasmus+ (KA2 Partnerstwa Współpracy / Szkolnictwo Wyższe), jest prowadzona przez Wyższą Szkołę Nauk Stosowanych SMK (Litwa) we współpracy z Lubelską Akademią WSEI (Polska), Veleučilište Aspira (Chorwacja), FHM – Fachhochschule des Mittelstands (Niemcy), Maiêutica / Universidade da Maia (Portugalia), SIA Biznesa Augstskola Turiba (Łotwa) i Uniwersytetem Batman (Turcja).

W ramach projektu partnerzy opracują kompleksowe ramy odniesienia, które będą zawierać zalecenia dotyczące etycznego i odpowiedzialnego wykorzystania narzędzi AI w nauczaniu, badaniach i usługach administracyjnych, w tym w procesach audytu. Głównymi beneficjentami będą pracownicy naukowcy, badacze, pracownicy administracyjni i studenci, którzy otrzymają wsparcie w postaci samouczków, szkoleń online oraz dedykowanego modułu poświęconego znajomości i etyce AI.

### Cele

- Przygotowanie instytucji szkolnictwa wyższego na transformacyjny wpływ sztucznej inteligencji na działalność instytucjonalną.
- Ustanowienie wytycznych etycznych i najlepsze praktyki dotyczące wykorzystania sztucznej inteligencji w nauczaniu, badaniach i funkcjach administracyjnych .
- Poprawa budowania potencjału poprzez wyposażenie kadry naukowej, badaczy i personelu administracyjnego w niezbędne umiejętności i wiedzę, aby mogli odpowiedzialnie wykorzystywać sztuczną inteligencję.

### Kluczowe rezultaty

- Rekomendacje dotyczące wdrażania narzędzi AI w szkołach wyższych.
- Zestaw narzędzi do wdrażania sztucznej inteligencji w usługach administracyjnych.
- Wytyczne dotyczące wykorzystania sztucznej inteligencji w procesach audytu.
- Samouczki i szkolenia online przeznaczone dla wykładowców, badaczy i personelu administracyjnego.
- Moduł edukacyjny poświęcony „Umiejętności korzystania ze sztucznej inteligencji i etyce”.

**Szczegóły projektu:** Zasięg geograficzny: Międzynarodowy

Czas trwania projektu: 01.09.2024 – 31.10.2026

Instytucja finansująca: Erasmus+ (Partnerstwa w ramach współpracy KA2 / Szkolnictwo wyższe);



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them. Grant no:2024-1-LT01-



# Struktura dokumentu

## 1. Wprowadzenie

Wprowadzenie przedstawia kontekst audytu AI w szkolnictwie wyższym, podkreślając zarówno potencjał transformacyjny, jak i nieodłączne ryzyko związane z zastosowaniami AI. Wyjaśnia również, dlaczego systematyczny proces audytu jest niezbędny dla zapewnienia uczciwości, przejrzystości oraz zgodności z normami prawnymi i etycznymi.

- Rola sztucznej inteligencji w transformacji nauczania, badań i administracji.
- Zagrożenia: stronniczość, naruszenia prywatności danych, nieprzejrzysty proces decyzyjny, nierówny dostęp.
- Uzasadnienie ustrukturyzowanego procesu audytu sztucznej inteligencji.
- Zalecenie: Należy opracować **oświadczenie dotyczące polityki audytu sztucznej inteligencji**, określające zakres i zasady zarządzania.

## 2. Koncepcje podstawowe

W tej sekcji zdefiniowano podstawowe koncepcje, zasady etyczne i ramy prawne niezbędne do zrozumienia i przeprowadzania audytów AI w instytucjach szkolnictwa wyższego. Łączy ona teorię z praktyką, pomagając instytucjom w dostosowywaniu się do najlepszych światowych praktyk.

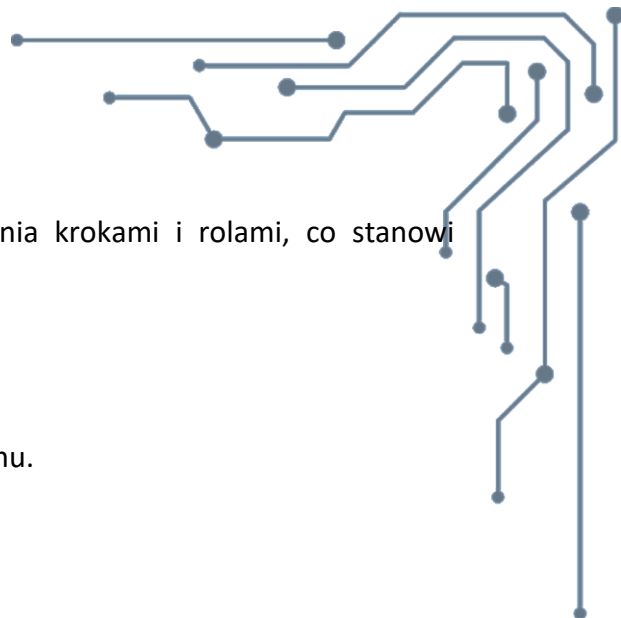
- **Definicja sztucznej inteligencji w szkolnictwie wyższym** : edukacja, badania naukowe i administracja.
- **Wyzwania dla zarządzania**: szybkie zmiany technologiczne kontra powolne dostosowywanie polityki.
- **Zasady etyczne**: odpowiedzialność, uczciwość, autonomia, prywatność, bezpieczeństwo, inkluzywność, przejrzystość.
- **Podstawy prawne**: RODO, FERPA, ustawa o sztucznej inteligencji, standardy ESG.
- **Struktura CRAFT**: zasady, dostęp, znajomość, zaufanie, kultura.
- Przełożenie etyki na mierzalne wskaźniki audytu.

## 3. Ramy procesu audytu AI

Niniejsze ramy przedstawiają siedmioetapową metodologię audytu, zapewniającą kompleksową ocenę od planowania po ciągły monitoring. Ma ona charakter cykliczny, wspierając ciągłe doskonalenie.

- **Faza 1**: Planowanie i określanie zakresu.
- **Faza 2**: Zaangażowanie interesariuszy.
- **Faza 3**: Ocena danych i modelu.
- **Faza 4**: Ocena ryzyka.
- **Faza 5**: Przegląd zgodności.
- **Faza 6**: Raportowanie i zalecenia.
- **Faza 7**: Wdrożenie i ciągły monitoring.





#### 4. Szczegółowe instrukcje krok po kroku

Każda faza jest opisana wraz z możliwymi do wykonania krokami i rolami, co stanowi przejrzysty plan operacyjny dla zespołów audytorskich.

- Określ zakres, cele i kryteria sukcesu.
- Zbierz interdyscyplinarny zespół audytorski.
- Przeprowadź warsztaty i ankiety wśród interesariuszy.
- Oceń jakość danych, stroniczość i wydajność algorytmu.
- Mapowanie zgodności z ramami prawnymi
- Ustal priorytety ryzyka według jego powagi i pilności.
- Przygotuj listy kontrolne dla poszczególnych faz.

#### 5. Kroki audytu

Procedury audytu specyficzne dla danej dziedziny zapewniają dokładną ocenę każdego obszaru zastosowania sztucznej inteligencji. Do użytku praktycznego dostępne są listy kontrolne i tabele ryzyka.

- **Procesy badawcze:** ocena stroniczości, powtarzalności i zgodności z wymogami etycznymi.
- **Procesy edukacyjne:** ocena uczciwości, dostępności i efektów uczenia się.
- **Procesy administracyjne:** przegląd przejrzystości, efektywności i równości usług.
- **Systemy/narzędzia AI:**
  - **Ocena techniczna :** wydajność, solidność, skalowalność.
  - **Ocena etyczna:** wykrywanie stroniczości, przejrzystość, wyjaśnialność.

#### 6. Ocena ogólnego procesu audytu

Na tym etapie wyniki ze wszystkich dziedzin konsolidowane są w instytucjonalny profil ryzyka. Multidyscyplinarne podejście zapewnia zrównoważony proces decyzyjny.

- Zbiorcze wyniki ze wszystkich obszarów audytu.
- Klasyfikuj ryzyko jako wysokie, średnie lub niskie.
- Korzystaj z paneli oceniających opartych na konsensusie.
- Ustal harmonogram działań naprawczych.
- Udokumentuj uzasadnienie decyzji.

#### 7. Ogólne sprawozdanie z audytu

Szablon raportu porządkuje wyniki audytu w dokumencie strategicznym dla kierownictwa instytucji. Identyfikuje on osiągnięcia, ryzyka i działania priorytetowe.

- Podsumuj ustalenia dla każdej dziedziny.
- Podkreśl mocne strony i najlepsze praktyki.
- Wymień obszary wymagające pilnej uwagi.
- Przedstaw rekomendacje, które można wdrożyć.
- Zintegruj z rocznym raportem jakości.



## Terminy i definicje

Poniżej przedstawiono podstawowe terminy i definicje stosowane w niniejszych wytycznych dotyczących audytu.

**Audyt AI:** systematyczna ocena systemów AI w celu zapewnienia zgodności z zasadami etycznymi, przepisami prawnymi i celami instytucjonalnymi, obejmująca zarówno kwestie wydajności technicznej, jak i aspekty zarządzania.

**Systemy AI:** Rozwiązania technologiczne wykorzystujące metody sztucznej inteligencji do wykonywania zadań wymagających inteligencji zbliżonej do ludzkiej. W szkolnictwie wyższym wspierają one nauczanie, badania i administrację za pomocą narzędzi takich jak algorytmy rekrutacyjne, automatyczne ocenianie, analityka uczenia się i modelowanie predykcyjne.

**Odpowiedzialność:** Zasada etyczna wymagająca jasno określonego ludzkiego nadzoru i odpowiedzialności za podejmowanie decyzji w odniesieniu do systemów AI, wraz z udokumentowanymi mechanizmami reagowania na ryzyka związane ze sztuczną inteligencją.

**Stronniczość i uczciwość:** wymóg, aby systemy AI były projektowane, testowane i monitorowane w sposób zapobiegający dyskryminacji i zapewniający równe wyniki różnym grupom użytkowników.

**Autonomia i sprawczość człowieka:** zasada, zgodnie z którą sztuczna inteligencja powinna wspierać proces podejmowania decyzji przez ludzi, a nie zastępować go lub podważać, zapewniając użytkownikom kontrolę nad kluczowymi wynikami.

**Prywatność i ochrona danych:** Ochrona danych osobowych w systemach sztucznej inteligencji, zapewnienie gromadzenia, przetwarzania i przechowywania zgodnie z obowiązującymi przepisami dotyczącymi prywatności, takimi jak np. RODO.

**Bezpieczeństwo:** zapewnienie niezawodnego działania systemów AI przy jednoczesnym minimalizowaniu ryzyka dla bezpieczeństwa fizycznego, psychologicznego i cyfrowego, w tym solidnej ochrony przed złośliwymi atakami.

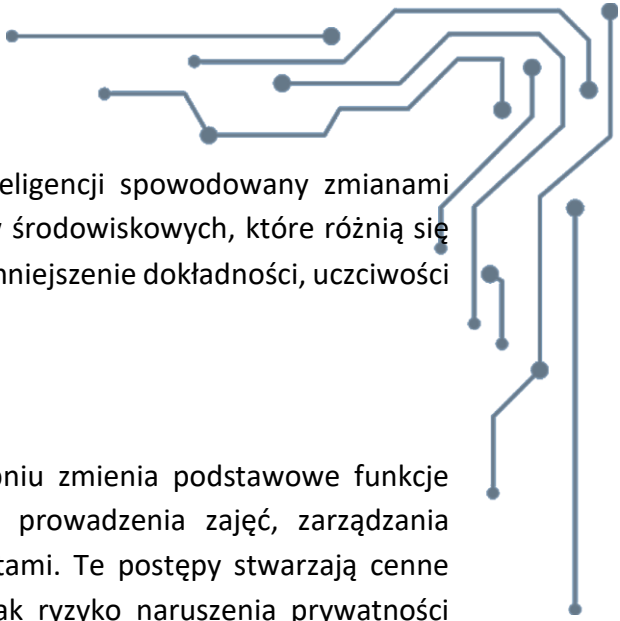
**Włączanie:** projektowanie i wdrażanie systemów AI w taki sposób, aby były dostępne i użyteczne dla różnych grup społecznych, w tym grup zmarginalizowanych i bezbronnych.

**Przejrzystość i możliwość wyjaśnienia:** zasada, zgodnie z którą procesy decyzyjne sztucznej inteligencji powinny być zrozumiałe dla użytkowników i interesariuszy, a także zawierać jasną dokumentację i jawną rolę sztucznej inteligencji w osiągnięciu wyników.



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them.  
Grant no:2024-1-LT01-KA220-HED-000251565



**Dryf modelu:** Spadek wydajności systemu sztucznej inteligencji spowodowany zmianami wzorców danych, zachowań użytkowników lub warunków środowiskowych, które różnią się od tych użytych podczas szkolenia modelu. Powoduje to zmniejszenie dokładności, uczciwości lub trafności wyników w dłuższej perspektywie.

## 1. Wstęp

Sztuczna inteligencja (AI) w coraz większym stopniu zmienia podstawowe funkcje instytucji szkolnictwa wyższego, wpływając na sposób prowadzenia zajęć, zarządzania administracją, prowadzenia badań i interakcji ze studentami. Te postępy stwarzają cenne możliwości, ale także wiążą się z wyzwaniami, takimi jak ryzyko naruszenia prywatności danych, stronniczość algorytmiczna, kwestie etyczne i zgodność z przepisami (Bates i in., 2020). Aby zapewnić odpowiedzialne wdrażanie technologii AI i ich zgodność z wartościami instytucji, uniwersytety muszą przeprowadzać regularne audyty wykraczające poza samą wydajność techniczną.

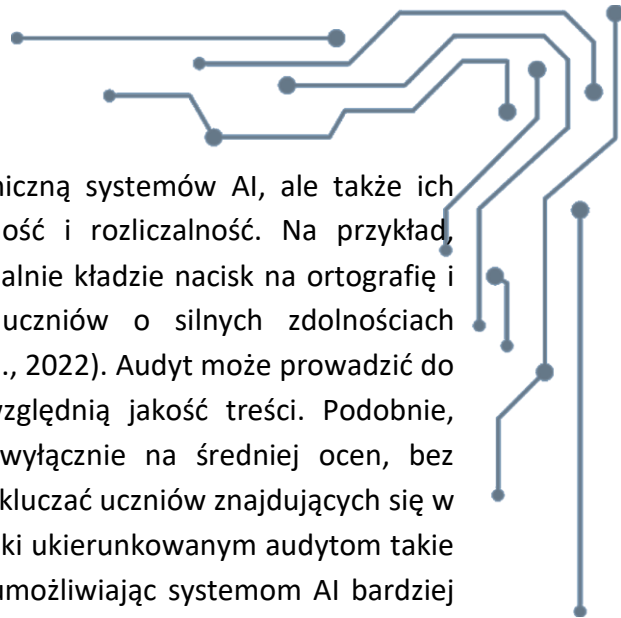
Niniejsze wytyczne stanowią przejrzyste i praktyczne źródło informacji dla liderów szkolnictwa wyższego, kadry akademickiej i personelu administracyjnego, w tym osób nieposiadających doświadczenia w dziedzinie sztucznej inteligencji. Zawierają one ustrukturyzowane, krok po kroku podejście do audytu systemów sztucznej inteligencji, uzupełnione listami kontrolnymi, przykładowymi narzędziami i prostymi wyjaśnieniami. Postępując zgodnie z tymi ramami, instytucje mogą promować etyczne, sprawiedliwe i przejrzyste wykorzystanie sztucznej inteligencji, wspierając w ten sposób bardziej inkluzywne, godne zaufania i efektywne środowisko uczenia się i pracy.

Sztuczna inteligencja jest już obecna w różnorodnych funkcjach akademickich, od oceny przyjęć na studia po personalizację treści dydaktycznych. Jednak wewnętrzne mechanizmy działania tych systemów – a zwłaszcza dane, na których się opierają, i sposób, w jaki kształtują wyniki – są często niejasne, co budzi poważne obawy dotyczące przejrzystości i rozliczalności (Cheong, 2024). Obawy te podkreślają konieczność systematycznego audytu sztucznej inteligencji, aby zapewnić, że instytucjonalne poleganie na sztucznej inteligencji pozostaje uczciwe, etyczne i zgodne z podstawowymi wartościami edukacyjnymi.

### 1.1. Cel i zadania

Technologie sztucznej inteligencji (AI) transformują szkolnictwo wyższe, usprawniając obsługę studentów, optymalizując procesy administracyjne i tworząc bardziej angażujące środowiska edukacyjne. Choć te zmiany oferują znaczne korzyści, niosą ze sobą również wyzwania, takie jak stronniczość, naruszenia prywatności i obawy dotyczące odpowiedzialności (Crompton i Burke, 2023). Niniejszy przewodnik wspiera instytucje szkolnictwa wyższego, prezentując ustrukturyzowane ramy przeprowadzania audytów AI, oferując narzędzia i strategie zapewniające zgodność systemów AI z celami etycznymi, prawnymi i instytucjonalnymi oraz zapewniając mechanizmy promujące zaufanie, równość i przejrzystość.





Proces audytu ocenia nie tylko wydajność techniczną systemów AI, ale także ich zgodność z zasadami takimi jak uczciwość, inkluzywność i rozliczalność. Na przykład, zautomatyzowany system oceniania, który nieproporcjonalnie kładzie nacisk na ortografię i gramatykę, może stawiać w niekorzystnej sytuacji uczniów o silnych zdolnościach poznawczych, ale słabszej mechanice pisania (Johnson i in., 2022). Audyt może prowadzić do bardziej zrównoważonych ram, które odpowiednio uwzględnią jakość treści. Podobnie, algorytm przyznawania stypendiów, który opiera się wyłącznie na średniej ocen, bez uwzględnienia sytuacji społeczno-ekonomicznej, może wykluczać uczniów znajdujących się w rzeczywistej potrzebie finansowej (Turahman, 2024). Dzięki ukierunkowanym audytom takie niedociągnięcia można zidentyfikować i wyeliminować, umożliwiając systemom AI bardziej sprawiedliwe, przejrzyste i odpowiedzialne działanie.

Niniejszy przewodnik został opracowany w celu zapewnienia efektywnego i zgodnego z zasadami etycznymi, prawnymi i instytucjonalnymi funkcjonowania systemów sztucznej inteligencji w szkolnictwie wyższym. Zakres jego działania obejmuje rekrutację, ocenianie, przyznawanie stypendiów, wsparcie badań, kształcenie na odległość oraz automatyzację administracyjną. Proces audytu obejmuje ocenę etyczną, analizę ryzyka, ochronę danych i doświadczenie użytkownika, wykraczając poza samą wydajność techniczną. Wszystkie oceny są zgodne ze Standardami i Wytycznymi Zapewnienia Jakości w Europejskim Obszarze Szkolnictwa Wyższego (ESG, 2015).

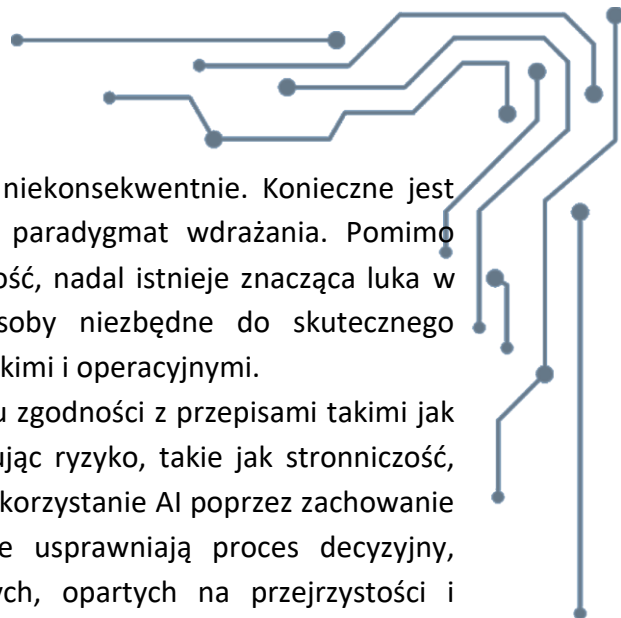
Do głównych grup odpowiedzialnych za wdrożenie tego przewodnika należą zespoły IT i rozwoju oprogramowania, biura ds. zapewnienia jakości, komisje etyczne, inspektorzy ochrony danych oraz właściwi liderzy akademicki i administracyjni. Przewodnik, wyposażając te jednostki w praktyczne narzędzia, listy kontrolne i przykładowe szablony, ma na celu wzmocnienie bezpiecznego, sprawiedliwego i przejrzystego zarządzania systemami sztucznej inteligencji (AI) w instytucjach szkolnictwa wyższego.

## 1.2. Znaczenie audytów AI w szkołach wyższych

Systematyczne integrowanie sztucznej inteligencji (AI) z nauczaniem, badaniami i działalnością okazało się trudne dla wielu instytucji szkolnictwa wyższego. Znacznej liczbie z nich brakuje personelu posiadającego niezbędne doświadczenie, aby skutecznie wdrażać i nadzorować AI (AI w edukacji: Raport specjalny Microsoft, b.d.). Przeszkodami są obawy dotyczące prywatności danych, stronniczość algorytmiczna, wykorzystywanie własności intelektualnej, zagrożenia dla uczciwości akademickiej oraz etyczne wykorzystanie AI zarówno przez studentów, jak i nauczycieli (Organizacja Narodów Zjednoczonych do spraw Oświaty, Nauki i Kultury [UNESCO], 2023a). Wykluczenie cyfrowe pogłębiają regionalne różnice regulacyjne i nierówny dostęp do narzędzi AI, szczególnie w krajach o niskich i średnich dochodach. Naukowcy i pedagodzy obawiają się również, że AI może zastąpić lub ograniczyć część ich obowiązków, zwiększając presję na i tak już duże obciążenie pracą.

Chociaż studenci cenią sobie wsparcie sztucznej inteligencji, nadal przywiązują większą wagę do ludzkich aspektów relacji nauczyciel-uczeń (UNESCO, 2023a). Podobnie jak w przypadku podejścia opartego na eksperymentach, stosowanego w przemyśle, instytucje





wdrażają sztuczną inteligencję generatywną ostrożnie i niekonsekwentnie. Konieczne jest jednak przejście na kompleksowy i dobrze wspierany paradygmat wdrażania. Pomimo rosnącej świadomości wartości takich jak etyka i uczciwość, nadal istnieje znacząca luka w wyposażeniu liderów, nauczycieli i studentów w zasoby niezbędne do skutecznego zintegrowania sztucznej inteligencji z procesami akademickimi i operacyjnymi.

Audyty AI odgrywają kluczową rolę w zapewnianiu zgodności z przepisami takimi jak RODO, polityka instytucjonalna, jednocześnie minimalizując ryzyko, takie jak stronniczość, dyskryminacja i naruszenia danych. Wspierają etyczne wykorzystanie AI poprzez zachowanie uczciwości, odpowiedzialności i inkluzywności, a także usprawniają proces decyzyjny, umożliwiając wyciąganie wniosków opartych na danych, opartych na przejrzystości i dokładności. W tym kontekście audyt AI odnosi się do systematycznej oceny systemów AI w instytucjach szkolnictwa wyższego w celu zapewnienia zgodności z zasadami etycznymi, prawnymi i instytucjonalnymi. Proces ten identyfikuje i eliminuje zagrożenia, takie jak naruszenia prywatności danych, stronniczość algorytmiczna, brak przejrzystości i potencjalne szkody dla użytkowników. Ocenia również, w jaki sposób aplikacje AI są zgodne z celami instytucji, wpływają na studentów i pracowników oraz przyczyniają się do podejmowania decyzji. Dzięki takim audytom systemy AI mogą być wdrażane sprawnie, sprawiedliwie i w sposób budujący zaufanie i rozliczalność.

Dobrze ustrukturyzowane audyty AI chronią prawa i prywatność studentów, wykładowców i pracowników, zapewniając zgodność z normami etycznymi i prawnymi. Budują zaufanie poprzez przejrzystość, rozliczalność i sprawiedliwe praktyki (Crompton i Burke, 2023). Wzmacniając proces podejmowania decyzji dzięki precyzyjnym i obiektywnym spostrzeżeniom, audyty pomagają instytucjom dostosować systemy AI do ich misji i wartości, minimalizować ryzyko, takie jak stronniczość i naruszenia danych, oraz zmniejszać odpowiedzialność reputacyjną i prawną. Ponadto audyty pomagają instytucjom utrzymać konkurencyjność poprzez odpowiedzialne wykorzystanie AI w celu poprawy wyników nauczania i efektywności operacyjnej, tworząc bardziej inkluzywne i efektywne środowiska edukacyjne (Fernsel i in., 2025).

Dzięki ustrukturyzowanym audytom AI instytucje mogą tworzyć bezpieczniejsze i bardziej przejrzyste systemy, zmniejszać ryzyko dyskryminacji algorytmicznej i budować większe zaufanie interesariuszy. Audyty umożliwiają również instytucjom wczesne wykrywanie słabości, udoskonalanie wewnętrznych polityk i zapewnianie zgodności z ewoluującymi ramami regulacyjnymi (Adeoye i in., 2025). Ostatecznie wspierają one rozwój etycznego, inkluzywnego i zrównoważonego ekosystemu AI w szkolnictwie wyższym, wzmacniając zaangażowanie instytucji w odpowiedzialną innowację i długoterminową doskonałość akademicką.





### 1.3. Przegląd kluczowych korzyści

W kontekście integracji AI, należy priorytetowo traktować kilka celów, aby zapewnić odpowiedzialne i skuteczne wdrożenie. Większa przejrzystość zapewnia jasne zrozumienie funkcjonowania systemów AI, umożliwiając interesariuszom zaufanie do ich procesów i rezultatów. Zwiększenie równości wymaga proaktywnej identyfikacji i ograniczania uprzedzeń, zapewniając sprawiedliwe i inkluzywne rezultaty. Budowanie zaufania interesariuszy jest kluczowe, wzmacniając zaufanie studentów, wykładowców i administratorów poprzez etyczne i odpowiedzialne praktyki. Wreszcie, wdrażanie zrównoważonych praktyk AI tworzy fundament ciągłego doskonalenia, gwarantując, że systemy pozostaną elastyczne, niezawodne i spójne z wartościami instytucji w czasie.

## 2. Koncepcje podstawowe

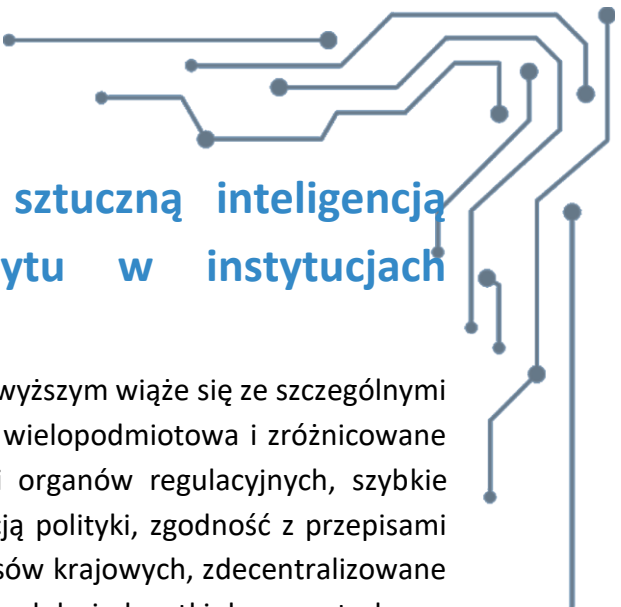
W tej sekcji przedstawiono teoretyczne i koncepcyjne podstawy audytu AI w instytucjach szkolnictwa wyższego. Wyjaśniono w niej kluczowe definicje, zasady przewodnie i kwestie zarządzania, niezbędne do odpowiedzialnej integracji, monitorowania i oceny systemów AI w edukacji, badaniach naukowych i administracji.

### 2.1. Definicja sztucznej inteligencji w szkolnictwie wyższym

Sztuczna inteligencja w szkolnictwie wyższym to systematyczna implementacja systemów obliczeniowych, które symulują ludzkie procesy poznawcze w celu ulepszenia, zautomatyzowania lub rozszerzenia funkcji instytucjonalnych w obszarach badań naukowych, edukacji i administracji, przy jednoczesnym zachowaniu standardów etycznych, prawnych i zgodności z instytucjami (Crompton i Song, 2021).

W szkołach wyższych zastosowania sztucznej inteligencji obejmują: (a) proces edukacyjny: adaptacyjne systemy nauczania, automatyczne oceny, analizę nauczania, spersonalizowane korepetycje; (b) proces badawczy: analizę danych, modelowanie predykcyjne, symulację, eksplorację literatury; (c) proces administracyjny: zarządzanie rekrutacją, alokacja zasobów, zasoby ludzkie, planowanie strategiczne.





## 2.2. Wyzwania w zarządzaniu sztuczną inteligencją i rozważania dotyczące audytu w instytucjach szkolnictwa wyższego

Integracja sztucznej inteligencji (AI) w szkolnictwie wyższym wiąże się ze szczególnymi wyzwaniami w zakresie zarządzania, takimi jak złożoność wielopodmiotowa i zróżnicowane interesy kadry naukowej, studentów, administratorów i organów regulacyjnych, szybkie zmiany technologiczne w porównaniu z powolną adaptacją polityki, zgodność z przepisami wymagającą dostosowania do RODO, ustawy o AI i przepisów krajowych, zdecentralizowane wdrażanie poprzez niezależne wdrażanie przez wydziały lub jednostki bez centralnego nadzoru oraz problemy transgraniczne związane z przepływami danych i konfliktami jurysdykcyjnymi we współpracy międzynarodowej. Sprostanie tym wyzwaniom w zakresie zarządzania wymaga skoordynowanego i proaktywnego podejścia, łączącego solidne ramy polityczne, współpracę międzywydziałową i adaptacyjne mechanizmy regulacyjne. Bez takiego dostosowania ryzyko etyczne, prawne i operacyjne związane ze sztuczną inteligencją (AI) w szkolnictwie wyższym może podważyć integralność instytucji i zaufanie interesariuszy (Al-Omari i in., 2025).

## 2.3. Zasady etyczne wdrażania sztucznej inteligencji w szkolnictwie wyższym

Etyczna integracja sztucznej inteligencji (AI) w szkolnictwie wyższym powinna opierać się na ustrukturyzowanym zestawie zasad, które chronią uczciwość akademicką, prawa człowieka i korzyści społeczne. Zasady te, zgodne z zaleceniami dotyczącymi etycznego wdrażania AI, są następujące:

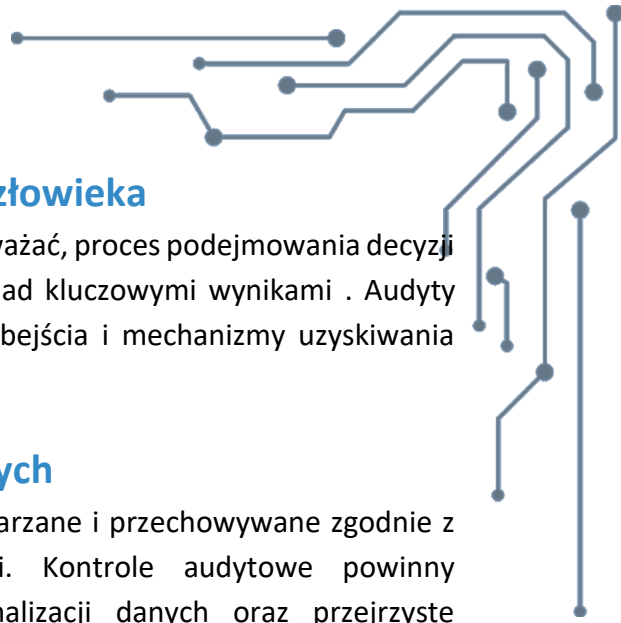
### 2.3.1. Odpowiedzialność i odpowiedzialność

Systemy AI powinny mieć jasno przypisany ludzki nadzór i odpowiedzialność za podejmowanie decyzji. Kryteria audytu powinny weryfikować obecność udokumentowanych mechanizmów odpowiedzialności, w tym kanałów raportowania w celu reagowania na ryzyka lub szkody związane z AI (Lazcoz i Hert, 2023).

### 2.3.2. Stronniczość i uczciwość

Sztuczna inteligencja musi być projektowana i monitorowana w sposób zapobiegający dyskryminacji i zapewniający równe traktowanie wszystkich grup użytkowników. Audyty powinny oceniać różnorodność zbiorów danych, procesy wykrywania stronniczości oraz środki naprawcze (Hasanzadeh i in., 2025).





### 2.3.3. Autonomia i sprawczość człowieka

Sztuczna inteligencja powinna wspierać, a nie podważać, proces podejmowania decyzji przez ludzi, zapewniając użytkownikom pełną kontrolę nad kluczowymi wynikami. Audyty powinny potwierdzać, że dostępne są opcje ręcznego obejścia i mechanizmy uzyskiwania zgody użytkownika ( Westphal i in., 2023).

### 2.3.4. Prywatność i ochrona danych

Dane osobowe powinny być gromadzone, przetwarzane i przechowywane zgodnie z obowiązującymi przepisami o ochronie prywatności. Kontrole audytowe powinny weryfikować stosowanie szyfrowania, praktyki minimalizacji danych oraz przejrzyste procedury uzyskiwania zgody ( Murdoch, 2021).

### 2.3.5. Bezpieczeństwo i ochrona

Oczekuje się, że systemy AI będą działać niezawodnie, minimalizując ryzyko dla bezpieczeństwa fizycznego, psychologicznego i cyfrowego. Audyty oceniają plany reagowania na incydenty, solidność systemu oraz ochronę przed atakami złośliwego oprogramowania ( Salhab i in., 2024).

### 2.3.6. Inkluzywność

Sztuczna inteligencja powinna być dostępna i użyteczna dla różnych grup społecznych, w tym grup marginalizowanych lub wrażliwych. Kryteria audytu powinny uwzględniać funkcje dostępności, obsługę wielojęzyczną oraz sprawiedliwy podział zasobów ( Acosta-Vargas i in., 2024).

### 2.3.7. Przejrzystość i wyjaśnialność

Zapewnienie, że procesy decyzyjne AI są zrozumiałe dla użytkowników i interesariuszy, jest kluczowe. Audyty weryfikują dostępność przejrzystej dokumentacji, narzędzi wyjaśniających i ujawniają rolę AI w osiągnięciu rezultatów ( Grimmelikhuisen, 2022).





## 2.4. Ramy prawne wdrażania sztucznej inteligencji w szkolnictwie wyższym

Uczelnie wyższe działają w złożonym środowisku regulacyjnym, regulującym etyczne i zgodne z prawem wykorzystanie systemów sztucznej inteligencji (AI). Ogólne Rozporządzenie o Ochronie Danych (RODO) ustanawia surowe wymogi dotyczące prywatności i bezpieczeństwa danych, a Ustawa o Prawach Edukacyjnych Rodziny i Prywatności (FERPA) chroni dokumentację studencką. Lokalne przepisy dodają obowiązki specyficzne dla danej jurysdykcji, tworząc wielowarstwowe systemy zgodności. Ustawa o AI klasyfikuje systemy według poziomu ryzyka – od niedopuszczalnego do minimalnego – i nakłada obowiązki na systemy wysokiego ryzyka, w tym transparentność, prowadzenie dokumentacji i ocenę ryzyka (Parlament Europejski i Rada, 2024; DataGuard, 2024).

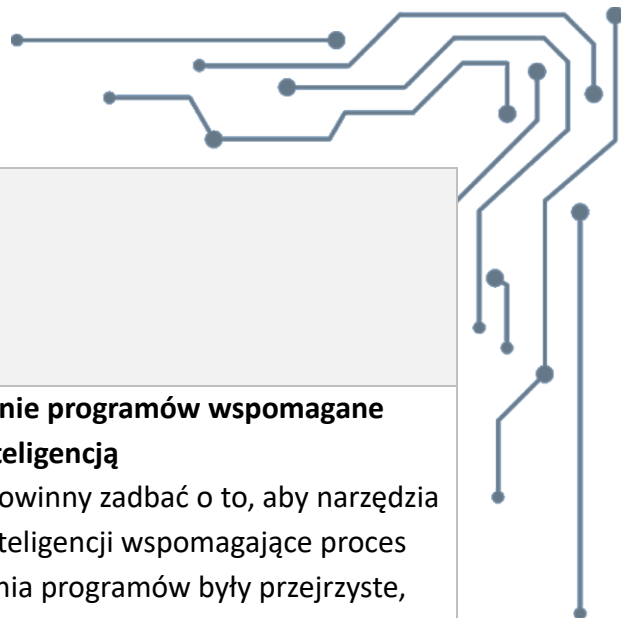
Dostosowanie regulacyjne powinno również objąć Standardy i Wytyczne Zapewnienia Jakości w Europejskim Obszarze Szkolnictwa Wyższego (ESG), które stanowią ramy zarówno dla wewnętrznego, jak i zewnętrznego zapewnienia jakości (ENQA, 2015). ESG nakłada na uczelnie obowiązek prowadzenia publicznie dostępnej, systematycznej polityki zapewniania jakości, zintegrowanej z zarządzaniem strategicznym. Kluczowe elementy obejmują projektowanie i zatwierdzanie programów, nauczanie i ocenę zorientowane na studenta, transparentny proces rekrutacji i certyfikacji, sprawiedliwą ocenę i rozwój kwalifikacji kadry dydaktycznej, odpowiednie zasoby edukacyjne i wsparcie dla studentów, efektywne zarządzanie informacją, transparentną komunikację instytucjonalną oraz regularne monitorowanie i przegląd programów.

Integracja sztucznej inteligencji (AI) ze szkolnictwem wyższym wymaga dostosowania tych zasad ESG, aby sprostać specyficznym dla AI wyzwaniom operacyjnym, etycznym i regulacyjnym, przy jednoczesnym zachowaniu jakości akademickiej. Ta adaptacja przekłada szeroko pojęte zasady zapewniania jakości na praktyczne wymagania dotyczące projektowania, wdrażania, monitorowania i ewaluacji AI. Dostosowane standardy koncentrują się na sprawiedliwości algorytmicznej, zarządzaniu danymi, przejrzystości, zaangażowaniu interesariuszy i ciągłym doskonaleniu jakości.

Tabela 1. Dostosowane do ESG standardy integracji sztucznej inteligencji dla instytucji szkolnictwa wyższego

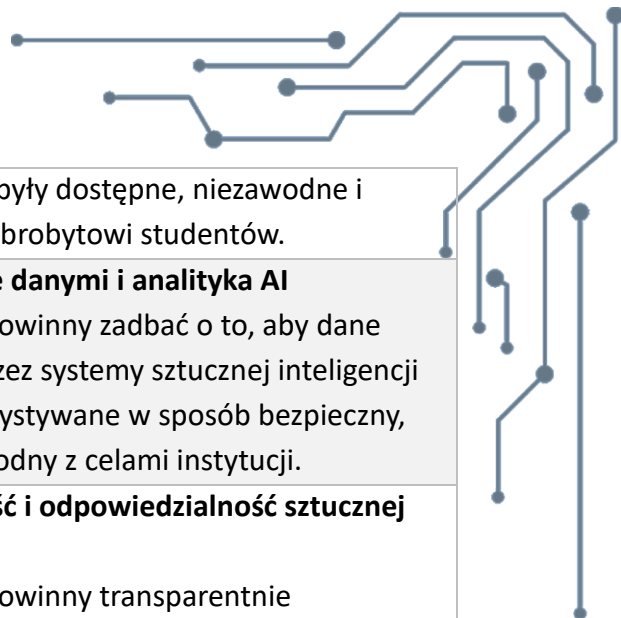
Standardy ESG	Standardy integracji sztucznej inteligencji dostosowane do ESG dla instytucji szkolnictwa wyższego
<b>1.1 Instytucje powinny posiadać politykę zapewniania jakości, która jest publicznie dostępna i stanowi element ich zarządzania strategicznego. Interesariusze wewnątrzni powinni opracować i wdrożyć tę politykę za pośrednictwem odpowiednich struktur i procesów, przy</b>	<b>Polityka zapewniania jakości sztucznej inteligencji</b> Instytucje powinny opracować jasne zasady dotyczące jakości, bezpieczeństwa i etycznego korzystania z systemów sztucznej inteligencji oraz włączyć te zasady do swojego zarządzania strategicznego.





<b>jednoczesnym zaangażowaniu interesariuszy zewnętrznych.</b>	
<b>1.2 Instytucje powinny posiadać procedury projektowania i zatwierdzania swoich programów. Programy powinny być projektowane w taki sposób, aby spełniały wyznaczone im cele, w tym zakładane efekty uczenia się. Kwalifikacje uzyskane w ramach programu powinny być jasno określone i zakomunikowane oraz odnosić się do właściwego poziomu krajowych ram kwalifikacji dla szkolnictwa wyższego, a w konsekwencji do Ram Kwalifikacji Europejskiego Obszaru Szkolnictwa Wyższego.</b>	<b>Projektowanie programów wspomagane sztuczną inteligencją</b> Instytucje powinny zadbać o to, aby narzędzia sztucznej inteligencji wspomagające proces projektowania programów były przejrzyste, obiektywne i zgodne z celami edukacyjnymi.
<b>1.3 Instytucje powinny zadbać o to, aby programy były realizowane w sposób zachęcający studentów do aktywnego uczestnictwa w procesie kształcenia, a ocena studentów powinna odzwierciedlać to podejście.</b>	<b>Spersonalizowane uczenie się wspomagane sztuczną inteligencją</b> Instytucje powinny zadbać o to, aby systemy nauczania oparte na sztucznej inteligencji były zorientowane na ucznia i zapewniały personalizację, jednocześnie chroniąc jego prywatność.
<b>1.4 Instytucje powinny konsekwentnie stosować ustalone wcześniej i opublikowane regulacje obejmujące wszystkie fazy „cyklu życia” studenta, np. przyjęcie, postępy, uznawanie i certyfikację.</b>	<b>Sztuczna inteligencja w zarządzaniu cyklem życia studenta</b> Uczelnie powinny zadbać o to, aby systemy sztucznej inteligencji stosowane w procesach rekrutacji, rozwoju i oceny studentów były uczciwe, przejrzyste i niedyskryminujące.
<b>1.5 Instytucje powinny upewnić się co do kompetencji swoich nauczycieli. Powinny stosować uczciwe i przejrzyste procedury rekrutacji i rozwoju kadry.</b>	<b>Rozwój kompetencji AI</b> Instytucje powinny wspierać kadre dydaktyczną w nabywaniu niezbędnych kompetencji umożliwiających etyczne i efektywne korzystanie z narzędzi sztucznej inteligencji.
<b>1.6 Instytucje powinny dysponować odpowiednimi środkami finansowymi na działalność dydaktyczną i edukacyjną oraz</b>	<b>Zarządzanie zasobami AI i wsparcie</b> Uczelnie powinny zadbać o to, aby usługi wsparcia studentów oparte na sztucznej





<b>zapewnić studentom odpowiednie i łatwo dostępne zasoby edukacyjne i wsparcie.</b>	inteligencji były dostępne, niezawodne i sprzyjały dobrobytowi studentów.
<b>1.7 Instytucje powinny zadbać o gromadzenie, analizowanie i wykorzystywanie stosownych informacji w celu efektywnego zarządzania swoimi programami i innymi działaniami.</b>	<b>Zarządzanie danymi i analityka AI</b> Instytucje powinny zadbać o to, aby dane zbierane przez systemy sztucznej inteligencji były wykorzystywane w sposób bezpieczny, etyczny i zgodny z celami instytucji.
<b>1.8 Instytucje powinny publikować informacje o swoich działaniach, w tym programach, które są jasne, dokładne, obiektywne, aktualne i łatwo dostępne.</b>	<b>Przejrzystość i odpowiedzialność sztucznej inteligencji</b> Instytucje powinny transparentnie udostępniać interesariuszom informacje na temat działania systemów sztucznej inteligencji, procesów decyzyjnych i zasad użytkowania.
<b>1.9 Instytucje powinny monitorować i okresowo weryfikować swoje programy, aby zapewnić, że osiągają one wyznaczone cele i odpowiadają potrzebom studentów i społeczeństwa. Weryfikacje te powinny prowadzić do ciągłego doskonalenia programu. Wszelkie działania planowane lub podejmowane w ich wyniku powinny być komunikowane wszystkim zainteresowanym.</b>	<b>Ciągły monitoring systemu AI</b> Instytucje powinny stale monitorować wydajność systemów sztucznej inteligencji, zgodność z zasadami etycznymi i dostosowanie do potrzeb społecznych, a następnie wdrażać odpowiednie ulepszenia.
<b>1.10 Instytucje powinny cyklicznie poddawać się zewnętrznemu zapewnianiu jakości zgodnie z ESG.</b>	<b>Zewnętrzny audyt i ocena AI</b> Instytucje powinny regularnie poddawać się niezależnym zewnętrznym audytom systemów sztucznej inteligencji i włączać je do procesów zapewniania jakości.

Dostosowane do wymogów ESG standardy integracji sztucznej inteligencji (AI) wypełniają lukę między tradycyjnymi ramami zapewniania jakości a złożonymi wymogami zarządzania sztuczną inteligencją, oferując instytucjom szkolnictwa wyższego sposób na wdrożenie technologii przy jednoczesnym zachowaniu zgodności z instytucjonalnymi standardami jakości i nowymi przepisami.

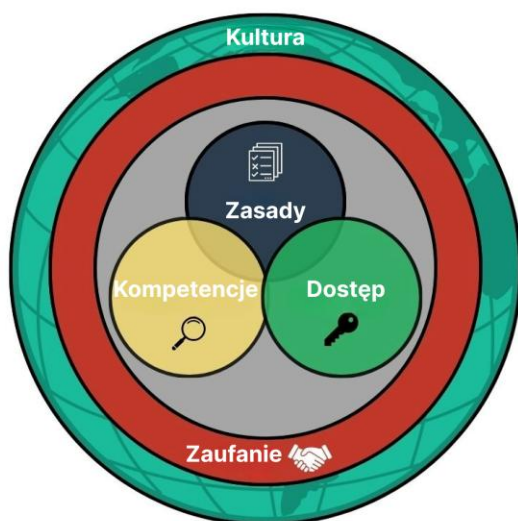


Co-funded by  
the European Union

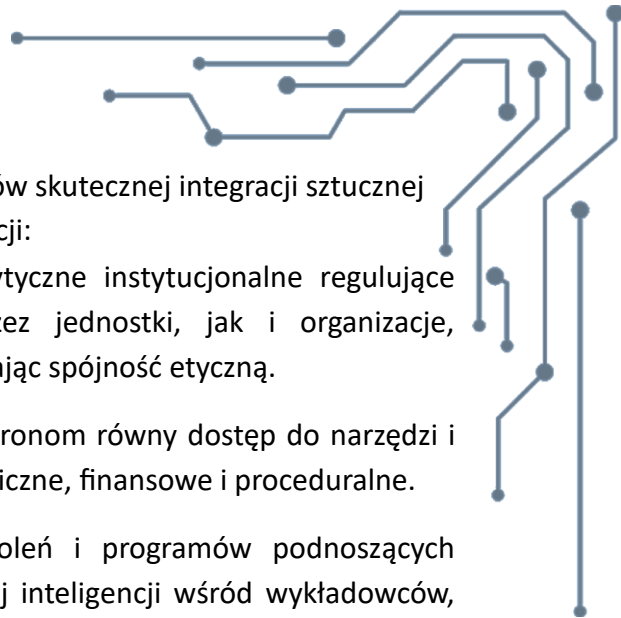
## 2.5. Zasady odpowiedzialnego korzystania ze sztucznej inteligencji

Odpowiedzialna integracja sztucznej inteligencji w instytucjach szkolnictwa wyższego wymaga nie tylko przestrzegania zasad etycznych, ale także stosowania praktycznych ram, które kierują codziennymi działaniami. Oprócz przestrzegania zasad przejrzystości, uczciwości, rozliczalności, prywatności i nadzoru ze strony człowieka, instytucje szkolnictwa wyższego korzystają ze strukturalnych podejść, które przekładają wartości na praktykę.

Jednym z takich podejść jest framework CRAFT (Liu i Bates, 2025), który identyfikuje pięć obszarów efektywnej integracji sztucznej inteligencji:



Rysunek 1 Struktura CRAFT – pięć podstawowych obszarów niezbędnych do bezpiecznego rozwiązania problemu Gen-AI w szkołach wyższych – dostosowane z (Liu & Bates, 2025)



W ramach tych ram wskazano pięć kluczowych obszarów skutecznej integracji sztucznej inteligencji w edukacji, badaniach naukowych i administracji:

- **Zasady** – ustanowienie jasnych zasad- zasady i wytyczne instytucjonalne regulujące korzystanie ze sztucznej inteligencji zarówno przez jednostki, jak i organizacje, zapobiegając niewłaściwemu wykorzystaniu i zapewniając spójność etyczną.
- **Dostęp** – zapewnienie wszystkim zainteresowanym stronom równy dostęp do narzędzi i zasobów sztucznej inteligencji, usuwając bariery techniczne, finansowe i proceduralne.
- **Kompetencje** – zapewnienie ukierunkowanych szkoleń i programów podnoszących świadomość w celu zwiększenia znajomości sztucznej inteligencji wśród wykładowców, pracowników i studentów, umożliwiając świadome i efektywne korzystanie z niej.
- **Zaufanie** – budowanie zaufania instytucjonalnego do systemów AI poprzez przejrzystość ich projektowania, wdrażania i wyników, wspieraną przez ciągłą ocenę
  - i zaangażowanie interesariuszy.
- **Kultura** – wspieranie pozytywnej kultury instytucjonalnej, która sprzyja innowacjom, a jednocześnie zachowuje uczciwość akademicką, wartości zorientowane na człowieka i inkluzywność.

Dzięki przyjęciu takich ustrukturyzowanych ram wdrażania wraz z wytycznymi etycznymi instytucje szkolnictwa wyższego mogą mieć pewność, że technologie sztucznej inteligencji będą integrowane w sposób zarówno skuteczny pod względem operacyjnym, jak i społecznie odpowiedzialny.

## 2.6. Przekształcanie zasad etycznych i ram prawnych w procesy audytowe

Audyt AI w szkolnictwie wyższym musi być kompleksowy i uwzględniać wskaźniki wydajności technicznej, zgodność z etyką, wymogi prawne oraz dostosowanie do polityki instytucjonalnej. To wielowymiarowe podejście gwarantuje, że AI wzmacnia, a nie podważa wartości szkolnictwa wyższego, przy jednoczesnym zachowaniu zaufania i integralności (Luo i in., 2025). Przełożenie zasad etycznych i wymogów prawnych na praktyczne procedury audytu jest kluczowym elementem zarządzania AI w instytucjach szkolnictwa wyższego. Cykliczne etapy rozwoju audytu przedstawiono na schemacie blokowym 1.





*Schemat blokowy 1. Proces rozwoju audytu*

Jak pokazano na schemacie blokowym 1, proces transformacji rozpoczyna się od określenia zasad etycznych, takich jak uczciwość, przejrzystość, odpowiedzialność i godność człowieka. Są one brane pod uwagę w kontekście wymogów prawnych wynikających z RODO, FERPA i innych, oraz Ustawy o sztucznej inteligencji (AI Act). Zasady te są następnie systematycznie przekształcane w mierzalne kryteria audytu. Rozwój narzędzi i metryk audytu umożliwia instytucjom wdrożenie tych zasad w praktyce, tworząc standardowe procedury oceny AI w różnych kontekstach.

Wdrożenie generuje dane empiryczne, które ujawniają luki w zakresie zgodności i ryzyka etyczne. Raportowanie ustaleń i rekomendacji generuje pętlę sprzężenia zwrotnego, rozwiązując bieżące problemy, a jednocześnie dopracowując zasady i kryteria na potrzeby przyszłych ocen (Brown i in., 2021). Ciągły monitoring i aktualizacja zapewniają, że ramy audytu pozostają elastyczne i dostosowane do zmieniających się standardów etycznych, obowiązków prawnych i możliwości technologicznych.

### 3. Ramy procesu audytu AI

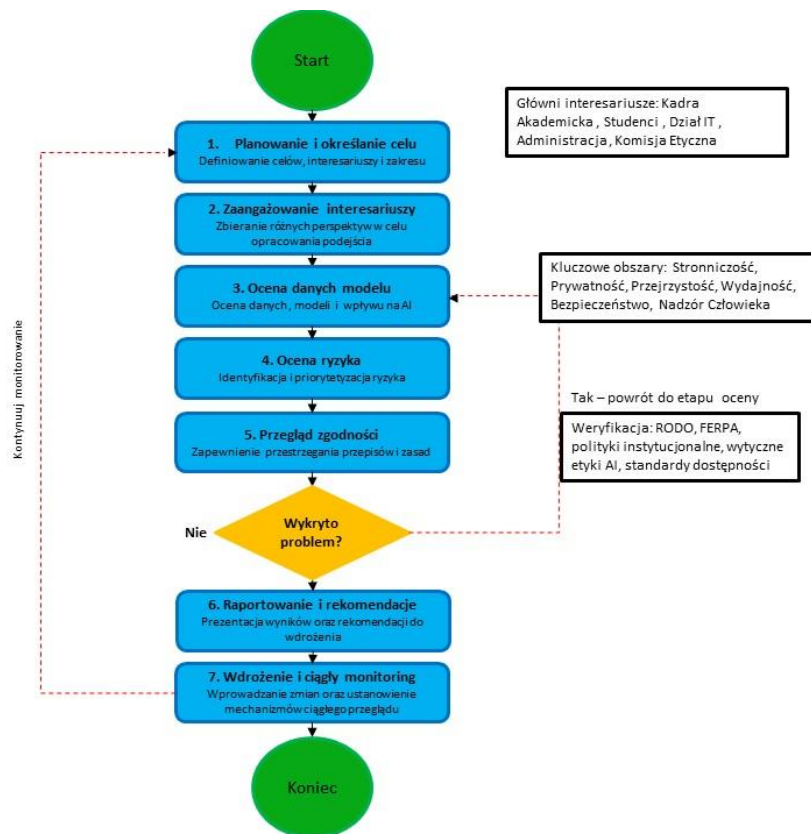
Wraz ze wzrostem znaczenia systemów sztucznej inteligencji (AI) w szkolnictwie wyższym, znacząco wzrosła potrzeba ustrukturyzowanej oceny i nadzoru. Systemy te – obejmujące algorytmy rekrutacyjne, platformy analityki nauczania, chatboty i narzędzia modelowania predykcyjnego – zmieniają funkcjonowanie instytucji, metodykę nauczania i wsparcie dla studentów. Choć takie innowacje przynoszą znaczne korzyści, wiążą się również z poważnymi zagrożeniami, takimi jak nasilenie uprzedzeń, naruszenia prywatności, nieprzejrzystość w procesie decyzyjnym oraz potencjalna niezgodność z wartościami instytucjonalnymi i ramami regulacyjnymi (Yan i Tang, 2025).

Audyt AI to ustrukturyzowana, multidyscyplinarna ocena, która bada nie tylko wydajność techniczną, ale także integralność etyczną, zgodność z prawem i zgodność z celami instytucjonalnymi. W przeciwieństwie do konwencjonalnych audytów IT, audyty AI dotyczą szerszych kwestii, takich jak prawa człowieka, uczciwość, przejrzystość, ochrona danych, integralność akademicka oraz przestrzeganie wymogów regulacyjnych, w tym ogólnego rozporządzenia o ochronie danych (RODO), ustawy o prawach edukacyjnych i prywatności rodziny (FERPA) oraz standardów dostępności (Farley i Lansang, 2025; Fernsel i in., 2025).

Biorąc pod uwagę złożoność audytu AI w szkolnictwie wyższym, spójna struktura wizualna jest niezbędna, aby przeprowadzić instytucje przez ten proces. Badania nad ramami audytowalności podkreślają znaczenie strukturyzacji ewaluacji wokół weryfikowalnych twierdzeń, dowodów potwierdzających i dostępnych mechanizmów walidacji (Fernsel i in., 2025). Szersze zarządzanie instytucjonalne AI często opiera się na takich zasadach, jak wyjaśnialność, rozliczalność, uczciwość i prywatność danych, odzwierciedlając modele proponowane w praktykach konsultingowych dotyczących odpowiedzialnego wdrażania AI w szkolnictwie wyższym (Huron, 2025).

Schemat blokowy 1 przedstawia siedmiofazową metodologię dostosowaną do potrzeb sektora. Ten ustrukturyzowany model wspiera zespoły audytorskie w poruszaniu się po kolejnych, powiązanych ze sobą etapach, identyfikowaniu krytycznych punktów decyzyjnych i utrzymywaniu spójności metodologicznej w różnych ocenach systemów AI. Dzięki wykorzystaniu pętli sprzężenia zwrotnego, model ten umożliwi instytucjom usuwanie niedociągnięć i wdrażanie mechanizmów ciągłego doskonalenia, zapewniając jednocześnie adaptację do różnych kontekstów instytucjonalnych i zastosowań AI.





Schemat blokowy 1. Fazy procesu audytu

Proces audytu AI przebiega w siedmiu powiązanych ze sobą fazach. Każda faza ma na celu zapewnienie kompleksowej oceny i ciągłego doskonalenia systemów AI w szkolnictwie wyższym. Rozpoczyna się od Planowania i Ustalania Zakresu, gdzie instytucje definiują cele audytu, określają systemy podlegające przeglądowi i identyfikują odpowiednich interesariuszy. Następnie następuje Zaangażowanie Interesariuszy, integrujące perspektywy studentów, wykładowców, administratorów, służb IT i komisji etycznych, aby uchwycić zróżnicowane obawy i oczekiwania. Faza Ewaluacji bada jakość danych, wydajność algorytmów, przejrzystość podejmowania decyzji oraz szerszy wpływ systemów AI na uczciwość akademicką, równość i efektywność instytucji. Analiza Ryzyka identyfikuje i priorytetyzuje potencjalne szkody lub zagrożenia w oparciu o ich prawdopodobieństwo i wagę, zapewniając jasne ramy dla działań łagodzących. Kontrola Zgodności weryfikuje zgodność z wymogami regulacyjnymi, takimi jak RODO, FERPA i Ustawa o AI, oraz zapewnia spójność z polityką instytucjonalną, obowiązkami dotyczącymi dostępności i wytycznymi etycznymi. Raportowanie syntetyzuje ustalenia w postaci praktycznych, dobrze ustrukturyzowanych rekomendacji dostosowanych do potrzeb kierownictwa i odpowiednich jednostek operacyjnych. Wreszcie, wdrażanie i monitorowanie ustanawia mechanizmy wdrażania działań korygujących, śledzenia wyników i tworzenia pętli sprzężenia zwrotnego, które pozwalają strukturalnym zarządcom na zachowanie adaptacji do zmieniających się technologii i priorytetów instytucjonalnych. Ten cykliczny proces gwarantuje, że audyty AI nie są jednorazowymi ocenami, lecz ciągłą, responsywną praktyką zarządzania.

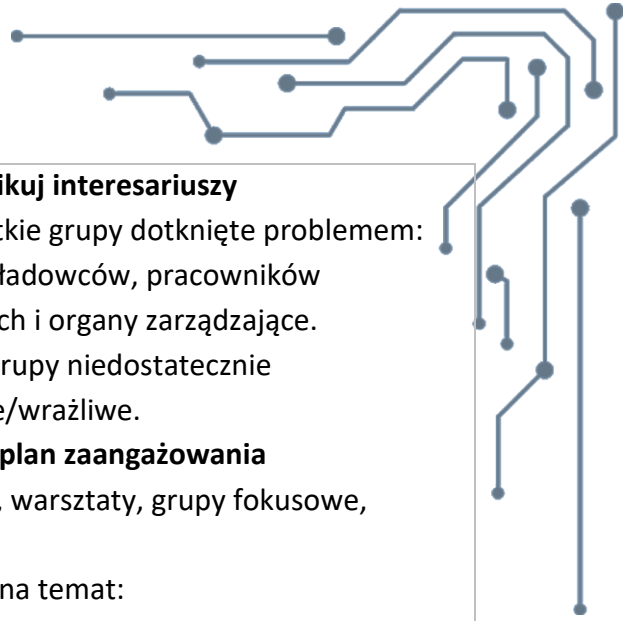
## 4. Szczegółowe instrukcje krok po kroku

W tej sekcji przedstawiono szczegółowe zadania związane z każdą fazą procesu audytu, jak pokazano na schemacie blokowym 1.

Tabela 2. Wskazówki krok po kroku

Faza	Zamiar	Szczegółowe kroki
<b>Faza 1 – Planowanie i określanie zakresu</b>	Określ zakres, cele i strukturę zarządzania audytem sztucznej inteligencji, aby zagwarantować przejrzystość, niezależność i zgodność z instytucjami.	<b>1. Zdefiniuj zakres</b> <ul style="list-style-type: none"><li>• Zidentyfikuj systemy sztucznej inteligencji, które należy poddać audytowi (np. algorytmy rekrutacyjne, narzędzia oceniające).</li><li>• Opisz cel każdego systemu, kontekst operacyjny i wpływ na instytucję.</li></ul> <b>2. Wyznacz cele</b> <ul style="list-style-type: none"><li>• Dostosuj cele do priorytetów instytucji.</li><li>• Weź pod uwagę standardy techniczne, etyczne, prawne i oparte na wartościach.</li></ul> <b>3. Zbierz zespół audytowy</b> <ul style="list-style-type: none"><li>• Uwzględnij przedstawicieli działów IT, prawnych, etyki i środowiska akademickiego.</li><li>• Zapewnienie niezależności: brak wcześniejszego zaangażowania w audytowany system.</li><li>• W miarę możliwości korzystaj z pomocy zewnętrznych ekspertów.</li><li>• Rotacja audytorów w audytach cyklicznych.</li></ul> <b>4. Opracuj plan projektu</b> <ul style="list-style-type: none"><li>• Harmonogramy, kamienie milowe, produkty końcowe.</li><li>• Koordynacja za pośrednictwem właściwych jednostek:<ul style="list-style-type: none"><li>○ Badania: B+R, Komisja Etyczna, Transfer Technologii.</li><li>○ Nauczanie: EdTech, Zapewnienie jakości akademickiej.</li><li>○ Administracja: HR, finanse, sprawy studenckie.</li></ul></li><li>• Zapewnienie odpowiedniego udziału ekspertów.</li></ul> <b>5. Zapewnij wsparcie kadry kierowniczej wyższego szczebla</b> <ul style="list-style-type: none"><li>• Uzyskaj zaangażowanie kierownictwa.</li><li>• Ustanowienie zasad zarządzania koordynacją międzywydziałową.</li></ul>





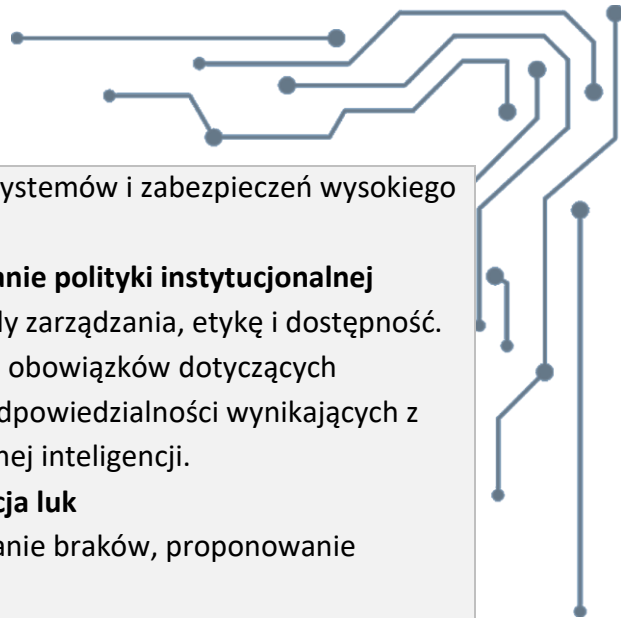
<b>Faza 2 – Zaangażowanie interesariuszy</b>	Upewnij się, że audyt odzwierciedla różne perspektywy, uwzględnia obawy i integruje doświadczenia użytkowników.	<b>1. Zidentyfikuj interesariuszy</b> <ul style="list-style-type: none"><li>• Zmapuj wszystkie grupy dotknięte problemem: studentów, wykładowców, pracowników administracyjnych i organy zarządzające.</li><li>• Uwzględniać grupy niedostatecznie reprezentowane/wrażliwe.</li></ul> <b>2. Opracuj plan zaangażowania</b> <ul style="list-style-type: none"><li>• Stosuj ankiety, warsztaty, grupy fokusowe, wywiady.</li><li>• Zbieraj opinie na temat:<ul style="list-style-type: none"><li>○ Uczciwość i przejrzystość.</li><li>○ Prywatność i bezpieczeństwo danych.</li><li>○ Dostępność i doświadczenie użytkownika.</li><li>○ Zrozumienie procesu decyzyjnego sztucznej inteligencji.</li><li>○ Oczekiwania co do nadzoru ze strony człowieka.</li><li>○ Wpływ na instytucje (wolność akademicka, zatrudnienie, kultura).</li><li>○ Konkretnie przypadki użycia i problemy.</li></ul></li></ul> <b>3. Przekształć opinię w działanie</b> <ul style="list-style-type: none"><li>• Doprecyzowanie zakresu i kryteriów audytu.</li><li>• Przejrzyj cele na podstawie priorytetów i ryzyka.</li><li>• Zaplanuj bieżącą komunikację i aktualizacje.</li></ul> <b>4. Opinie na temat dokumentów</b> <ul style="list-style-type: none"><li>• Bezpieczne, ustrukturyzowane rejestry.</li><li>• Powiąż opinie z kategoriami interesariuszy.</li></ul>
<b>Faza 3 – Ocena danych i modelu</b>	Ocenianie danych wejściowych (danych) i mechanizmów przetwarzania (modeli) pod kątem uczciwości, dokładności, przejrzystości i zgodności.	<b>1. Zdefiniuj zakres</b> <ul style="list-style-type: none"><li>• Dane: rejestry studentów, dane dotyczące wyników, dane demograficzne, analiza LMS, zestawy danych badawczych, rejestry administracyjne, dane szkoleniowe i operacyjne.</li><li>• Modele: modele ML, statystyczne, oparte na regułach, NLP, algorytmy rekomendacji.</li></ul> <b>2. Ocena jakości danych</b> <ul style="list-style-type: none"><li>• Sprawdź dokładność, kompletność i różnorodność.</li><li>• Zapewnienie reprezentacji demograficznej.</li><li>• Poprawianie brakujących/nieaktualnych/niespójnych danych.</li></ul>





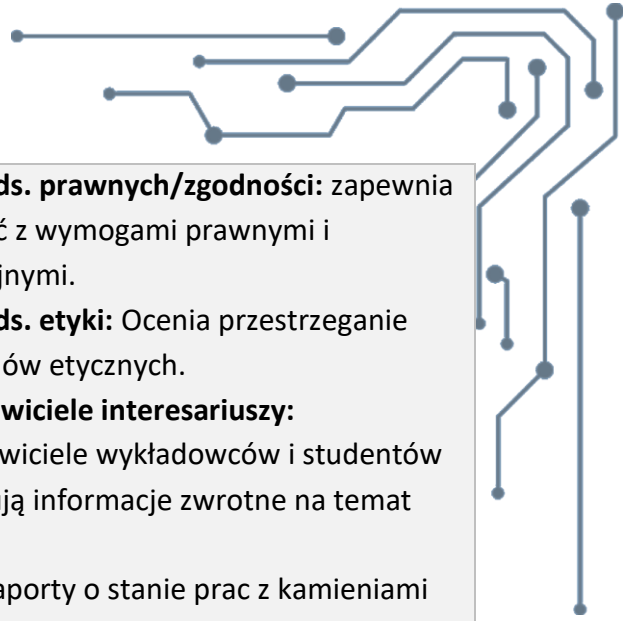
		<p><b>3. Analiza wydajności modelu</b></p> <ul style="list-style-type: none"><li>• Wykrywanie stronniczości w różnych grupach demograficznych.</li><li>• Weryfikacja dokładności, wskaźników błędów i wskaźników uczciwości.</li><li>• Testowanie wytrzymałości w różnych warunkach.</li><li>• Przejrzyj przejrzystość/możliwość wyjaśnienia.</li></ul> <p><b>4. Przegląd zarządzania i zgodności</b></p> <ul style="list-style-type: none"><li>• Sprawdź zgodność z RODO i FERPA.</li><li>• Zapewnienie zarządzania zgodami, limitów przechowywania, anonimizacji.</li><li>• Zgodność ze standardami etycznymi i dostępnością.</li></ul>
<b>Faza 4 – Ocena ryzyka</b>	Przekształć ustalenia w priorytetowe i możliwe do podjęcia działania ryzyka.	<p><b>1. Metodologia konsensusu ekspertów</b></p> <ul style="list-style-type: none"><li>• Przegląd interdyscyplinarny (techniczny, etyczny, prawny, operacyjny).</li><li>• Klasyfikacja kolaboracyjna.</li><li>• Zapewnienie niezależności od twórców systemu.</li></ul> <p><b>2. Ramy klasyfikacji</b></p> <ul style="list-style-type: none"><li>• Wysokie: zagrożenia dla reputacji, kwestii prawnych, bezpieczeństwa, integralności (naprawa trwa ≤30 dni).</li><li>• Średnie: nieefektywność, umiarkowane problemy z przestrzeganiem zasad, niezadowolenie (60–90 dni).</li><li>• Niskie: luki w użyteczności/dokumentacji (≥90 dni).</li></ul> <p><b>3. Rejestr Ryzyka</b></p> <ul style="list-style-type: none"><li>• Źródło danych, lista kontrolna błędów, uzasadnienie, interesariusze, wpływ, oceny ekspertów.</li></ul> <p><b>4. Planowanie łagodzenia skutków</b></p> <ul style="list-style-type: none"><li>• Określ działania, przypisz odpowiedzialność, ustal harmonogramy, sporządź listę zasobów.</li></ul> <p><b>5. Monitorowanie i eskalacja</b></p> <ul style="list-style-type: none"><li>• Punkty kontrolne postępu, wskaźniki sukcesu, zasady eskalacji.</li></ul>
<b>Faza 5 – Przegląd zgodności</b>	Sprawdź zgodność z przepisami, zasadami i etyką.	<p><b>1. Zgodność z przepisami prawnymi i regulacyjnymi</b></p> <ul style="list-style-type: none"><li>• RODO, FERPA, ustawa o sztucznej inteligencji (kategorie ryzyka).</li></ul>





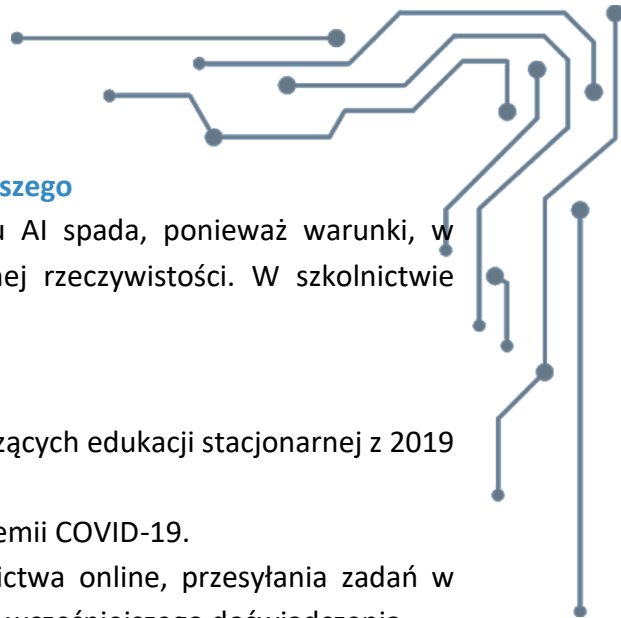
		<ul style="list-style-type: none"><li>• Identyfikacja systemów i zabezpieczeń wysokiego ryzyka.</li></ul> <p><b>2. Dostosowanie polityki instytucjonalnej</b></p> <ul style="list-style-type: none"><li>• Sprawdź zasady zarządzania, etykę i dostępność.</li><li>• Zintegrowanie obowiązków dotyczących przejrzystości/odpowiedzialności wynikających z ustawy o sztucznej inteligencji.</li></ul> <p><b>3. Identyfikacja luk</b></p> <ul style="list-style-type: none"><li>• Dokumentowanie braków, proponowanie aktualizacji.</li><li>• Zapewnienie zgodności z priorytetami prawnymi i strategicznymi.</li></ul> <p><b>4. Dokumentacja</b></p> <ul style="list-style-type: none"><li>• Przechowuj szczegółowe zapisy zgodności w bezpiecznym miejscu (ochrona prywatności w fazie projektowania).</li></ul>
<b>Faza 6 – Raportowanie i rekomendacje</b>	Komunikuj wyniki i kieruj ulepszeniami.	<p><b>1. Kompilacja wyników</b></p> <ul style="list-style-type: none"><li>• Ryzyka, luki w zgodności, wyniki wydajności.</li></ul> <p><b>2. Praktyczne rekomendacje</b></p> <ul style="list-style-type: none"><li>• Techniczne (dane, modele) + zarządzanie (polityki, nadzór).</li><li>• Ustal priorytety i harmonogramy.</li></ul> <p><b>3. Streszczenie</b></p> <ul style="list-style-type: none"><li>• Strategiczne skupienie się na przywództwie.</li><li>• Podkreśl najważniejsze ryzyka i zapotrzebowanie na zasoby.</li></ul> <p><b>4. Ochrona danych</b></p> <ul style="list-style-type: none"><li>• Przechowywanie, szyfrowanie, kontrola dostępu, ścieżki audytu zgodne z przepisami RODO.</li></ul>
<b>Faza 7 – Wdrożenie i ciągły monitoring</b>	Wdrażaj i utrzymuj ulepszenia w czasie.	<p><b>1. Działania korygujące</b></p> <ul style="list-style-type: none"><li>• Przypisz role za pomocą macierzy RACI, aby określić, kto jest odpowiedzialny, rozliczany, konsultowany i informowany za każde zadanie, a także śledzić kamienie milowe. Na przykład:<ul style="list-style-type: none"><li>○ <b>Kierownik projektu:</b> koordynuje działania i raportuje postępy.</li><li>○ <b>Ekspert techniczny:</b> wdraża aktualizacje sztucznej inteligencji, w tym ponowne szkolenie i zmiany kodu.</li></ul></li></ul>





		<ul style="list-style-type: none"><li>○ <b>Ekspert ds. prawnych/zgodności:</b> zapewnia zgodność z wymogami prawnymi i regulacyjnymi.</li><li>○ <b>Ekspert ds. etyki:</b> Ocenia przestrzeganie standardów etycznych.</li><li>○ <b>Przedstawiciele interesariuszy:</b> Przedstawiciele wykładowców i studentów przekazują informacje zwrotne na temat zmian.</li></ul> <ul style="list-style-type: none"><li>• Tygodniowe raporty o stanie prac z kamieniami kontrolnymi, wczesnymi wskaźnikami ryzyka i monitorowaniem harmonogramu w celu zapobiegania opóźnieniom.</li></ul> <p style="text-align: center;"><b>2. Monitorowanie i metryki</b></p> <ul style="list-style-type: none"><li>• Zdefiniuj kluczowe wskaźniki efektywności (KPI) dotyczące uczciwości, przejrzystości, dokładności i zgodności.</li><li>• Stosuj wskaźniki ryzyka.</li></ul> <p style="text-align: center;"><b>3. Częstotliwość przeglądu</b></p> <ul style="list-style-type: none"><li>• Wysoki: 6 miesięcy.</li><li>• Średni: 1 rok.</li><li>• Niski: 2 lata.</li><li>• Nadzwyczajne kontrole w razie konieczności.</li></ul> <p style="text-align: center;"><b>4. Wykrywanie dryfu modelu</b></p> <ul style="list-style-type: none"><li>• Monitorowanie dryfu koncepcji/danych.</li><li>• W razie potrzeby przetrenować model lub dokonać zmian.</li><li>• Przykładowe czynniki wyzwajające: nauka hybrydowa, zmiany gospodarcze, oszukiwanie przy pomocy sztucznej inteligencji, zmiany na rynku pracy.</li></ul> <p style="text-align: center;"><b>5. Eskalacja</b></p> <ul style="list-style-type: none"><li>• Niskie: problem zostanie rozwiązany w ciągu 30 dni.</li><li>• Średni: naprawa w ciągu 90 dni.</li><li>• Wysoki: pilna interwencja.</li></ul> <p style="text-align: center;"><b>6. Ciągłe doskonalenie</b></p> <ul style="list-style-type: none"><li>• Wykorzystaj wyniki do udoskonalenia zasad, wzmocnienia nadzoru i poprawy zabezpieczeń.</li></ul>
--	--	---





## Scenariusze dryfu modeli w instytucjach szkolnictwa wyższego

Dryf modelu występuje, gdy wydajność systemu AI spada, ponieważ warunki, w których był on trenowany, nie odpowiadają już obecnej rzeczywistości. W szkolnictwie wyższym może się to objawiać na kilka sposobów:

### Scenariusz 1 – System przewidywania sukcesów uczniów

**Sytuacja początkowa:** Model trenowany na danych dotyczących edukacji stacjonarnej z 2019 r.

**Zmiana:** Przejście na nauczanie hybrydowe w czasie pandemii COVID-19.

**Skutek:** Model nie jest w stanie interpretować uczestnictwa online, przesyłania zadań w formie cyfrowej i egzaminów wirtualnych z powodu braku wcześniejszego doświadczenia.

**Wynik:** Wskaźnik błędów w przewidywaniach wzrasta o 30–40%.

### Scenariusz 2 – Algorytm przydziału stypendiów

**Sytuacja początkowa:** Model trenowany w okresie stabilności gospodarczej.

**Zmiana:** Kryzys gospodarczy powodujący wzrost bezrobocia i zmiany w strukturze dochodów rodzin.

**Skutek:** Model nie rozpoznaje nowych wskaźników potrzeb finansowych.

**Wynik:** niesprawiedliwy podział stypendiów i nieefektywna alokacja zasobów.

### Scenariusz 3 – System wykrywania uczciwości akademickiej

**Sytuacja początkowa:** Model wytrenowany w celu wykrywania tradycyjnych metod oszukiwania.

**Zmiana:** Powszechne stosowanie narzędzi sztucznej inteligencji, takich jak ChatGPT.

**Skutek:** Model nie jest w stanie wykryć plagiatu ani ghostwritingu wspomaganego przez sztuczną inteligencję.

**Wynik:** Istotne braki w przestrzeganiu uczciwości akademickiej.

### Scenariusz 4 – System rekomendacji kursów

**Sytuacja początkowa:** Model wytrenowany na podstawie preferencji studentów sprzed pandemii.

**Zmiana:** Zmiana popytu na nowe kursy technologiczne, odpowiadające zmieniającym się potrzebom rynku pracy.

**Skutek:** Model nadal rekomenduje przestarzałe wzorce kursów.

**Wynik:** Zalecenia nie odpowiadały obecnym wymaganiom w zakresie umiejętności.

### Zalecenie:

W ramach ciągłego monitorowania instytucje powinny ustanowić mechanizm wykrywania dryfu modelu, aby szybko identyfikować i reagować na spadki wydajności. Gwarantuje to, że systemy AI zachowają dokładność, rzetelność i znaczenie dla instytucji. Trójstopniowe podejście do monitorowania zostało szczegółowo opisane w rozdziale 5.3.1 (Ocena techniczna systemów/narzędzi AI).



## 5. Kroki audytu

Poniższe ramy przedstawiają systematyczne podejście do przeprowadzania audytów AI w instytucjach szkolnictwa wyższego. Przedstawiają one fazy przygotowawcze, które stanowią podstawę skutecznej ewaluacji. Proces audytu rozpoczyna się od jasnego zdefiniowania celów i oczekiwań instytucji wobec audytu, a następnie następują kompleksowe działania przygotowawcze, obejmujące określenie zakresu, identyfikację interesariuszy i powołanie zespołu audytowego.

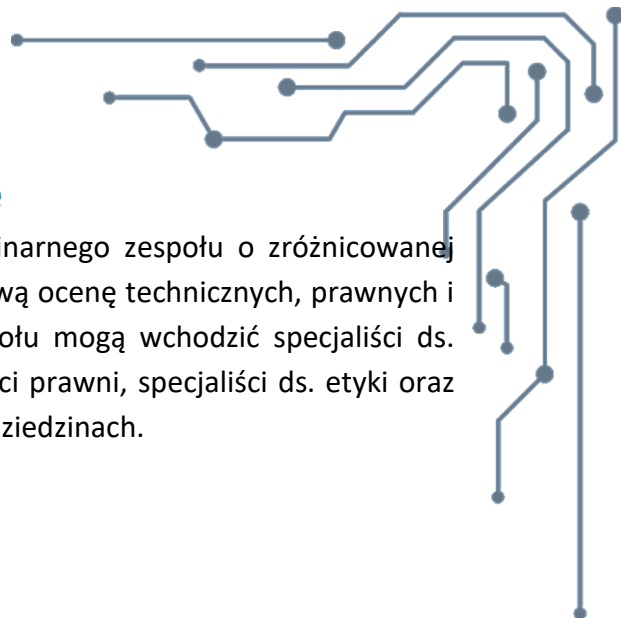
### 5.1. Przygotowanie do audytu

Przygotowanie do audytu polega przede wszystkim na zdefiniowaniu celów audytu i stworzeniu zespołów z jasno przypisanymi zadaniami.

#### 5.1.1. Określanie celów audytu

- Jakie konkretne cele stawia sobie Twoja instytucja w ramach tego audytu?
- Jaki jest charakter audytu:
  - Audyt koncentruje się na sprawdzeniu, czy korzystanie przez instytucję ze sztucznej inteligencji jest zgodne z zasadami etycznymi i wymogami prawnymi;
  - Audyt koncentruje się na ocenie, czy instytucja dysponuje strukturami, politykami, zasobami i kompetencjami umożliwiającymi etyczne i skuteczne korzystanie ze sztucznej inteligencji;
  - Audyt ma na celu ocenę stopnia zrozumienia przez studentów, pracowników, badaczy i administratorów sposobu wykorzystania sztucznej inteligencji w danej instytucji.
  - Wszystko powyższe  
Inny:.....
- Dostosuj cele audytu do:  
Kodeksu etyki instytucji – Wewnętrznych systemów zapewnienia jakości –  
Standardów ESG, w szczególności:
  - ESG 1.1 (Polityka zapewnienia jakości)
  - ESG 1.3 (Nauka i ocena skoncentrowane na uczniu)
  - ESG 1.5 (Kadra dydaktyczna – kompetencje i kompetencje w zakresie sztucznej inteligencji)
  - ESG 1.7 (Zarządzanie informacją i wykorzystanie danych generowanych przez sztuczną inteligencję)





### 5.1.2. Zbierz zespoły audytorskie

Skuteczny audyt AI wymaga multidyscyplinarnego zespołu o zróżnicowanej wiedzy specjalistycznej, który zapewni kompleksową ocenę technicznych, prawnych i etycznych aspektów systemów AI. W skład zespołu mogą wchodzić specjaliści ds. zapewnienia jakości, specjaliści techniczni, eksperci prawni, specjaliści ds. etyki oraz specjaliści ds. AI specjalizujący się w konkretnych dziedzinach.

#### Skład zespołu:

- Specjalista w zakresie jakości
- Specjalista IT
- Doradcy prawni
- Ekspert – Komisja Etyczna
- Eksperci ds. systemów AI/badań/edukacji/administracji do konkretnych celów audytowych

Przypisz role swojemu zespołowi:

1. Imię: \_\_\_\_\_ Rola: \_\_\_\_\_  
2. Imię: \_\_\_\_\_ Rola: \_\_\_\_\_



## 5.2. Audyt etycznego wykorzystania sztucznej inteligencji

Te ramy audytu etycznego obejmują trzy specjalistyczne listy kontrolne audytu, mające na celu ocenę etycznej integracji systemów sztucznej inteligencji w podstawowych funkcjach szkolnictwa wyższego: listę kontrolną audytu badań wspomaganych sztuczną inteligencją, listę kontrolną audytu edukacji wspomaganej sztuczną inteligencją oraz listę kontrolną audytu procesów administracyjnych wspomaganych sztuczną inteligencją. Każda lista kontrolna odpowiada na kluczową potrzebę systematycznego nadzoru nad wdrażaniem sztucznej inteligencji, zapewniając zgodność postępu technologicznego z wartościami edukacyjnymi i odpowiedzialnością instytucji.

Listy kontrolne opierają się na siedmiu fundamentalnych zasadach etycznych: odpowiedzialności i obowiązkowości, stronniczości i uczciwości, autonomii i sprawczości człowieka, prywatności i ochrony danych, bezpieczeństwa, inkluzywności oraz przejrzystości i zrozumiałości. Każdy element oceny jest systematycznie powiązany ze standardami ESG, specjalnie dostosowanymi do integracji sztucznej inteligencji (AI) w instytucjach szkolnictwa wyższego, zapewniając solidne ramy do oceny ryzyka, opracowywania polityk, monitorowania zgodności i ciągłego doskonalenia praktyk wdrażania AI.

### Instrukcje dotyczące typowej listy kontrolnej

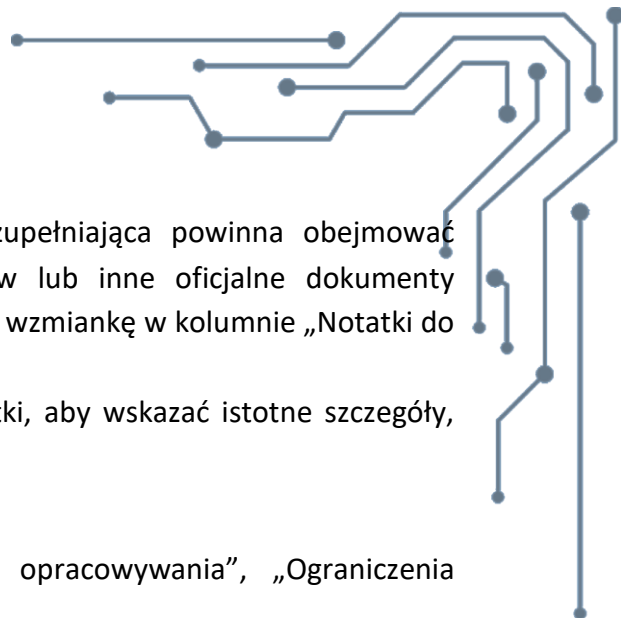
Dla każdej pozycji wybierz tylko jedną opcję: „Tak”, „Nie” lub „N/D” (nie dotyczy).

- **„Tak”**: Praktyka/polityka istnieje i jest udokumentowana weryfikowalnymi dowodami
- **„Nie”**: Praktyka/polityka nie istnieje lub nie posiada odpowiedniej dokumentacji
- **„N/D”**: Pozycja nie ma zastosowania do obecnego wykorzystania sztucznej inteligencji w Twojej instytucji

### Wymagania listy kontrolnej

- Opieraj wszystkie odpowiedzi na udokumentowanych faktach i bieżących praktykach, a nie na założeniach lub planowanych działaniach
- W przypadku odpowiedzi „Nie” należy dołączyć wewnętrzne notatki wskazujące planowane działania naprawcze
- Odpowiedz na wszystkie pytania, które mają zastosowanie; żadna pozycja nie powinna pozostać pusta
- Po zakończeniu przeprowadź pełną kontrolę pod kątem dokładności i spójności





## Weryfikacja i notatki

W przypadku odpowiedzi „Tak”: Dokumentacja uzupełniająca powinna obejmować polityki, procedury, zapisy szkoleń, raporty z audytów lub inne oficjalne dokumenty instytucjonalne potwierdzające zgodność. Proszę o krótką wzmiankę w kolumnie „Notatki do listy kontrolnej”.

W przypadku odpowiedzi „Nie”: Użyj kolumny Notatki, aby wskazać istotne szczegóły, takie jak:

- Zaplanowane działania naprawcze i harmonogramy
- Powody obecnych luk (np. „Polityka w trakcie opracowywania”, „Ograniczenia budżetowe”, „Ograniczenia techniczne”)
- Poziom priorytetu w celu rozwiązania problemu niedoboru
- Dział/osoba odpowiedzialna za działania następcze
- Czy obowiązują jakiegokolwiek środki tymczasowe?

W przypadku odpowiedzi „N/D” można dodać krótkie wyjaśnienia, jeśli będą przydatne w przyszłości.

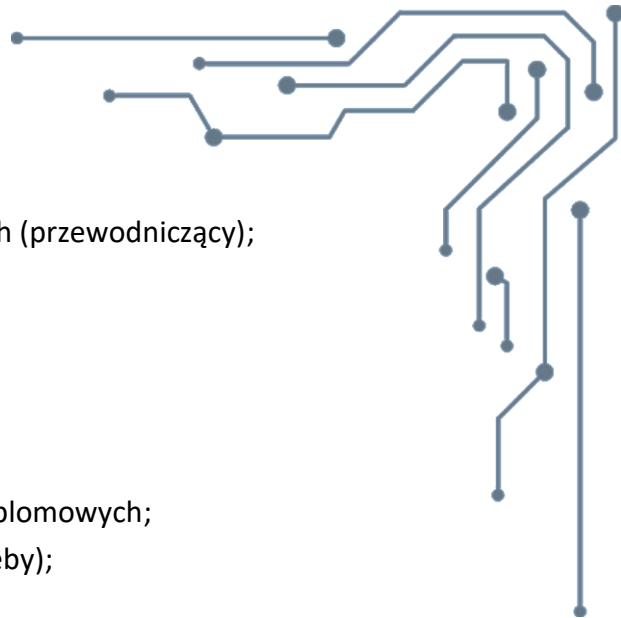
### 5.2.1 Audyt procesu badawczego wspomaganego sztuczną inteligencją

W tej sekcji dokonano oceny integracji i zarządzania technologiami AI w ramach instytucjonalnych działań badawczych. Audyt sprawdza, czy narzędzia AI wykorzystywane w procesach badawczych są zgodne ze standardami rzetelności akademickiej, wytycznymi etycznymi w zakresie badań oraz przepisami o ochronie danych. Kluczowe obszary zainteresowania obejmują dostępność polityk instytucjonalnych dotyczących wykorzystania AI w badaniach, dostępność i inkluzywność narzędzi AI dla wszystkich środowisk badawczych, przejrzystość metodologii badawczych opartych na AI oraz wdrożenie odpowiednich mechanizmów szkoleniowych i informacji zwrotnej. Gwarantuje to, że AI przyczynia się do produktywności badań, jednocześnie zachowując rygor naukowy i standardy etyczne. Poniższe zalecenia mogą służyć jako punkt odniesienia przy określaniu częstotliwości audytów i składu zespołu audytującego badania wspomagane AI.

#### Zalecane okresy audytu

- Systemy AI o znaczeniu krytycznym dla badań: rocznie;
- Narzędzia ogólnego wsparcia badań: co 2 lata;
- Audyty specyficzne dla projektu: na początku i na końcu projektu;
- Audyty inicjowane: w przypadku wątpliwości etycznych, naruszeń danych lub poważnych zarzutów dotyczących niewłaściwego postępowania badawczego;





### Zaleceni członkowie zespołu audytowego

- Przewodniczący Komisji ds. Etyki Badań Naukowych (przewodniczący);
- Przedstawiciel Biura Rozwoju Badań;
- Inspektor Ochrony Danych;
- Starszy członek wydziału badawczego;
- Specjalista ds. bezpieczeństwa informacji;
- Doradca ds. praw własności intelektualnej;
- Przedstawiciel ds. badań studentów studiów podyplomowych;
- Zewnętrzny ekspert ds. etyki badań (w razie potrzeby);

### LISTA KONTROLNA DOTYCZĄCA AUDYTU BADAŃ WSPOMAGANYCH SZTUCZNĄ INTELIGENCJĄ

Ta kompleksowa lista kontrolna audytu składa się z trzech kolejnych sekcji, które należy wypełnić zgodnie z podanymi wytycznymi:

- **Informacje o audycie:** Udokumentuj szczegóły instytucjonalne, skład zespołu audytorskiego, zakres i grupy docelowe.
- **Lista kontrolna audytu badań wspomaganych sztuczną inteligencją:** Systematyczna ocena integracji sztucznej inteligencji w sześciu kluczowych obszarach (odpowiedzialność, stronniczość i uczciwość, autonomia i sprawczość człowieka, ochrona prywatności i danych, bezpieczeństwo, inkluzywność oraz przejrzystość i wyjaśnialność).
- **Wyniki oceny badań wspomaganych sztuczną inteligencją i priorytetyzacja analizy ryzyka :** ocena zidentyfikowanych ryzyk i opracowanie strategii działań korygujących.

**Instrukcja:** Uzupełnij każdą sekcję w podanej kolejności, upewniając się, że wszystkie odpowiednie pola są wypełnione, a właściwe pola wyboru są zaznaczone, zanim przejdziesz do następnej sekcji.

#### 1. INFORMACJE O AUDYCIE

Instytucja/Organizacja: \_\_\_\_\_

Audyt przeprowadzony przez:

- Instytucja wewnętrzna/Dział: \_\_\_\_\_
- Instytucja/Dział zewnętrzny: \_\_\_\_\_

Lider zespołu audytorskiego:

\_\_\_\_\_

Stanowisko/tytuł:

\_\_\_\_\_

Członkowie zespołu audytorskiego: \_\_\_\_\_

\_\_\_\_\_

Okres audytu: Od \_\_\_\_\_ do \_\_\_\_\_

Data zakończenia audytu: \_\_\_\_\_



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them.  
Grant no:2024-1-LT01-KA220-HED-000251565

Grupy docelowe/objęte wydziały: \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

Przed wypełnieniem listy kontrolnej instytucje powinny najpierw rozważyć następujące pytanie dotyczące warunku wstępnego.

**Pytanie wstępne:** Czy Twoja instytucja wykorzystuje obecnie zintegrowane systemy sztucznej inteligencji w procesach badawczych, stosując ustalone procedury lub zasady?

Tak  Nie

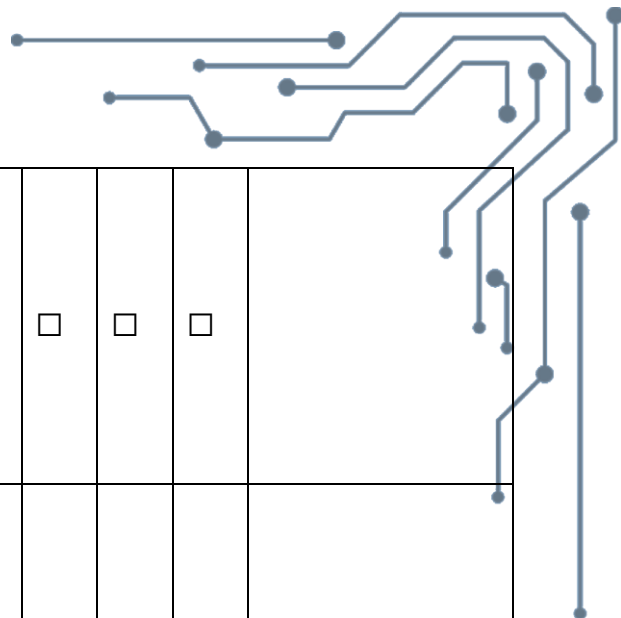
Jeśli „Nie”, ta lista kontrolna nie ma zastosowania. Jeśli „Tak”, przejdź do pełnej oceny .

## 2. LISTA KONTROLNA AUDYTU BADAŃ WSPOMAGANYCH SZTUCZNĄ INTELIGENCJĄ

Tabela 3. Lista kontrolna audytu badań wspomaganých sztuczną inteligencją

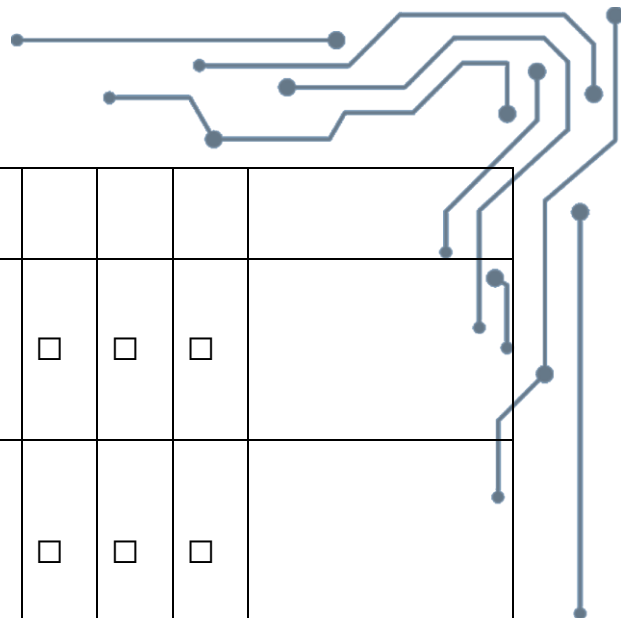
Elementy		Powiązane standardy ESG	Tak	Nie	Nie dotyczy	Notatki
Zakres	Czy istnieją strategie i polityki instytucjonalne mające na celu integrację narzędzi sztucznej inteligencji z procesami badawczymi?	1.1, 1.2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Oceniane są potencjalne zagrożenia (np. ochrona danych, stronniczość, kwestie własności intelektualnej) związane z wykorzystaniem sztucznej inteligencji?	1.4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy zasady dotyczące praw własności intelektualnej w kontekście korzystania z narzędzi sztucznej inteligencji w procesie badawczym są jasno określone?	1.1, 1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Stronniczość i uczciwość	Czy istnieje udokumentowana procedura instytucjonalna dotycząca przeglądu zbiorów danych pod kątem różnorodności i reprezentacji przed ich wykorzystaniem w badaniach wspieranych przez sztuczną inteligencję?	1.1, 1.7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





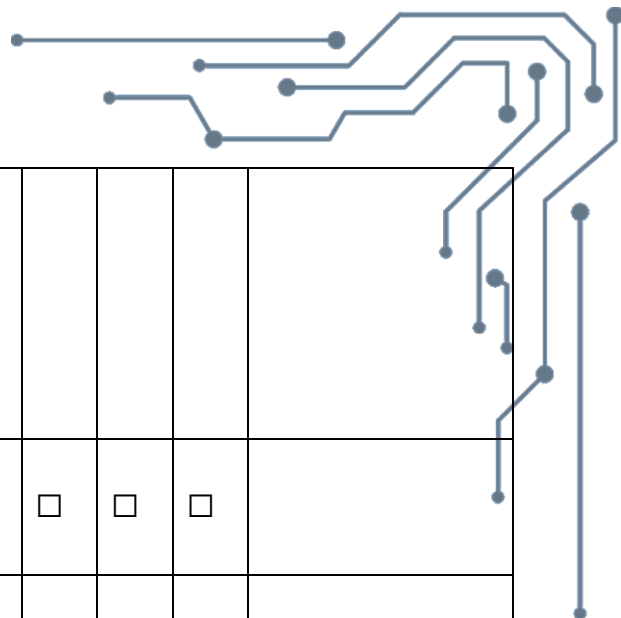
	Czy badacze są zobowiązani do dokumentowania i zgłaszania wszelkich zaobserwowanych błędów systematycznych lub obaw dotyczących uczciwości podczas badań wspomaganych przez sztuczną inteligencję?	<b>1.1, 1.8</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy istnieje proces monitorowania i dokumentowania działań podejmowanych w przypadku zidentyfikowania problemów związanych z uprzedzeniami lub uczciwością w badaniach wspomaganych sztuczną inteligencją? (nowe)	<b>1.1, 1.9</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Autonomia i sprawczość człowieka</b>	Czy istnieje udokumentowana procedura zapewniająca, że zgoda lub interwencja człowieka jest obowiązkowa przed zastosowaniem w badaniach rekomendacji wygenerowanych przez sztuczną inteligencję?	<b>1.8</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy badacze mają obowiązek dokumentować, w jaki sposób ostateczne decyzje w badaniach wspomaganych przez sztuczną inteligencję pozostają pod nadzorem człowieka?	<b>1.8</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy instytucja zapewnia systematyczne szkolenia, które pozwalają badaczom krytycznie oceniać i, w razie potrzeby, ignorować wyniki badań AI?	<b>1,5</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Prywatność i ochrona danych</b>	Czy wprowadzono zasady bezpieczeństwa danych dotyczące wykorzystania narzędzi sztucznej inteligencji w badaniach naukowych?	<b>1.1</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy dane osobowe lub poufne przekazane sztucznej inteligencji są anonimizowane lub w inny sposób	<b>1.7</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





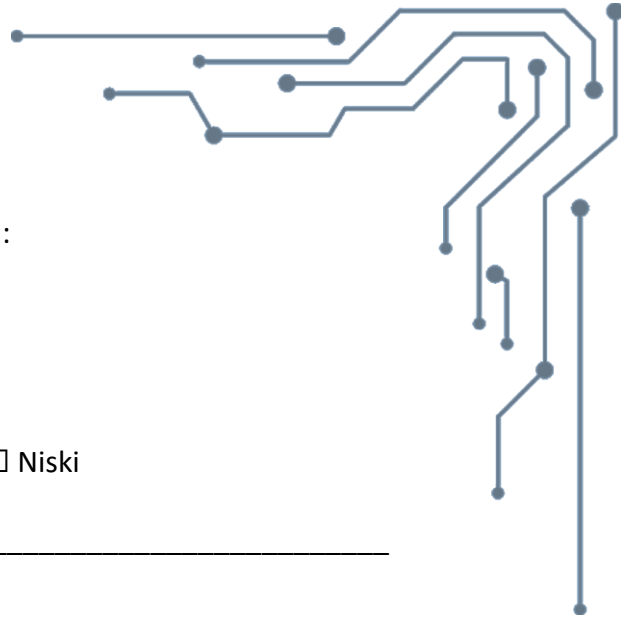
	chronione odpowiednią dokumentacją?					
	Czy zasady przechowywania i usuwania danych badawczych przetwarzanych przez sztuczną inteligencję są jasno określone?	<b>1.1, 1.7</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Bezpieczeństwo i ochrona</b>	Czy modele sztucznej inteligencji zostały przetestowane pod kątem dokładności, niezawodności i ryzyka nieoczekiwanych wyników przed wdrożeniem w badaniach?	<b>1.9</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy incydenty wpływające na wiarygodność i wiarygodność wyników badań wspomaganych sztuczną inteligencją są dokumentowane i czy podejmowane są działania naprawcze?	<b>1.7, 1.8</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy instytucja opracowała z góry określone kroki mające na celu rozwiązanie potencjalnych problemów związanych ze sztuczną inteligencją w badaniach?	<b>1.7</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy wdrożono procedury reagowania na potencjalne naruszenia bezpieczeństwa w badaniach wspomaganych sztuczną inteligencją?	<b>1.7</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy istnieje protokół zapobiegający lub łagodzący potencjalną utratę danych w badaniach wspomaganych sztuczną inteligencją?	<b>1.7</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy narzędzia AI są regularnie aktualizowane, aby zachować standardy bezpieczeństwa i stawić czoła nowym zagrożeniom w procesach badawczych?	<b>1.9</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	<b>Inkluzywność</b>	Czy narzędzia sztucznej inteligencji wykorzystywane w badaniach naukowych są dostępne dla grup defaworyzowanych (np. studentów niepełnosprawnych, osób znajdujących się w niekorzystnej	<b>1.3</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>





	sytuacji społeczno-ekonomicznej, mniejszości językowych, studentów będących uchodźcami lub migrantami oraz osób z obszarów wiejskich lub oddalonych) zgodnie z zasadą inkluzywności?				
	Czy oceniono inkluzywność narzędzi AI pod kątem obsługi różnych języków?	<b>1.3, 1.6</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Czy narzędzia AI są oceniane pod kątem dostępności, zwłaszcza dla użytkowników niepełnosprawnych lub z innymi utrudnieniami?	<b>1.3, 1.6</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Czy przyjazność dla użytkownika narzędzi AI jest systematycznie oceniana w celu zapewnienia inkluzywności?	<b>1.3, 1.6</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Czy oceniono inkluzywność wyników generowanych przez sztuczną inteligencję pod kątem tego, czy wzmacniają one czy łagodzą istniejące nierówności w badaniach naukowych?	<b>1.7</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>Przejrzystość i wyjaśnialność</b>	Czy udostępniono jasne informacje na temat źródeł danych wykorzystanych do opracowania narzędzi sztucznej inteligencji i ich ogólnej wydajności?	<b>1.7, 1.8</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Czy wyniki generowane przez sztuczną inteligencję są jasno udokumentowane w sposób zrozumiały dla badaczy?	<b>1.8</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Czy udokumentowano ograniczenia i znane słabości modeli sztucznej inteligencji?	<b>1.2, 1.5, 1.6</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Czy te udokumentowane ograniczenia i słabości są udostępniane odpowiednim interesariuszom badań?	<b>1.2, 1.5, 1.6</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>





Wymagany plan analizy i łagodzenia ryzyka:  Tak  Nie

Jeśli tak, szczegóły analizy ryzyka i planu łagodzenia ryzyka:

- Odpowiedzialny za opracowanie planu:  
\_\_\_\_\_
- Docelowa data zakończenia działań łagodzących:  
\_\_\_\_\_
- Poziom priorytetu ogólnego:  Wysoki  Średni  Niski

Data przeglądu kontrolnego: \_\_\_\_\_

### 3. WYNIKI OCENY BADAŃ WSPOMAGANE PRZEZ SZTUCZNĄ INTELIGENCJĘ I PRIORYTETYZACJA ANALIZY RYZYKA

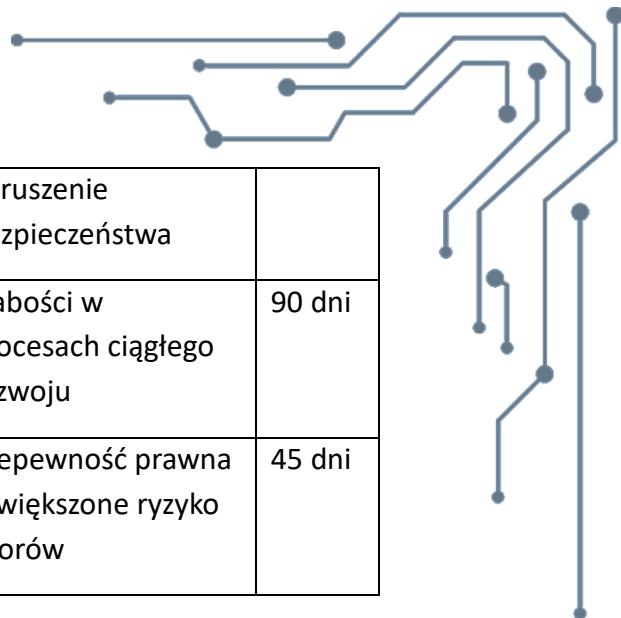
Po wypełnieniu listy kontrolnej audytu badań wspomaganym sztuczną inteligencją, kolejnym kluczowym krokiem jest analiza i priorytetyzacja wszelkich zidentyfikowanych luk lub niedociągnięć. Poniższa tabela priorytetyzacji ryzyka przekłada ustalenia z listy kontrolnej na priorytety, kategoryzując problemy według ich potencjalnego wpływu i pilności dla integralności i zgodności badań.

Dzięki takiemu systematycznemu podejściu najpoważniejsze luki w zabezpieczeniach są eliminowane w pierwszej kolejności, a jednocześnie wyznaczany jest jasny harmonogram kompleksowego usprawniania ram zarządzania badaniami wspomaganymi przez sztuczną inteligencję.

Tabela 4. Priorytetyzacja ryzyka – procesy badawcze

Zidentyfikowany problem	Obszar	Poziom ryzyka	Racjonalne uzasadnienie	Pilność
Brak wytycznych etycznych dotyczących korzystania ze sztucznej inteligencji	Etyka badań	Wysoki	Zagrożenie dla rzetelności badań i reputacji akademickiej	Pilny
Brak szkoleń w zakresie sztucznej inteligencji dla badaczy	Rozwój potencjału	Średni	Ryzyko niewłaściwego wykorzystania i obniżenia jakości badań	60 dni
Niewystarczające bezpieczeństwo danych badawczych	Bezpieczeństwo danych	Wysoki	Utrata własności intelektualnej i potencjalne	Pilny





dotyczących sztucznej inteligencji			naruszenie bezpieczeństwa	
Nie zebrano opinii od badaczy	Ciągłe doskonalenie	Niski	Słabości w procesach ciągłego rozwoju	90 dni
Niejasne zasady dotyczące własności intelektualnej	Ramy prawne	Średni	Niepewność prawna i zwiększone ryzyko sporów	45 dni

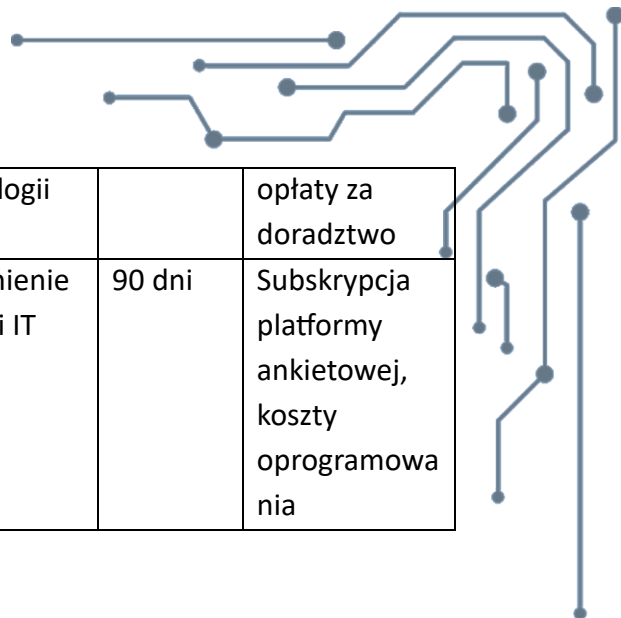
### Analiza i ograniczanie ryzyka procesu badawczego wspomaganego sztuczną inteligencją

Najpoważniejsze zagrożenia zidentyfikowane w obszarze badań to brak wytycznych etycznych i niedociągnięcia w zakresie bezpieczeństwa danych. Słabości te bezpośrednio podważają uczciwość akademicką i reputację instytucji.

Tabela 5. Plan łagodzenia skutków – procesy badawcze

Wydanie	Poziom docelowy	Strategia interwencji	Odpowiedzialny	Termin ostateczny	Wymagania dotyczące zasobów
Niedobór wytycznych etycznych dotyczących sztucznej inteligencji	Kompleksowy dokument wytycznych etycznych	Powołanie komisji interdyscyplinarnej i opracowanie wytycznych	Komisja Etyczna i Biuro Badań Naukowych	30 dni	Zewnętrzna ekspertyza etyczna, opłaty za konsultacje
Bezpieczeństwo danych badawczych	Struktura bezpieczeństwa zgodna z normą ISO 27001	Wzmocnienie infrastruktury bezpieczeństwa	Biuro Bezpieczeństwa Informatycznego i Badań	25 dni	Licencje na oprogramowanie zabezpieczające, koszty infrastruktury
Szkolenie AI badaczy	≥ 90% wskaźnik uczestnictwa	Obowiązkowy program edukacji w zakresie sztucznej inteligencji	Rozwój i edukacja badaczy	60 dni	Materiały szkoleniowe, koszty opracowania programu
Niepewność własności intelektualnej	Przejrzysta i aktualna polityka IP	Doradztwo prawne i rewizja polityki	Biuro Radców Prawnych i Transferu	45 dni	Usługi doradztwa prawnego,





			Technologii (TTO)		opłaty za doradztwo
Mechanizm sprzężenia zwrotnego	System systematycznego sprzężenia zwrotnego	System ankiet online, protokół analizy	Zapewnienie jakości i IT	90 dni	Subskrypcja platformy ankietowej, koszty oprogramowania

### 5.2.2. Audyt procesu edukacji wspomaganej sztuczną inteligencją

W tej sekcji dokonano oceny wdrażania i zarządzania systemami sztucznej inteligencji (AI) w środowiskach nauczania i uczenia się. Audyt kładzie szczególny nacisk na strategię instytucjonalną dotyczącą integracji AI w edukacji, równy dostęp do możliwości uczenia się wspomaganego AI oraz ochronę bezpieczeństwa i prywatności danych studentów. Kluczowe obszary analizy obejmują ustanowienie jasnych wytycznych dotyczących wykorzystania AI w działalności akademickiej, zapewnienie szkoleń z zakresu etyki zarówno dla pracowników, jak i studentów oraz skuteczność mechanizmów informacji zwrotnej w celu ciągłego doskonalenia. Celem tej oceny jest zapewnienie, że technologie AI wspierają doskonałość edukacyjną, jednocześnie chroniąc dobro studentów i uczciwość akademicką. Poniższe zalecenia mogą służyć jako punkt odniesienia przy ustalaniu częstotliwości audytów i odpowiedniego składu zespołów audytorskich w przypadku procesów edukacyjnych wspomaganego przez sztuczną inteligencję.

#### Zalecane okresy audytu

- Systemy sztucznej inteligencji do oceny uczniów: Rocznie
- Systemy zarządzania nauczaniem i analityki/Narzędzia wspomagające edukację: co dwa lata
- Audyty inicjowane: w odpowiedzi na skargi studentów, obawy dotyczące uczciwości akademickiej lub problemy z dostępnością

#### Zaleceni członkowie zespołu audytowego

- Przedstawiciel Biura Spraw Akademickich (kierownik);
- Specjalista ds. technologii edukacyjnych;
- Rzecznik ds. Studenckich;
- Koordynator ds. rozwoju wydziału;
- Inspektor Ochrony Danych;
- Przedstawiciel ds. Dostępności;
- Przedstawiciel studentów;

#### LISTA KONTROLNA DO AUDYTU EDUKACJI WSPOMAGANEJ SZTUCZNĄ INTELIGENCJĄ



Co-funded by  
the European Union



Ta kompleksowa lista kontrolna audytu składa się z trzech kolejnych sekcji, które należy wypełnić zgodnie z podanymi wytycznymi:

- **Informacje o audycie** : Udokumentuj szczegóły instytucjonalne, skład zespołu audytorskiego, zakres i grupy docelowe.
- **Lista kontrolna audytu edukacji wspomaganej sztuczną inteligencją** : systematyczna ocena integracji sztucznej inteligencji w sześciu kluczowych obszarach (odpowiedzialność, stronniczość i uczciwość, autonomia i sprawczość człowieka, ochrona prywatności i danych, bezpieczeństwo, inkluzywność oraz przejrzystość i wyjaśnialność).
- **Wyniki oceny edukacji wspomaganej sztuczną inteligencją i priorytetyzacja ryzyka** : ocena zidentyfikowanych ryzyk i opracowanie strategii działań naprawczych, skupiających się na uczciwości akademickiej, równości w edukacji i jakości nauczania.

**Instrukcja:** Uzupełnij każdą sekcję w podanej kolejności, upewniając się, że wszystkie odpowiednie pola są wypełnione, a właściwe pola wyboru są zaznaczone, zanim przejdziesz do następnej sekcji.

## 1. INFORMACJE O AUDYCIE

Instytucja/Organizacja: \_\_\_\_\_

Audyt przeprowadzony przez:

- Instytucja wewnętrzna/Dział: \_\_\_\_\_
- Instytucja/Dział zewnętrzny: \_\_\_\_\_

Lider zespołu audytorskiego:

\_\_\_\_\_

Stanowisko/tytuł:

\_\_\_\_\_

Członkowie zespołu audytorskiego: \_\_\_\_\_

\_\_\_\_\_

Okres audytu: Od \_\_\_\_\_ do \_\_\_\_\_

Data zakończenia audytu: \_\_\_\_\_

Grupy docelowe/objęte wydziały: \_\_\_\_\_

\_\_\_\_\_

Przed wypełnieniem listy kontrolnej instytucje powinny najpierw rozważyć następujące pytanie dotyczące warunku wstępnego .

**Pytanie wstępne:** Czy Twoja instytucja wykorzystuje obecnie zintegrowane systemy sztucznej inteligencji w procesach edukacyjnych, stosując ustalone procedury lub zasady?

Tak  Nie

Jeśli „Nie”, ta lista kontrolna nie ma zastosowania. Jeśli „Tak”, przejdź do pełnej oceny.

## 2. LISTA KONTROLNA AUDYTU EDUKACJI WSPOMAGANEJ SZTUCZNĄ INTELIENCJĄ



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them.  
Grant no:2024-1-LT01-KA220-HED-000251565

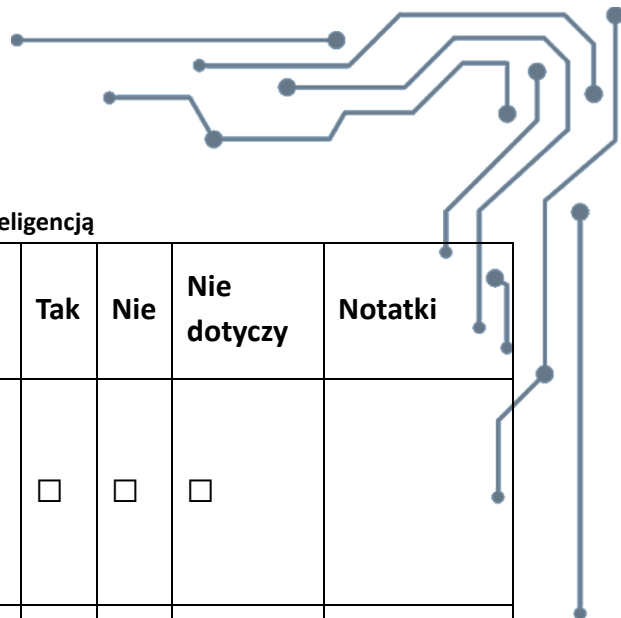
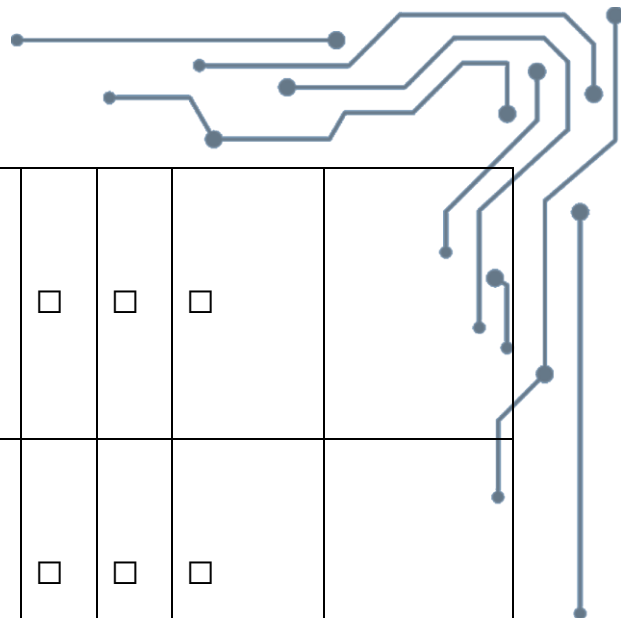


Tabela 6. Lista kontrolna audytu edukacji wspomaganej sztuczną inteligencją

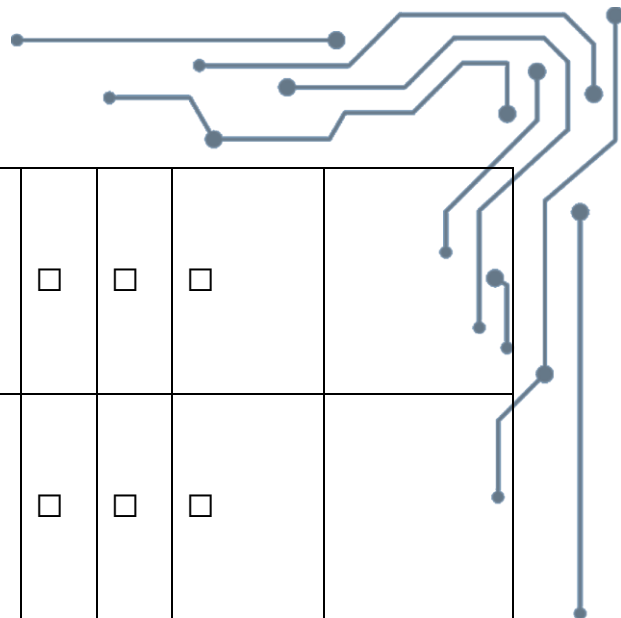
Zakres		Powiązany standard ESG.	Tak	Nie	Nie dotyczy	Notatki
Odpowiedzialność	Czy istnieją strategie i polityki instytucjonalne mające na celu integrację narzędzi sztucznej inteligencji z procesami nauczania i uczenia się?	1.1, 1.2, 1.4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy ustalono jasne i wyraźne zasady dotyczące wykorzystania sztucznej inteligencji podczas zadań i zajęć lekcyjnych ?	1.6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy są dostępne programy szkoleniowe mające na celu podniesienie świadomości etycznej kadry naukowej i studentów w zakresie korzystania ze sztucznej inteligencji?	1,5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy studenci mają dostęp do mechanizmu zgłaszania problemów napotkanych podczas korzystania z narzędzi AI?	1.3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy nauczyciele dysponują mechanizmem zgłaszania problemów napotkanych podczas korzystania z narzędzi AI?	1.3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Stronniczość i uczciwość	Czy treści edukacyjne i rekomendacje generowane przez sztuczną inteligencję są regularnie sprawdzane pod kątem tego, czy nie utrwalają stereotypów lub uprzedzeń kulturowych?	1.2, 1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





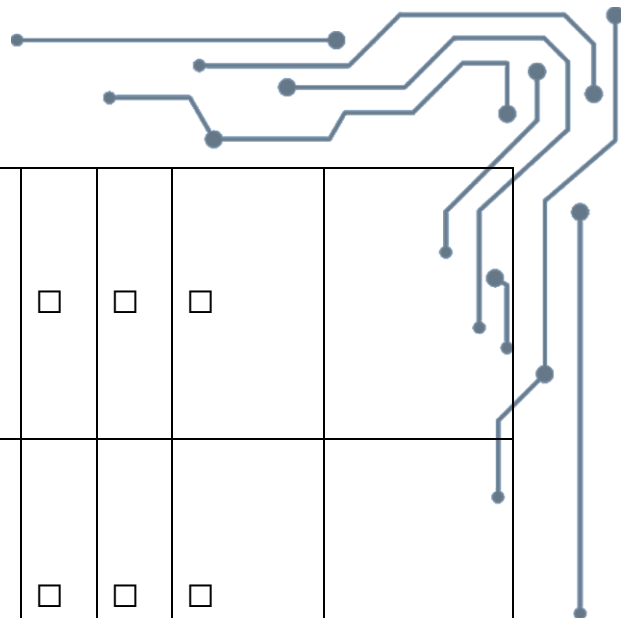
	Czy narzędzia sztucznej inteligencji stosowane w edukacji zapewniają równe szanse uczniom pochodzącym z różnych środowisk społeczno-ekonomicznych i kulturowych?	1.3, 1.4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Autonomia i sprawczość człowieka</b>	Czy uczniowie są jasno informowani o prawie akceptowania lub odrzucania sugestii generowanych przez sztuczną inteligencję podczas zajęć edukacyjnych?	1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy uczniom zapewnia się w trakcie zajęć edukacyjnych wyraźne możliwości krytycznej oceny i kwestionowania treści generowanych przez sztuczną inteligencję?	1,5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Prywatność i ochrona danych</b>	Czy wprowadzono zasady bezpieczeństwa i poufności danych dotyczące wykorzystania narzędzi sztucznej inteligencji w edukacji?	1.6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy uczniowie są jasno informowani o tym, w jaki sposób ich dane osobowe będą gromadzone, wykorzystywane i przechowywane, gdy korzystają z narzędzi edukacyjnych wspomaganych przez sztuczną inteligencję?	1.7, 1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy odpowiedni pracownicy (np. kadra naukowa i administracyjna) są jasno informowani o tym, w jaki sposób ich dane osobowe będą gromadzone, wykorzystywane i przechowywane, gdy korzystają z narzędzi edukacyjnych wspomaganych przez sztuczną inteligencję?	1.7, 1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





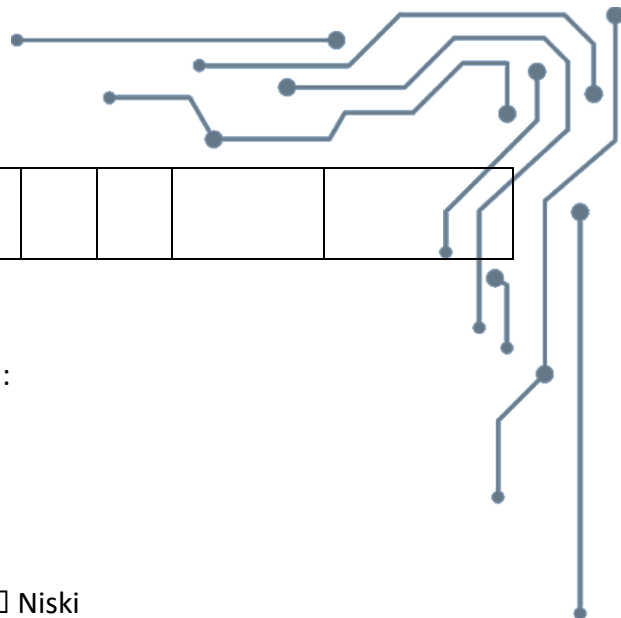
	Czy uczniowie wyrażają wyraźną zgodę na wykorzystanie swoich danych osobowych w zajęciach edukacyjnych wspomaganych sztuczną inteligencją?	1.7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Bezpieczeństwo i ochrona</b>	Czy istnieją ustalone procedury zapewniające ciągłość nauczania i uczenia się w przypadku awarii narzędzi AI lub przerw w świadczeniu usług?	1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy wdrożono środki mające na celu ochronę procesów edukacyjnych przed potencjalnym niewłaściwym wykorzystaniem narzędzi sztucznej inteligencji przez studentów i kadrę dydaktyczną?	1.1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Inkluzywność</b>	Czy narzędzia AI są dostępne dla grup defaworyzowanych (np. uczniów niepełnosprawnych, osób znajdujących się w niekorzystnej sytuacji społeczno-ekonomicznej, mniejszości językowych, uczniów będących uchodźcami lub migrantami oraz osób z obszarów wiejskich lub oddalonych) zgodnie z zasadą inkluzywności?	1.3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy są prowadzone programy szkoleniowe mające na celu zwiększenie świadomości etycznej studentów w zakresie korzystania ze sztucznej inteligencji?	1,5, 1,6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy są prowadzone programy szkoleniowe mające na celu podniesienie świadomości etycznej kadry akademickiej w zakresie korzystania ze sztucznej inteligencji?	1,5, 1,6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





	Czy materiały edukacyjne wspomagane sztuczną inteligencją są projektowane tak, aby uwzględniać zróżnicowane potrzeby i style uczenia się?	1.3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy przykłady, studia przypadków i treści zawarte w materiałach edukacyjnych wspomaganych sztuczną inteligencją są reprezentatywne dla różnorodnych środowisk kulturowych i społecznych?	1.2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Przejrzystość i wyjaśnialność</b>	Czy zbierane są opinie od studentów na temat skuteczności i przydatności narzędzi sztucznej inteligencji w procesie uczenia się?	1.7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy zbierane są opinie kadry akademickiej na temat aspektów pedagogicznych wykorzystania narzędzi sztucznej inteligencji w edukacji?	1.7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy zbierane są opinie kadry akademickiej na temat aspektów etycznych stosowania narzędzi sztucznej inteligencji w edukacji?	1.7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy podejmowane są działania pedagogiczne mające na celu ulepszenie środowisk edukacyjnych wspomaganych przez sztuczną inteligencję?	1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy podejmowane są działania etyczne mające na celu ulepszenie środowisk edukacyjnych wspieranych przez sztuczną inteligencję?	1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy rola i wkład narzędzi sztucznej inteligencji w proces uczenia się są jasno	1.8, 1.2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





	komunikowane uczestnikom na początku kursu lub zajęć?					
--	---	--	--	--	--	--

Wymagany plan analizy i łagodzenia ryzyka:  Tak  Nie

Jeśli tak, szczegóły analizy ryzyka i planu łagodzenia ryzyka:

- Odpowiedzialny za opracowanie planu:

\_\_\_\_\_

- Docelowa data zakończenia działań łagodzących:

\_\_\_\_\_

- Poziom priorytetu ogólnego:  Wysoki  Średni  Niski

Data przeglądu kontrolnego: \_\_\_\_\_

### 3. WYNIKI OCENY EDUKACJI WSPOMAGANEJ SZTUCZNĄ INTELIGENCJĄ I PRIORYTYZACJA RYZYKA

Po wypełnieniu listy kontrolnej audytu edukacji wspomaganej sztuczną inteligencją (AI-Assisted Education Checklist), niezbędna jest ocena i priorytyzacja zidentyfikowanych niedociągnięć lub luk w instytucjonalnym zarządzaniu AI w edukacji. Poniższa tabela priorytyzacji ryzyka przekłada ustalenia audytu na strategiczny plan działania, klasyfikując problemy według ich potencjalnego wpływu na uczciwość akademicką, równość studentów i jakość edukacji.

Tabela 7. Priorytyzacja ryzyka – procesy edukacyjne

Zidentyfikowany problem	Obszar	Poziom ryzyka	Racjonalne uzasadnienie	Pilność
Brak świadomości uczniów na temat zasad korzystania ze sztucznej inteligencji	Integralność akademicka	Wysoki	Ryzyko nieuczciwości akademickiej, niesprawiedliwego oceniania	Pilny
Nierówny dostęp uczniów do narzędzi AI	Równość edukacyjna	Wysoki	Ryzyko nierówności szans i dyskryminacji	Pilny
Brak szkoleń kadry w zakresie etyki sztucznej inteligencji	Jakość nauczania	Średni	Niewystarczające wskazówki, niespójne stosowanie	60 dni
Niedociągnięcia w zakresie bezpieczeństwa danych uczniów	Ochrona danych	Wysoki	Naruszenie FERPA, naruszenia prywatności	Pilny



Niewystarczająca kontrola jakości treści kursów wspomaganych sztuczną inteligencją	Jakość akademicka	Średni	Spadek standardów nauczania	45 dni
--	-------------------	--------	-----------------------------	--------

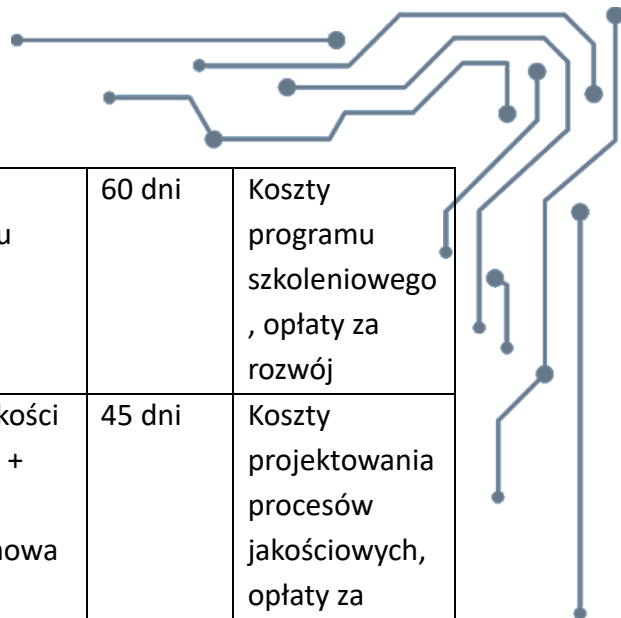
### Analiza i ograniczanie ryzyka procesu edukacyjnego wspomagane go sztuczną inteligencją

Najpoważniejsze zagrożenia w edukacji dotyczą uczciwości akademickiej, równości w dostępie do edukacji oraz bezpieczeństwa danych uczniów. Obszary te bezpośrednio wpływają na prawa uczniów i jakość edukacji.

Tabela 8. Plan łagodzenia skutków – procesy edukacyjne

Wydanie	Poziom docelowy	Strategia interwencji	Odpowiedzialny	Termin ostateczny	Wymagania dotyczące zasobów
Zasady korzystania ze sztucznej inteligencji przez uczniów	Jasne wytyczne + $\geq$ 90% świadomości	Opracowanie Polityki uczciwości akademickiej w zakresie sztucznej inteligencji, obowiązkowe szkolenie studentów	Sprawy akademickie + Sprawy studenckie	30 dni	Koszty rozwoju polityki, opłaty za szkolenia
Nierówności w dostępie do narzędzi AI	100% równy dostęp	Licencje na narzędzia instytucjonalnej sztucznej inteligencji, program wsparcia dla studentów	IT + Sprawy finansowe + Sprawy studenckie	25 dni	Opłaty licencyjne za oprogramowanie, koszty licencjonowania
Bezpieczeństwo danych uczniów	Pełna zgodność z FERPA	Wzmocnienie infrastruktury ochrony danych, obowiązkowy program szkoleniowy	Urząd Ochrony Danych + IT + Edukacja	20 dni	Koszty infrastruktury bezpieczeństwa, opłaty systemowe





Szkolenie z etyki AI na wydziale	≥ 90% uczestnictwa i certyfikacji	Kompleksowy program rozwoju kadry	Rozwój Wydziału	60 dni	Koszty programu szkoleniowego, opłaty za rozwój
Kontrola jakości treści kursów AI	Standaryzowany protokół jakości	System recenzji eksperckich, ustrukturyzowana lista kontrolna jakości	Biuro Jakości Edukacji + Komisja Programowa	45 dni	Koszty projektowania procesów jakościowych, opłaty za wdrożenie

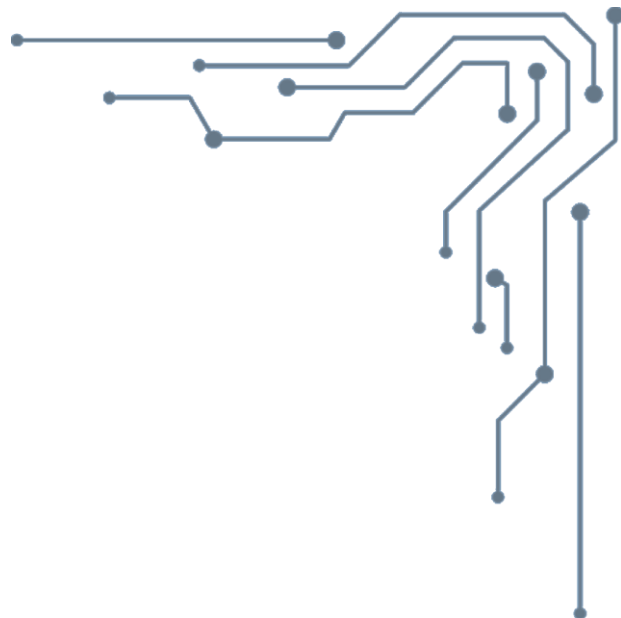
### 5.2.3 Audyt procesów administracyjnych wspomaganych przez sztuczną inteligencję

W tej sekcji analizowane jest wdrażanie i nadzór nad systemami sztucznej inteligencji (AI) w instytucjonalnych operacjach administracyjnych. Audyt ocenia ramy zarządzania wykorzystaniem AI w administracji, zgodność z przepisami o ochronie danych i prywatności oraz skuteczność programów szkoleniowych dla personelu w zakresie etycznego wdrażania AI. Kluczowe obszary oceny obejmują instytucjonalne polityki AI w zakresie funkcji administracyjnych, procedury zarządzania ryzykiem oraz standardy dostępności. Ponadto, analizowane są mechanizmy ciągłego doskonalenia, oparte na opiniach personelu. Niniejsza ocena zapewnia, że AI zwiększa efektywność administracyjną, jednocześnie zachowując przejrzystość, rozliczalność i jakość usług dla wszystkich interesariuszy instytucji. Poniższe rekomendacje mogą służyć jako punkt odniesienia do określania częstotliwości audytów i składu zespołu audytującego procesy administracyjne wspomagane przez AI.

#### Zalecane okresy audytu

- Systemy sztucznej inteligencji do zarządzania zasobami ludzkimi i płacami: co 6 miesięcy;
- Krytyczne systemy decyzyjne dla studentów (przyjęcia, pomoc finansowa, ocena akademicka): co 6 miesięcy;
- Ogólne usługi studenckie (systemy informacyjne, planowanie, wsparcie): Rocznie
- Ogólna automatyzacja administracyjna: co 2 lata;
- Audyty inicjowane: w przypadku skarg pracowników, awarii operacyjnych lub naruszeń przepisów;





### Zalecany skład zespołu audytowego:

- Dyrektor ds. administracyjnych (kierownik);
- Menedżer ds. Zasobów Ludzkich;
- Specjalista ds. usług finansowych;
- Menedżer systemów informatycznych;
- Przedstawiciel ds. Prawnych;
- Specjalista ds. audytu wewnętrznego;
- Przedstawiciel personelu administracyjnego;
- Zewnętrzny audytor zgodności (w razie potrzeby);

### LISTA KONTROLNA DO AUDYTU ADMINISTRACJI WSPOMAGANEJ SZTUCZNĄ INTELIGENCJĄ

Ta kompleksowa lista kontrolna audytu składa się z trzech kolejnych sekcji, które należy wypełnić zgodnie z podanymi wytycznymi:

- **Informacje o audycie:** Udokumentuj szczegóły instytucjonalne, skład zespołu audytorskiego, zakres i grupy docelowe.
- **Lista kontrolna audytu administracji wspomaganej sztuczną inteligencją:** systematyczna ocena integracji sztucznej inteligencji w sześciu kluczowych obszarach (odpowiedzialność, stronniczość i uczciwość, autonomia i sprawczość człowieka, ochrona prywatności i danych, bezpieczeństwo, inkluzywność oraz przejrzystość i wyjaśnialność).
- **Wyniki oceny administracji wspomaganej sztuczną inteligencją i priorytetyzacja analizy ryzyka:** ocena zidentyfikowanych ryzyk i opracowanie strategii działań naprawczych, skupiających się na zgodności zasobów ludzkich, bezpieczeństwie systemu finansowego i wydajności operacyjnej.

**Instrukcja:** Uzupełnij każdą sekcję w podanej kolejności, upewniając się, że wszystkie odpowiednie pola są wypełnione, a właściwe pola wyboru są zaznaczone, zanim przejdziesz do następnej sekcji.



Co-funded by  
the European Union

## 1. INFORMACJE O AUDYCIE

Instytucja/Organizacja: \_\_\_\_\_

Audyt przeprowadzony przez:

- Instytucja wewnętrzna/Dział: \_\_\_\_\_
- Instytucja/Dział zewnętrzny: \_\_\_\_\_

Lider zespołu audytorskiego: \_\_\_\_\_

Stanowisko/tytuł: \_\_\_\_\_

Członkowie zespołu audytorskiego (jeśli dotyczy): \_\_\_\_\_

Okres audytu: Od \_\_\_\_\_ do \_\_\_\_\_

Data zakończenia audytu: \_\_\_\_\_

Grupy docelowe/objęte działą: \_\_\_\_\_

---

---

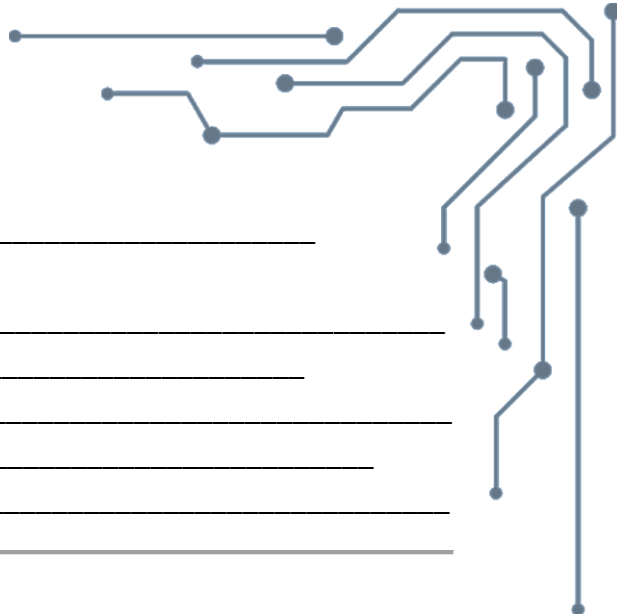
---

Przed wypełnieniem tej listy kontrolnej instytucje powinny najpierw odpowiedzieć na następujące pytanie wstępne:

**Pytanie wstępne:** Czy w Twojej instytucji funkcjonują obecnie zintegrowane systemy sztucznej inteligencji w procesach administracyjnych, z ustalonymi zasadami lub procedurami?

Tak  Nie

Jeśli „Nie”, ta lista kontrolna nie ma zastosowania. Jeśli „Tak”, przejdź do pełnej oceny.

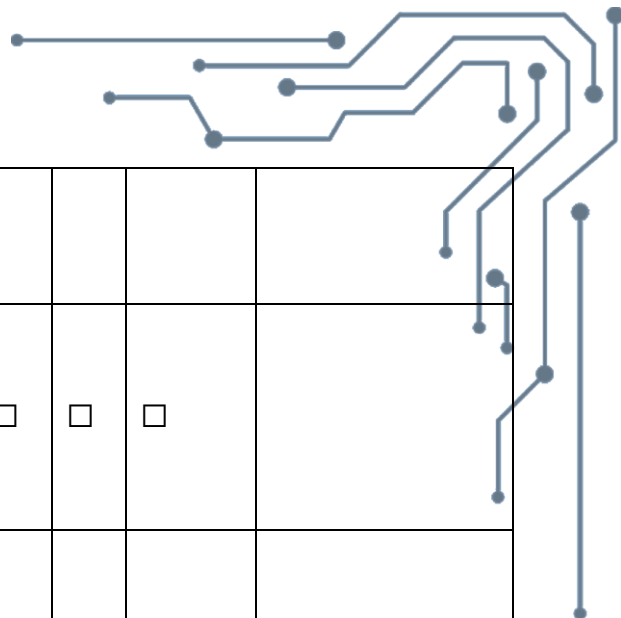


Co-funded by  
the European Union

## 2. LISTA KONTROLNA AUDYTU ADMINISTRACJI WSPOMAGANEJ SZTUCZNĄ INTELIGENCJĄ

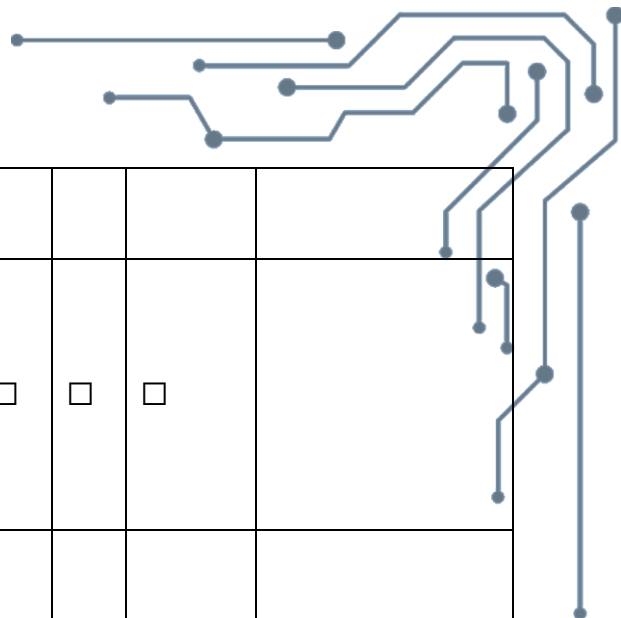
Tabela 9. Lista kontrolna audytu administracji wspomaganej przez sztuczną inteligencję

Zakres		Powiązany standard ESG.	Tak	Nie	Nie dotyczy	Notatki
Odpowiedzialność	Czy istnieją strategie i polityki instytucjonalne mające na celu integrację narzędzi sztucznej inteligencji z procesami administracyjnymi?	1.1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy w procesach administracyjnych wspomaganych przez sztuczną inteligencję wdrożono środki rozliczalności uwzględniające opinie użytkowników?	1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy role i obowiązki w zakresie monitorowania i zarządzania zadaniami administracyjnymi wspomaganymi przez sztuczną inteligencję są jasno określone?	1.1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy są dostępne programy szkoleniowe wspierające etyczne wykorzystanie sztucznej inteligencji w procesach administracyjnych?	1,5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Stronniczość i uczciwość	Czy procesy rekrutacji i zatrudniania wspomagane sztuczną inteligencją są regularnie kontrolowane pod kątem potencjalnych uprzedzeń związanych z cechami demograficznymi (np. płcią, wiekiem, pochodzeniem etnicznym, niepełnosprawnością)?	1.4, 1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy istnieje procedura przeglądu decyzji administracyjnych wspomaganych przez sztuczną	1.4, 1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	



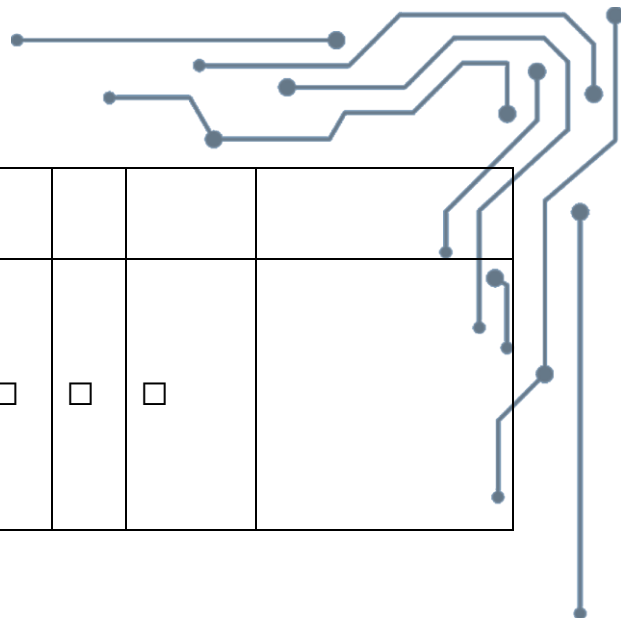
	inteligencję w celu zapewnienia, że nie dyskryminują one systematycznie żadnej grupy?					
<b>Autonomia i sprawczość człowieka</b>	Czy personel administracyjny dysponuje mechanizmem zgłaszania problemów napotkanych podczas korzystania z narzędzi AI?	1.7, 1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy systematycznie zbierane są opinie od personelu administracyjnego dotyczące korzystania z narzędzi sztucznej inteligencji?	1.7, 1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy pracownicy są uprawnieni i upoważnieni do odrzucania sugestii lub decyzji generowanych przez sztuczną inteligencję w procesach administracyjnych, gdy zachodzi taka potrzeba?	1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Prywatność i ochrona danych</b>	Czy wprowadzono zasady bezpieczeństwa danych i prywatności w odniesieniu do stosowania narzędzi sztucznej inteligencji w procesach administracyjnych?	1.1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy wdrożono procedury bezpiecznego usuwania lub anonimizowania danych administracyjnych przetwarzanych przez narzędzia sztucznej inteligencji po ich zamierzonym wykorzystaniu?	1.1, 1.7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Bezpieczeństwo i ochrona</b>	Czy personel administracyjny jest regularnie informowany o swoich obowiązkach związanych z ochroną poufnych danych podczas korzystania z narzędzi sztucznej inteligencji?	1,5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy potencjalne ryzyko związane z wykorzystaniem sztucznej	1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





	inteligencji jest regularnie oceniane?					
	Czy określono środki awaryjne mające na celu utrzymanie kluczowych funkcji administracyjnych na wypadek awarii systemu AI lub naruszenia bezpieczeństwa?	1.1, 1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy istnieje regularny proces przeglądu mający na celu zapewnienie ochrony narzędzi sztucznej inteligencji przed nowymi zagrożeniami bezpieczeństwa?	1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Inkluzywność</b>	Czy istnieją polityki instytucjonalne mające na celu zapewnienie, że narzędzia sztucznej inteligencji wykorzystywane w procesach administracyjnych spełniają zasady dostępności i inkluzywności?	1.1, 1.6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy narzędzia administracyjne oparte na sztucznej inteligencji zostały przetestowane w celu potwierdzenia ich przydatności dla osób o różnym poziomie kompetencji cyfrowych?	1.6, 1.9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Przejrzystość i wyjaśnialność</b>	Czy pracownicy administracyjni są informowani o tym, jakie narzędzia sztucznej inteligencji są wykorzystywane w ich procesach pracy i jak narzędzia te wpływają na ich codzienne zadania?	1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	Czy narzędziom administracyjnym opartym na sztucznej inteligencji towarzyszą jasne, nietechniczne wyjaśnienia dotyczące ich	1.8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	





działania i sposobu interpretacji wyników?					
Czy ograniczenia i potencjalne błędy narzędzi AI w procesach administracyjnych są jasno udokumentowane i zakomunikowane odpowiednim pracownikom?	1,5, 1,8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Wymagany plan analizy i łagodzenia ryzyka:  Tak  Nie

Jeśli tak, szczegóły analizy ryzyka i planu łagodzenia ryzyka:

- Odpowiedzialny za opracowanie planu: \_\_\_\_\_
- Docelowa data zakończenia działań łagodzących: \_\_\_\_\_
- Poziom priorytetu ogólnego:  Wysoki  Średni  Niski

Data przeglądu kontrolnego: \_\_\_\_\_

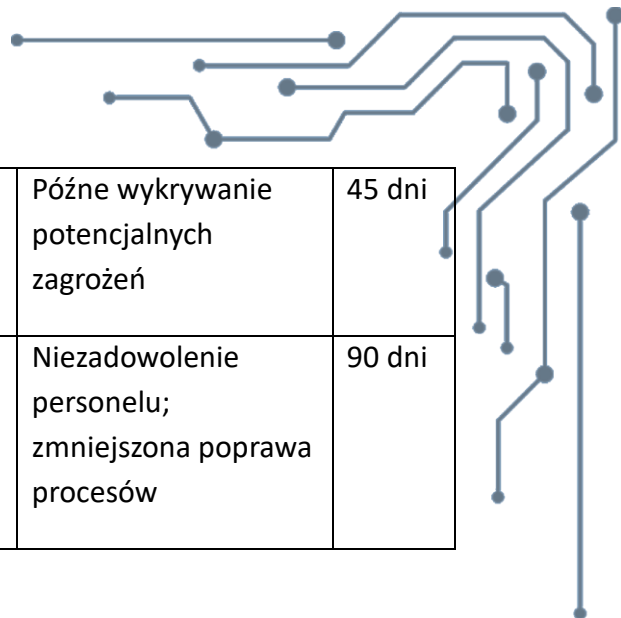
### 3. WYNIKI OCENY PROCESU ADMINISTRACYJNEGO WSPOMAGANE PRZEZ SZTUCZNĄ INTELIGENCJĘ I PRIORYTYZACJA RYZYKA

Najpoważniejsze zagrożenia w procesach administracyjnych wiążą się z uprzedzeniami w systemach kadrowych oraz niedociągnięciami w bezpieczeństwie systemu finansowego. Kwestie te wymagają natychmiastowej interwencji, ponieważ stwarzają znaczne ryzyko sankcji prawnych i odpowiedzialności instytucjonalnej.

Tabela 10. Priorytyzacja ryzyka – procesy administracyjne

Zidentyfikowany problem	Obszar	Poziom ryzyka	Racjonalne uzasadnienie	Pilność
Wykryto błędy w procesach decyzyjnych HR	Zasoby ludzkie	Wysoki	Potencjalne naruszenie prawa pracy; ryzyko dyskryminacji	Pilny
Niedociągnięcia w zakresie bezpieczeństwa danych w systemach finansowych	Sprawy finansowe	Wysoki	Ryzyko strat finansowych i niezgodności z przepisami	Pilny
Niewystarczające przeszkolenie personelu w zakresie sztucznej inteligencji	Rozwój potencjału	Średni	Spadek wydajności operacyjnej	60 dni





Nie przeprowadzono oceny ryzyka systemu AI	Zarządzanie ryzykiem	Średni	Późne wykrywanie potencjalnych zagrożeń	45 dni
Brak mechanizmu składania skarg przez pracowników	Odpowiedzialność	Niski	Niezadowolenie personelu; zmniejszona poprawa procesów	90 dni

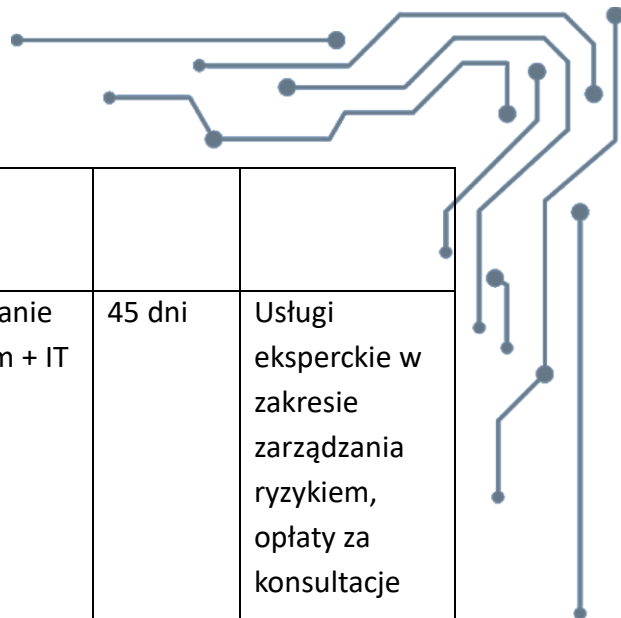
### Analiza i ograniczanie ryzyka procesów administracyjnych wspomaganie sztuczną inteligencją

Najpoważniejsze zagrożenia w procesach administracyjnych wiążą się z uprzedzeniami w systemach kadrowych oraz niedociągnięciami w bezpieczeństwie systemu finansowego. Kwestie te wymagają natychmiastowej interwencji, ponieważ stwarzają znaczne ryzyko sankcji prawnych i odpowiedzialności instytucjonalnej.

Tabela 11. Plan łagodzenia skutków – procesy administracyjne

Wydanie	Poziom docelowy	Strategia interwencji	Odpowiedzia Iny	Termin ostateczny	Wymagania dotyczące zasobów
Błąd systemu HR	≥ 95% wskaźnik dobrych wyników	Wdrażanie aplikacji do wykrywania stronniczości; przegląd algorytmów	HR + IT + Dział prawny	20 dni	Narzędzia do analizy uprzedzeń, koszty testowania
Bezpieczeństwo w systemie finansowego	Bezpieczeństwo na poziomie bankowym	Zaawansowane szyfrowanie; uwierzytelnianie wieloskładnikowe	Sprawy finansowe + Bezpieczeństwo w IT	25 dni	Koszty infrastruktury bezpieczeństwa, opłaty systemowe
Szkolenie personelu AI	≥ 90% wskaźnika przeszkolenia personelu	Program nauczania w zakresie sztucznej inteligencji w	Szkolenia HR + IT	60 dni	Materiały szkoleniowe, koszty programu





		departamencie			
Ocena ryzyka AI	Protokół systematycznego ryzyka	Ustanowienie ram oceny ryzyka	Zarządzanie ryzykiem + IT	45 dni	Usługi eksperckie w zakresie zarządzania ryzykiem, opłaty za konsultacje
System skarg pracowników	Skuteczny kanał informacji zwrotnej	Wdrożenie platformy do przekazywania opinii cyfrowych	Wsparcie HR + IT	90 dni	Subskrypcja platformy opinii, koszty oprogramowania

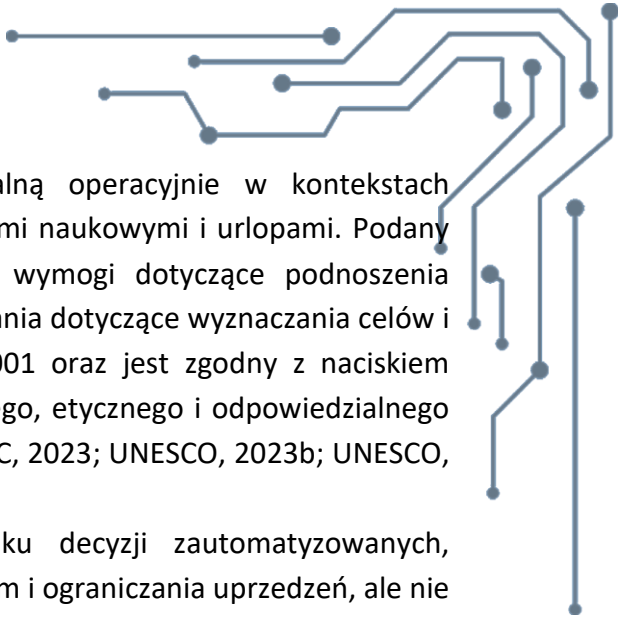
#### 5.2.4. Uzasadnienie wartości odniesienia w wytycznych

W niniejszych wytycznych zarówno wartości procentowe, jak i nieprocentowe przyjęto jako wartości odniesienia ustalane przez instytucję – zdefiniowane w polityce, podlegające audytowi cele, które przekładają zasady zarządzania na mierzalne cele dla instytucji szkolnictwa wyższego. Przywołane ramy kładą nacisk na wyznaczanie celów, świadomość, zarządzanie ryzykiem, monitorowanie i ciągłe doskonalenie, jednak zazwyczaj nie określają progów liczbowych. W związku z tym wartości podane w niniejszych wytycznych służą jako instytucjonalne punkty odniesienia, które przekładają te wymagania na przejrzystą i sprawdzoną praktykę (Międzynarodowa Organizacja Normalizacyjna i Międzynarodowa Komisja Elektrotechniczna [ISO/IEC], 2022; ISO/IEC, 2023; Narodowy Instytut Norm i Technologii [NIST], 2023; UNESCO, 2023b; UNESCO, 2023c).

Dostępność i inkluzywność są traktowane jako kwestie binarnej zgodności z prawem. Usługi internetowe i mobilne sektora publicznego muszą być dostępne na mocy dyrektywy 2016/2102 Parlamentu Europejskiego i Rady, wdrożonej za pośrednictwem zharmonizowanej normy EN 301 549 Europejskiego Instytutu Norm Telekomunikacyjnych (ETSI). Ponieważ zgodności z prawem nie można interpretować w sposób ciągły, cele dotyczące dostępu cyfrowego ustalono na poziomie 100% równego dostępu, co oznacza, że częściowa zgodność nie może stanowić akceptowalnego stanu równowagi (Parlament Europejski i Rada, 2016; ETSI, 2021).

W zakresie szkoleń, podnoszenia świadomości i certyfikacji, ramy zarządzania wymagają od instytucji określenia celów, zapewnienia interesariuszom świadomości ról i obowiązków oraz wdrożenia mechanizmów ciągłego doskonalenia, ale nie określają wskaźników uczestnictwa. W związku z tym przyjęto punkt odniesienia w postaci co najmniej 90% pokrycia dla studentów, pracowników, wykładowców i badaczy. Wartość ta przybliża „niemal





powszechny” zasięg, pozostając jednocześnie wykonalną operacyjnie w kontekstach akademickich charakteryzujących się rotacją kadr, urlopami naukowymi i urlopami. Podany procent stanowi wybór instytucjonalny, który wdraża wymogi dotyczące podnoszenia świadomości określone w normie ISO/IEC 27001, oczekiwania dotyczące wyznaczania celów i ciągłego doskonalenia określone w normie ISO/IEC 42001 oraz jest zgodny z naciskiem UNESCO na budowanie potencjału w zakresie bezpiecznego, etycznego i odpowiedzialnego wykorzystania sztucznej inteligencji (ISO/IEC, 2022; ISO/IEC, 2023; UNESCO, 2023b; UNESCO, 2023c).

Aby zapewnić sprawiedliwe wyniki w przypadku decyzji zautomatyzowanych, instrumenty autorytatywne wymagają zarządzania ryzykiem i ograniczania uprzedzeń, ale nie wyznaczają jednego progu sprawiedliwego wyniku. W związku z tym niniejsza wytyczna przyjmuje bardziej rygorystyczny instytucjonalny punkt odniesienia, wynoszący co najmniej 95% wskaźnik sprawiedliwego wyniku, aby zminimalizować zróżnicowane skutki i zmniejszyć zmienność pomiarów w kontekstach takich jak rekrutacja pracowników czy rekrutacja studentów. „Zasada czterech piątych” (80%) ustanowiona przez Jednolite Wytyczne Komisji ds. Równych Szans Zatrudnienia (EEOC) Stanów Zjednoczonych jest uznawana za heurystykę służącą do identyfikacji potencjalnie negatywnych skutków, ale nie jest traktowana jako wiążący próg w tym kontekście (Parlament Europejski i Rada, 2024; EEOC, 1978).

W przypadku wydajności technicznej w modelach wspomagania decyzji, ramy cyklu życia wymagają zdefiniowania i zarządzania metrykami adekwatnymi do celu, nie ustanawiają jednak uniwersalnych progów. Aby zapewnić porównywalność i przejrzystość eskalacji w systemach heterogenicznych, przyjęto punkt odniesienia na poziomie co najmniej 80% dla dokładności, precyzji i wykrycia. W przypadku aplikacji wysokiego ryzyka można ustalić bardziej rygorystyczne progi. Wartość ta jest wyraźnie określona jako instytucjonalny punkt odniesienia, mający na celu promowanie przejrzystości zarządzania i spójnego monitorowania (NIST, 2023; ISO/IEC, 2023).

Jakość danych i monitorowanie dryfu są przedstawiane jako kontrole diagnostyczne, a nie jako zewnętrznie narzucone progi. Wskaźniki takie jak 15% braków danych lub 25% wzrost liczby wartości odstających są interpretowane jako wartości obserwowane, a nie jako limity normatywne. Ponieważ akceptowalność braków danych zależy od mechanizmu leżącego u ich podstaw – takiego jak braki całkowicie losowe, braki losowe lub braki nielosowe – oraz od zastosowanego rozwiązania analitycznego, w literaturze statystycznej nie zidentyfikowano uniwersalnego „akceptowalnego odsetka”. Powszechnie cytowane dążenie do braku danych poniżej 5% jest uznawane za punkt odniesienia w zapewnianiu jakości, a nie za ścisłą regułę. Przesunięcie rozkładu jest monitorowane za pomocą Wskaźnika Stabilności Populacji (PSI), gdzie heurystyki oparte na praktyce zazwyczaj interpretują wartości poniżej 0,10 jako niewielką zmianę, a wartości 0,20–0,25 lub powyżej jako znaczny dryf uzasadniający zbadanie lub odświeżenie modelu. Uważa się, że te punkty odcięcia stanowią raczej heurystykę opartą na praktyce niż wymogi prawne (Little i Rubin, 2019; du Pisane, Allison i Visagie, 2023; Yurdakul i Naranjo, 2020).



W przypadku, gdy obserwowana wydajność spada poniżej wartości odniesienia (np. 65% dokładności bazowej w porównaniu z progami wdrożenia  $\geq 80\%$ ), wdrażane są plany poprawy z ograniczonym czasem realizacji (np. 12–24 miesiące). Gwarantuje to, że progi będą pełniły funkcję progresywnych punktów odniesienia, a nie statycznych kryteriów zaliczenia/zaliczenia. Podejście to jest zgodne z oczekiwaniami systemu zarządzania w zakresie wyznaczania celów, wykazywania ciągłego doskonalenia oraz dokumentowania decyzji zarządczych i postępów (ISO/IEC, 2023).

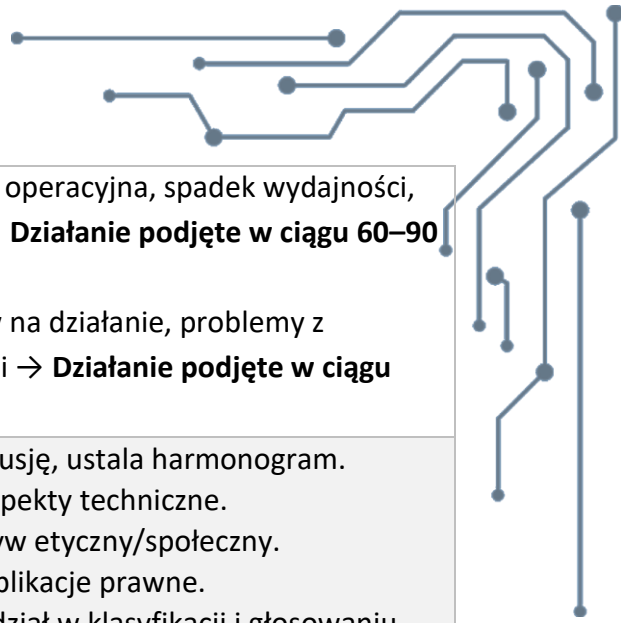
## 6. Ocena ogólnego procesu audytu

W tej sekcji przedstawiono ustrukturyzowane ramy oceny audytów AI w szkolnictwie wyższym, określania profili ryzyka instytucjonalnego i formułowania strategii doskonalenia. Ustalenia z czterech obszarów audytu – systemów/narzędzi AI, badań naukowych, edukacji i administracji – zostały skonsolidowane w celu oceny dojrzałości zarządzania AI. Metodologia wykorzystuje wielowymiarową analizę wpływu i prawdopodobieństwa ryzyka, wspartą wyraźnymi kryteriami priorytetyzacji. Każdy obszar jest oceniany oddzielnie, a wyniki są integrowane z profilem instytucjonalnym, który uwzględnia plany działania i mechanizmy monitorowania ciągłego doskonalenia. Poniższa tabela podsumowuje podstawowe elementy metodologii oceny ryzyka, szczegółowo opisując jej cel, podejście, etapy procesu, poziomy klasyfikacji ryzyka oraz integrację z instytucjonalnymi planami doskonalenia. Stanowi ona praktyczne odniesienie dla zespołów audytorskich, zapewniając spójność, przejrzystość i wykonalne rezultaty we wszystkich obszarach audytu AI.

Tabela 12. Składniki oceny ryzyka

Część	Bliższe dane
Zamiar	Zapewnienie spójnej, prowadzonej przez ekspertów oceny ryzyka we wszystkich obszarach audytu sztucznej inteligencji, wspierającej podejmowanie decyzji instytucjonalnych.
Zbliżyć się	Standardowa metodologia wykorzystująca konsensus ekspertów, a nie czysto matematyczną punktację.
Proces	<ol style="list-style-type: none"> <li><b>Identyfikacja</b> – wyszczególniono elementy z list kontrolnych (sekcje 5.2.1, 5.2.2, 5.3, 5.4, 5.5), które nie zostały zrealizowane.</li> <li><b>Przegląd interdyscyplinarny</b> – eksperci techniczni, etyczni, prawni i operacyjni oceniają kwestie ze swojej perspektywy.</li> <li><b>Spotkanie konsensusowe</b> – zespół omawia różnice i uzgadnia klasyfikację ryzyka.</li> <li><b>Rejestrowanie decyzji</b> – Głosowanie większością głosów; udokumentowane uzasadnienie.</li> </ol>
Poziomy ryzyka i wymagany czas działania	<p><b>Wysoki priorytet</b> – bezpośrednie zagrożenia dla reputacji, zgodności, bezpieczeństwa, integralności akademickiej lub finansów</p> <p>→ <b>Podjęcie działań w ciągu 30 dni.</b></p>





	<p><b>Średni priorytet</b> – nieefektywność operacyjna, spadek wydajności, drobne problemy ze zgodnością → <b>Działanie podjęte w ciągu 60–90 dni.</b></p> <p><b>Niski priorytet</b> – minimalny wpływ na działanie, problemy z użytecznością, luki w dokumentacji → <b>Działanie podjęte w ciągu ponad 90 dni.</b></p>
<b>Role decyzyjne</b>	<ul style="list-style-type: none"><li>• <b>Lider zespołu</b> – moderuje dyskusję, ustala harmonogram.</li><li>• <b>Ekspert techniczny</b> – ocenia aspekty techniczne.</li><li>• <b>Ekspert ds. etyki</b> – ocenia wpływ etyczny/społeczny.</li><li>• <b>Ekspert prawny</b> – analizuje implikacje prawne.</li><li>• <b>Wszyscy członkowie</b> – biorą udział w klasyfikacji i głosowaniu.</li></ul>
<b>Dokumentacja</b>	Wszystkie klasyfikacje zostały zarejestrowane wraz z uzasadnieniem, odpowiedzialnymi stronami, realistycznymi harmonogramami i zapotrzebowaniem na zasoby.
<b>Integracja z planem łagodzenia skutków</b>	Klasyfikacje ryzyka bezpośrednio wpływają na działania korygujące, alokację zasobów i harmonogramy monitorowania. Dzięki temu w pierwszej kolejności rozpatrywane są ryzyka o najwyższym priorytecie, przy jednoczesnym zachowaniu przejrzystości i rozliczalności.

Ta struktura oceny ryzyka zapewnia ustrukturyzowany, opracowany przez ekspertów proces oceny niezaliczonych audytów we wszystkich obszarach zarządzania sztuczną inteligencją. Łącząc wielodyscyplinarny przegląd, priorytetyzację opartą na konsensusie oraz jasne przypisanie ról, zapewnia spójną klasyfikację ryzyk na poziomy o wysokim, średnim i niskim priorytecie. Poza funkcją techniczną, metodologia stanowi fundament zaufania i ładu instytucjonalnego, chroniąc integralność akademicką, zapewniając zgodność z normami prawnymi i etycznymi oraz wzmacniając zaufanie interesariuszy. Jej przejrzyste i oparte na rozliczalności podejście umożliwia instytucjom szkolnictwa wyższego skuteczne reagowanie na bezpośrednie ryzyka, jednocześnie wzmacniając długoterminową zdolność adaptacji i stale podnosząc dojrzałość w zakresie zarządzania sztuczną inteligencją.



## 7. Ogólny raport z audytu

Szablon ogólnego raportu z audytu jest sporządzany na podstawie wkładu wszystkich zespołów audytorskich. W ten sposób powstaje wspólny raport podsumowujący, integrujący informacje zwrotne każdego zespołu dotyczące jego obszaru audytu. Ta skonsolidowana struktura ułatwia decydującym monitorowanie i zapewnia obiektywną podstawę do podejmowania świadomych decyzji instytucjonalnych. Poniższy szablon raportu z audytu podkreśla mocne strony instytucji, obszary wymagające poprawy oraz kwestie krytyczne, a także zawiera oceny końcowe i zalecenia. Uzupełnione listy kontrolne audytu, tabele priorytetyzacji ryzyka oraz plany łagodzenia ryzyka są dołączone do tego ogólnego raportu z audytu, który jest następnie przekazywany najwyższemu organowi zarządzającemu instytucji (np. rektorowi, prorektorowi lub przewodniczącemu komisji ds. jakości).

### Szablon raportu audytu

W ramach ustaleń uzyskanych podczas audytu, przedstawiono oceny dla każdego obszaru audytu, identyfikując mocne strony, obszary wymagające poprawy oraz kwestie krytyczne. Ustalenia te podsumowano w poniższej tabeli.

Tabela 13. Wnioski oparte na wynikach audytu

Obszar audytu	Mocne strony	Obszary do poprawy	Krytyczne problemy
Audyt etyczny/prawny sztucznej inteligencji			
Audyt procesów badawczych wspomaganych sztuczną inteligencją			
Audyt procesów edukacyjnych wspomaganych sztuczną inteligencją			
Audyt procesów administracyjnych wspomaganych sztuczną inteligencją			
Audyt techniczny systemów i narzędzi AI			
Audyt etyczny/prawny systemów i narzędzi AI			



## Rekomendacje i planowanie działań

W oparciu o kompleksowe ustalenia przedstawione powyżej, poniższa tabela przedstawia przełożenie zidentyfikowanych obszarów wymagających poprawy i kluczowych problemów na konkretne, możliwe do wdrożenia strategie. Rekomendacje te zostały uszeregowane pod względem priorytetów na podstawie wyników oceny ryzyka i mają na celu wzmocnienie instytucjonalnego zarządzania AI, przy jednoczesnym rozwiązaniu bieżących problemów związanych z zgodnością z przepisami i kwestiami operacyjnymi.

Tabela 14. Rekomendacje dotyczące usprawnień instytucjonalnych i kwestii krytycznych

Obszar audytu	Rekomendacje dotyczące obszarów wymagających poprawy	Zalecenia dotyczące kluczowych problemów
Audyt etyczny/prawny sztucznej inteligencji		
Audyt procesów badawczych wspomaganych sztuczną inteligencją		
Audyt procesów edukacyjnych wspomaganych sztuczną inteligencją		
Audyt procesów administracyjnych wspomagany sztuczną inteligencją		
Audyt techniczny systemów i narzędzi AI		
Audyt etyczny/prawny systemów i narzędzi AI		

Ocena końcowa:

---

---

---

---

Dodatkowe uwagi

---

---

---

---



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them.  
Grant no:2024-1-LT01-KA220-HED-000251565

# ZAŁĄCZNIK

## Audyt techniczny systemów/narzędzi AI

W tej sekcji przedstawiono dwa uzupełniające się podejścia do kompleksowej oceny systemów i narzędzi sztucznej inteligencji (AI) wykorzystywanych w instytucjach szkolnictwa wyższego. Po pierwsze, systemy są oceniane na podstawie mierzalnych kryteriów, takich jak wydajność techniczna, wskaźniki dokładności, wskaźniki szybkości i wydajność operacyjna. Po drugie, ocena etyczna jest przeprowadzana w ramach zasad przejrzystości algorytmicznej, uczciwości, inkluzywności i rozliczalności. To dwuwymiarowe podejście ma na celu zapewnienie, że systemy AI działają zarówno technicznie niezawodnie, jak i etycznie odpowiedzialnie. Poniższe zalecenia mogą służyć jako punkt odniesienia przy określaniu częstotliwości audytów i składu zespołów ds. systemów i narzędzi AI.

### Zalecane okresy audytu

- **Systemy wysokiego ryzyka:** Roczne (algorytmy przyjęć, systemy oceniania, decyzje dotyczące pomocy finansowej);
- **Systemy o średnim ryzyku:** co 2 lata (analityka uczenia się, chatboty, systemy rekomendacji);
- **Systemy niskiego ryzyka:** co 3 lata z pełnym audytem (narzędzia informacyjne ogólne, podstawowa automatyzacja);
- **Audyty inicjowane:** w przypadku aktualizacji systemu, problemów z wydajnością, incydentów bezpieczeństwa lub zmian w przepisach;

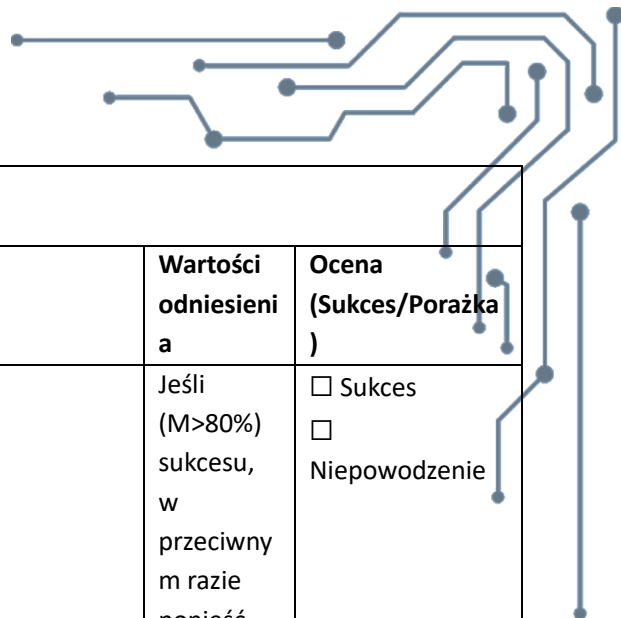
### Zaleceni członkowie zespołu audytowego

- Administrator systemów informatycznych (kierownik);
- Specjalista ds. bezpieczeństwa danych;
- Ekspert techniczny ds. sztucznej inteligencji i uczenia maszynowego;
- Specjalista ds. zapewnienia jakości;
- Doradca ds. zgodności z prawem;
- Przedstawiciel Komisji Etycznej;
- Przedstawiciel użytkownika końcowego (wykładowcy/pracownicy/studenci);

## Ocena techniczna systemów/narzędzi AI

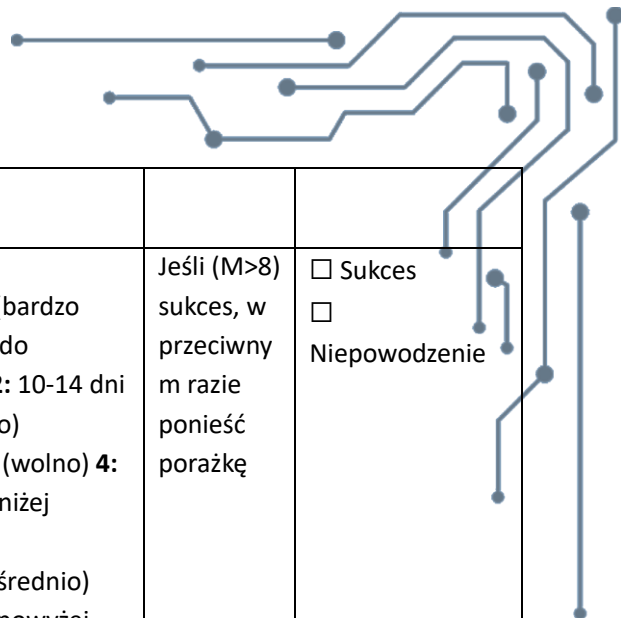
Ta lista kontrolna pozwala ocenić podstawowe wskaźniki wydajności technicznej systemów AI, aby upewnić się, że spełniają one standardy instytucjonalne dotyczące dokładności, wydajności i wymagań operacyjnych.





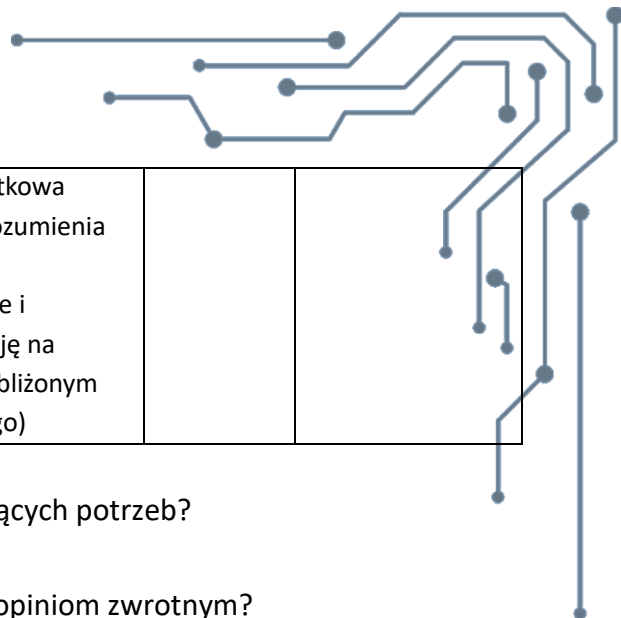
LISTA KONTROLNA WSKAŹNIKÓW WYDAJNOŚCI				
Metryczny	Opis	Miara (M)	Wartości odniesienia	Ocena (Sukces/Porażka)
Dokładność	Dokładność to stosunek prawidłowo przewidzianych przykładów do całkowitej liczby przykładów (jeśli masz 100 przykładów, a Twój model prawidłowo przewiduje 85 z nich, wówczas wartość dokładności wynosi 85%).	(1 do 100)	Jeśli (M>80%) sukces, w przeciwnym razie ponieść porażkę	<input type="checkbox"/> Sukces <input type="checkbox"/> Niepowodzenie
Niezawodność (precyzja)	Precyzja to stosunek liczby prawdziwie pozytywnych przykładów do liczby przykładów przewidzianych przez model jako pozytywne. Odpowiada na pytanie: „Jak dokładne są pozytywne przewidywania modelu?” (Jeśli model wygenerował „pozytywne” przewidywania dla 100 przykładów, a 80 z nich jest faktycznie pozytywnych, wówczas wartość precyzji wynosi 80%).	(1 do 100)	Jeśli (M>80%) sukces, w przeciwnym razie ponieść porażkę	<input type="checkbox"/> Sukces <input type="checkbox"/> Niepowodzenie
Przypomnienie sobie czegoś	Współczynnik rozpoznawalności (recall) to stosunek prawidłowo przewidzianych przykładów pozytywnych do wszystkich faktycznie pozytywnych. Odpowiada on na pytanie: jaki odsetek faktycznie pozytywnych wyników został wykryty przez model? (Jeśli zbiór danych zawiera 100 przykładów, które są faktycznie pozytywne, a model poprawnie zidentyfikuje 75 z nich jako pozytywne, wówczas współczynnik rozpoznawalności wynosi 75%).	(1 do 100)	Jeśli (M>80%) sukces, w przeciwnym razie ponieść porażkę	<input type="checkbox"/> Sukces <input type="checkbox"/> Niepowodzenie
<b>Efektywność</b>				
<ul style="list-style-type: none"> <li>Czas wnioskowania</li> </ul>	Czas wnioskowania to czas potrzebny wytrenowanemu modelowi na wygenerowanie prognozy dla nowego wejścia (przykład). Innymi słowy, jest to czas potrzebny modelowi na ukończenie procesu przewidywania (jeśli model przetwarzania języka naturalnego potrzebuje 0,2 sekundy na	<b>Pomyślnie, jeśli poniżej X ms</b>	Jeśli (M<200ms) sukces, w przeciwnym razie ponieść porażkę	<input type="checkbox"/> Sukces <input type="checkbox"/> Niepowodzenie





	analizę zdania, czas wnioskowania wynosi 200 ms).			
<ul style="list-style-type: none"> <li>• Czas szkolenia</li> </ul>	Czas potrzebny modelowi na nauczenie się znaczących wzorców, struktur językowych i relacji semantycznych z danych tekstowych	<b>od 1 do 10</b> <b>1:</b> 14+ dni (bardzo wolno, nie do przyjęcia) <b>2:</b> 10-14 dni (dość wolno) <b>3:</b> 7-10 dni (wolno) <b>4:</b> 5-7 dni (poniżej średniej) <b>5:</b> 3-5 dni (średnio) <b>6:</b> 2-3 dni (powyżej średniej) <b>7:</b> 1-2 dni (Dobry) <b>8:</b> 12-24 godzin (Bardzo dobry) <b>9:</b> 6-12 godzin (prawie idealnie) <b>10:</b> <6 godzin (Doskonały)	Jeśli (M>8) sukces, w przeciwnym razie ponieść porażkę	<input type="checkbox"/> Sukces <input type="checkbox"/> Niepowodzenie
<ul style="list-style-type: none"> <li>• Rozumienie języka naturalnego (NLU)</li> </ul>	NLU odnosi się do zdolności systemów sztucznej inteligencji do rozumienia, interpretowania i wyciągania znaczeń z języka ludzkiego.	<b>od 1 do 10</b> <b>1-2:</b> Bardzo słaba zdolność rozumienia (trudności ze zrozumieniem nawet podstawowych słów) <b>3-4:</b> Słaba zdolność rozumienia (rozumie proste konstrukcje zdań, ma problemy ze złożonymi wyrażeniami) <b>5-6:</b> Zdolność rozumienia na średnim poziomie (rozumie ogólne wyrażenia, nie dostrzega niektórych niuansów) <b>7-8:</b> Dobra zdolność rozumienia (poprawne rozumienie większości kontekstów, rozumienie większości złożonych wyrażen)	Jeśli (M>8) sukces, w przeciwnym razie ponieść porażkę	<input type="checkbox"/> Sukces <input type="checkbox"/> Niepowodzenie





		9-10: Wyjątkowa zdolność rozumienia (wykazuje zrozumienie i interpretację na poziomie zbliżonym do ludzkiego)		
--	--	---	--	--

**Skalowalność i konserwacja:**

Czy systemy/narzędzia AI potrafią dostosować się do rosnących potrzeb?

Tak  Nie

Czy systemy sztucznej inteligencji są udoskonalane dzięki opiniom zwrotnym?

Tak  Nie

Czy sztuczna inteligencja przestrzega zasad minimalizacji danych?

Tak  Nie

Czy wdrożono bezpieczne praktyki przechowywania danych?

Tak  Nie

Czy przeprowadzane są regularne testy naruszeń danych?

Tak  Nie

**Analiza ryzyka technicznego systemów/narzędzi AI**

Po zakończeniu oceny wskaźników wydajności technicznej, zidentyfikowane niedociągnięcia są systematycznie priorytetyzowane w oparciu o ich potencjalny wpływ na funkcjonowanie instytucji, wyniki studentów i niezawodność systemu. Ta analiza ryzyka przekłada luki w wydajności na priorytety działań naprawczych.

**Tabela 15. Priorytetyzacja ryzyka – wydajność techniczna**

Zidentyfikowany problem	System/Obszar	Poziom ryzyka	Racjonalne uzasadnienie	Pilność
Algorytm przyjmowania studentów: dokładność 65% (punkt odniesienia 80%)	Sprawy studenckie	Wysoki	Błędne decyzje o przyjęciu/odrzuconiu, odpowiedzialność prawna	Pilny
Wydajność chatbota NLU: ocena 4/10	Ogólny System Informacyjny	Średni	Niezadowolenie użytkowników, strata czasu	60 dni
Precyzja systemu oceny zadań $\geq 80\%$	Ocena akademicka	Wysoki	Niesprawiedliwe ocenianie, skargi dotyczące nauczania	Pilny



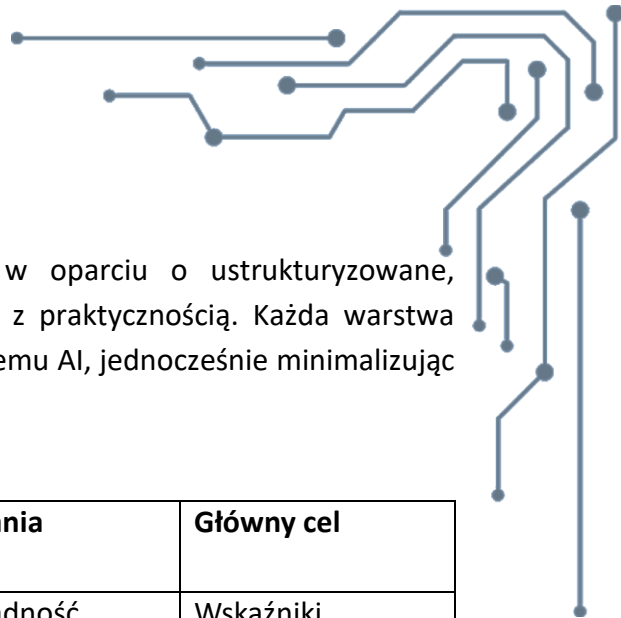
Czas wnioskowania systemu rekomendacji bibliotecznych: 250 ms	Usługi biblioteczne	Niski	Drobny problem z doświadczeniem użytkownika	90 dni
Zdolność NLU do zrozumienia systemu finansowego 6/10	Sprawy finansowe	Średni	Błędy transakcyjne, opóźnienia operacyjne	45 dni

### Analiza i ograniczanie ryzyka systemów/narzędzi AI

Jak wynika z tych przykładów, awarie techniczne w kluczowych procesach decyzyjnych, takich jak rekrutacja studentów i ocena akademicka, są klasyfikowane jako wysokiego ryzyka, ponieważ bezpośrednio wpływają na prawa jednostki i obowiązki instytucjonalne. Problemy w systemach wsparcia, takich jak ogólne chatboty informacyjne, można zaliczyć do kategorii średniego ryzyka.

Tabela 16. Plan łagodzenia skutków – wydajność techniczna

Wydanie	Poziom docelowy	Strategia interwencji	Odpowiedzialny	Termin ostateczny	Wymagania dotyczące zasobów
Niska dokładność algorytmu przyjęć studentów	≥ 80% dokładności	Ponowne trenowanie modelu, poprawa jakości danych	Dział IT + Zespół ds. nauki o danych	30 dni	Wsparcie naukowców zajmujących się danymi, koszty rozwoju
Błąd w systemie oceny zadań	≥ 80% precyzji, uczciwa punktacja	Przegląd algorytmów, rozszerzenie zestawu testowego	Jednostka IT + Ocena Akademicka	25 dni	Usługi konsultantów zewnętrznych, opłaty za konsultacje
Problem ze zrozumieniem chatbota	8/10+ wynik NLU	Aktualizacja bazy wiedzy, szkolenie modelu NLP	Zespół ds. wsparcia IT i treści	60 dni	Wsparcie ekspertów NLP, koszty szkoleń
Pojemność NLU systemu finansowego	Ocena zrozumienia 9/10+	Leksykon dziedzinowy, model specjalistyczny	Sprawy finansowe + IT	45 dni	Opłaty licencyjne za oprogramowanie, koszty licencjonowania
Prędkość rekomendacji bibliotecznych	Wnioskowanie <100 ms	Optymalizacja serwera, mechanizm pamięci podręcznej	Infrastruktura informatyczna	90 dni	Koszty modernizacji sprzętu, opłaty za sprzęt



## Mechanizm wykrywania dryfu modelu AI

Mechanizm wykrywania dryfu modelu działa w oparciu o ustrukturyzowane, trójwarstwowe podejście, które równoważy dokładność z praktycznością. Każda warstwa służy określonemu celowi, utrzymując niezawodność systemu AI, jednocześnie minimalizując nakłady ze strony personelu instytucji.

Warstwa	Częstotliwość	Czas trwania	Kluczowe działania	Główny cel
Warstwa 1	Cotygodniowa	5 minut	Monitoruj dokładność, skargi, czas przetwarzania	Wskaźniki wydajności
Warstwa 2	Miesięczny	30 minut	Kontrole jakości danych, obliczenia PSI	Integralność danych
Warstwa 3	Kwartalny	2 godziny	Pełna ocena, opinie interesariuszy	Ocena strategiczna

### Pierwsza warstwa: cotygodniowa ocena wyników

Cotygodniowa kontrola wydajności stanowi pierwszą linię obrony przed degradacją modelu. Podczas tej oceny instytucje monitorują trzy kluczowe wskaźniki: wskaźnik dokładności systemu, liczbę skarg użytkowników oraz średni czas przetwarzania. Spadek dokładności systemu poniżej 80% w stosunku do ustalonego poziomu bazowego wynoszącego 90% lub więcej wskazuje na znaczny spadek wydajności. Wzrost liczby skarg użytkowników o 50% sugeruje problemy techniczne lub zmieniające się potrzeby użytkowników, a podwojenie czasu przetwarzania często sygnalizuje, że model ma problemy z obsługą nowych typów danych.

### Druga warstwa: Miesięczna ocena integralności danych

Miesięczna kontrola stanu danych koncentruje się na zapewnieniu jakości i spójności danych. Ta ocena porównuje bieżące dane wejściowe z oryginalnymi danymi treningowymi w celu zidentyfikowania istotnych zmian. Podstawą tej warstwy jest Wskaźnik Stabilności Populacji (PSI) – gdy PSI przekracza 0,2, oznacza to znaczny dryft danych wymagający korekty modelu.

Monitorowanie stanu danych:

Brakujące wartości: [██████████] 15% (Norma: <5%)

Wykryto wartości odstające: [██████████] wzrost o 25%

Nowe kategorie: [██████████] Znalaziono 8 nowych typów

Wynik PSI: [██████████] 0,25 (Próg: 0,2) ⚠



Co-funded by  
the European Union

### Trzecia warstwa: kompleksowa kwartalna ocena

Kwartalny przegląd stanowi najbardziej dogłębną ocenę, obejmującą pełną analizę wydajności modelu, systematyczne zbieranie opinii od interesariuszy i strategiczne podejmowanie decyzji dotyczących aktualizacji modelu w oparciu o zgromadzone dowody.

### System klasyfikacji alertów oparty na ryzyku

Mechanizm wykrywania klasyfikuje stan systemu na trzy różne poziomy na podstawie powagi i zakresu zidentyfikowanych problemów.

Poziom statusu	Kryteria	Wymagane działania	Czas reakcji
<b>Stabilne operacje</b>	Wszystkie wskaźniki w zakresie	Kontynuuj normalne monitorowanie	Nie dotyczy
<b>Podwyższony poziom monitorowania</b>	1 metryka przekracza próg	Zwiększ częstotliwość kontroli do cotygodniowych i zbadaj	24-48 godzin
<b>Krytyczna interwencja</b>	Wiele błędów metryk	Natychmiastowe dochodzenie, ręczne obejście	0-24 godzin

W warunkach stabilnego funkcjonowania instytucje kontynuują normalną działalność, utrzymując jednocześnie regularne harmonogramy monitorowania. Wzmoczony monitoring ma miejsce, gdy jeden z wskaźników wydajności przekracza próg ostrzegawczy, co wymaga wzmoczonego nadzoru i wstępnego dochodzenia. Krytyczna interwencja pojawia się, gdy wiele wskaźników jednocześnie zawodzi, co wymaga natychmiastowego zbadania przyczyn źródłowych i podjęcia pilnych działań naprawczych.

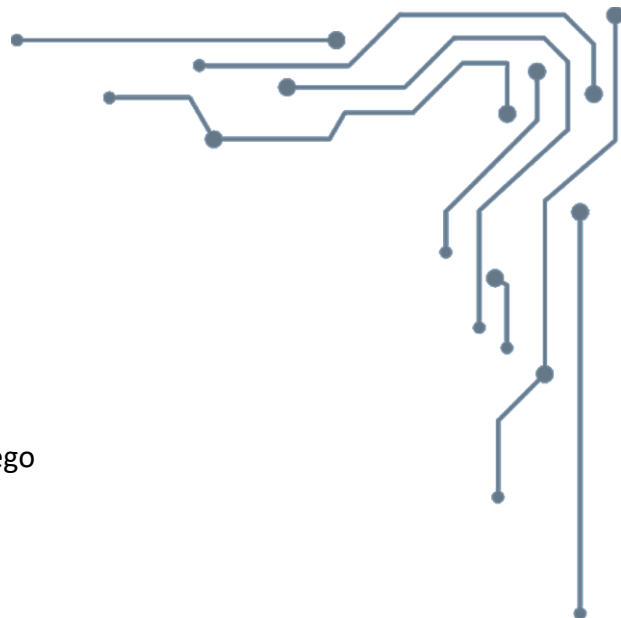
### Strategia wdrażania

Wdrożenie rozpoczyna się od zainstalowania podstawowego panelu monitorującego i ustalenia bazowych wskaźników na podstawie bieżących danych o wydajności systemu. Instytucja wyznacza jednego pracownika IT jako głównego monitora odpowiedzialnego za cotygodniowe kontrole.



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them.  
Grant no:2024-1-LT01-KA220-HED-000251565



Harmonogram wdrożenia:

Tydzień 1-2: Faza konfiguracji

- └─ Zainstaluj panel monitorujący
- └─ Ustal metryki bazowe
- └─ Przypisz personel odpowiedzialny

Tydzień 3+: Operacje

- └─ Co tydzień: 5-minutowa kontrola pulpitu nawigacyjnego
- └─ Miesięcznie: 30-minutowe spotkanie interesariuszy
- └─ Kwartalnik: 2-godzinny kompleksowy przegląd

### Faza operacyjna

Każdego tygodnia (np. w poniedziałkowe poranki) wyznaczony pracownik poświęca pięć minut na sprawdzenie pulpitu nawigacyjnego pod kątem alertów i problemów z rejestrowaniem. Comiesięczne spotkania gromadzą interesariuszy w celu przeglądu trendów wydajności i planowania działań. Kwartalne przeglądy umożliwiają kompleksową ocenę systemu i podejmowanie strategicznych decyzji dotyczących aktualizacji modelu.

### Wymagania dotyczące zasobów i protokół reakcji

Zapotrzebowanie na zasoby ludzkie obejmuje jednego pracownika IT poświęcającego pięć minut tygodniowo oraz jednego przełożonego poświęcającego 30 minut miesięcznie na nadzór. Potrzeby w zakresie infrastruktury technicznej pozostają skromne i obejmują jedynie podstawowe oprogramowanie pulpitu nawigacyjnego, funkcje automatycznego eksportu danych oraz prostą funkcję powiadomień e-mail.

Gdy system monitorowania wykryje niepokojące zmiany, instytucje postępują zgodnie ze zorganizowanym 15-dniowym harmonogramem reakcji:

Protokół odpowiedzi ( $PSI > 0,2$  LUB dokładność  $< 80\%$ ): Zbadaj przyczyny źródłowe → Zbierz bieżące próbki danych → Ponowne przeszkolenie modelu (stare + nowe dane) → Wdróż zaktualizowany model → Zweryfikuj ulepszenie

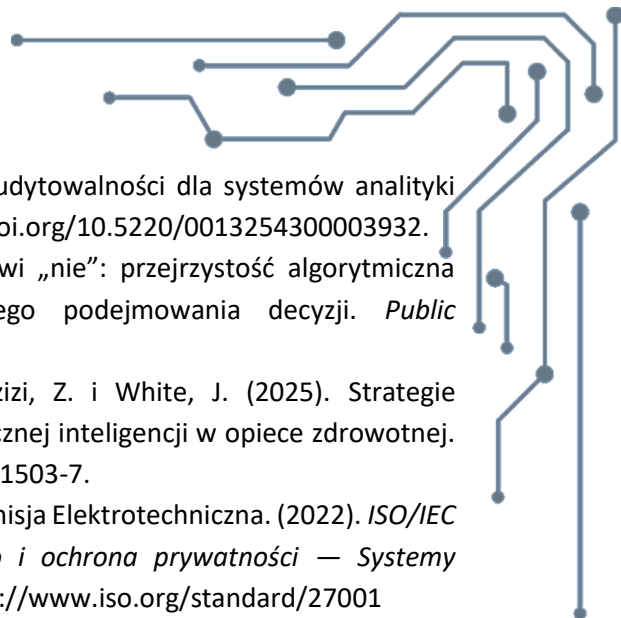
Takie podejście zapewnia szybką reakcję przy jednoczesnym zachowaniu kontroli jakości w całym procesie naprawczym. Usprawniona metodologia zapewnia znaczące korzyści w zakresie wykrywania dryftu przy minimalnej złożoności, co czyni ją niezwykle praktyczną dla instytucji szkolnictwa wyższego działających w warunkach typowych ograniczeń zasobów.



## REFERENCJE

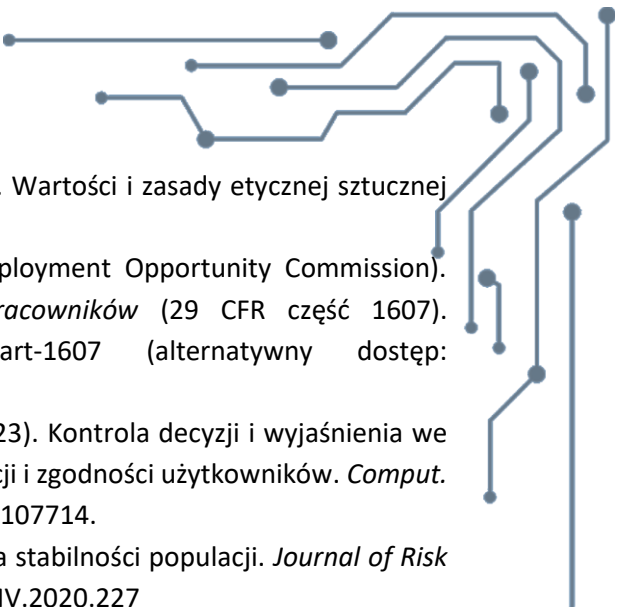
- Acosta-Vargas, P., Salvador-Acosta, B., Novillo-Villegas, S., Sarantis, D. i Salvador-Ullauri, L. (2024). Generatywna sztuczna inteligencja i dostępność sieci: ku inkluzywnej i zrównoważonej przyszłości. *Emerging Science Journal* . <https://doi.org/10.28991/esj-2024-08-04-021>.
- Adeoye, O., Alimi, A., Agboola, O., Akindele, A., Arulogun, O. i Adigun, G. (2025). Rozwój szkolnictwa wyższego dzięki sztucznej inteligencji (AI): ramy integracji nauczania, oceny i badań. *East African Journal of Education Studies* . <https://doi.org/10.37284/eajes.8.2.2946>.
- Sztuczna inteligencja w edukacji: specjalny raport firmy Microsoft, nd).
- Al-Omari, O., Alyousef, A., Fati, S., Shannaq, F. i Omari, A. (2025). Zarządzanie i ramy etyczne integracji sztucznej inteligencji w szkolnictwie wyższym: doskonalenie spersonalizowanego uczenia się i zgodności z przepisami prawa. *Journal of Ecohumanism* . <https://doi.org/10.62754/joe.v4i2.5781>.
- Bates, T., Cobo, C., Mariño, O. i Wheeler, S. (2020). Czy sztuczna inteligencja może przekształcić szkolnictwo wyższe? *International Journal of Educational Technology in Higher Education* , 17. <https://doi.org/10.1186/s41239-020-00218-x>.
- Brown, S., Davidović, J. i Hasan, A. (2021). Audyt algorytmów: ocena algorytmów, które nas oceniają. *Big Data i społeczeństwo* , 8. <https://doi.org/10.1177/2053951720983865>.
- Cheong, BC (2024). Przejrzystość i rozliczalność w systemach AI: Ochrona dobrostanu w dobie algorytmicznego podejmowania decyzji. *Frontiers in Human Dynamics*, 6 , 1421273. <https://doi.org/10.3389/fhumd.2024.1421273>
- Crompton, H. i Burke, D. (2023). Sztuczna inteligencja w szkolnictwie wyższym: stan wiedzy w tej dziedzinie. *International Journal of Educational Technology in Higher Education* , 20, 1-22. <https://doi.org/10.1186/s41239-023-00392-8>
- Crompton, H. i Song, D. (2021). Potencjał sztucznej inteligencji w szkolnictwie wyższym. *Revista Virtualny Universidad Católica del Norte* . <https://doi.org/10.35575/RVUCN.N62A1>.
- DataGuard. (20 sierpnia 2024). Ustawa UE o sztucznej inteligencji i obowiązki dostawców. DataGuard.
- du Pisanie, J., Allison, JS i Visagie, J. (2023). Proponowana technika symulacyjna do testowania stabilności populacji w kartach scoringowych ryzyka kredytowego. *Mathematics*, 11 (2), 492. <https://doi.org/10.3390/math11020492>
- ENQA. (2015). Standardy i wytyczne dotyczące zapewniania jakości w Europejskim Obszarze Szkolnictwa Wyższego (ESG). Europejskie Stowarzyszenie na rzecz Zapewniania Jakości w Szkolnictwie Wyższym, Bruksela, Belgia.
- Parlament Europejski i Rada. (2 grudnia 2016 r.). *Dyrektywa Parlamentu Europejskiego i Rady (UE) 2016/2102 z dnia 26 października 2016 r. w sprawie dostępności stron internetowych i aplikacji mobilnych organów sektora publicznego* . *Dziennik Urzędowy Unii Europejskiej* , L 327, s. 1–15. <https://eur-lex.europa.eu/eli/dir/2016/2102/oj/eng>
- Parlament Europejski i Rada. (2024, 12 lipca). *Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (ustawa o sztucznej inteligencji)* . *Dziennik Urzędowy Unii Europejskiej* . <http://data.europa.eu/eli/reg/2024/1689/oj>
- Europejski Instytut Norm Telekomunikacyjnych. (10 marca 2021 r.). *EN 301 549 V3.2.1: Wymagania dotyczące dostępności produktów i usług ICT* . [https://www.etsi.org/deliver/etsi\\_en/301500\\_301599/301549/03.02.01\\_60/en\\_301549v030201p.pdf](https://www.etsi.org/deliver/etsi_en/301500_301599/301549/03.02.01_60/en_301549v030201p.pdf) (zobacz także stronę przeglądowną: <https://www.etsi.org/human-factors-accessibility/en-301-549-v3-the-harmonized-european-standard-for-ict-accessibility> )





- Fernsel, L., Kalff, Y. i Simbeck, K. (2025). Audyty zaufania: ramy audytowalności dla systemów analityki uczenia się opartych na sztucznej inteligencji. , 51-62. <https://doi.org/10.5220/0013254300003932>.
- Grimmelikhuijsen, S. (2022). Wyjaśnienie, dlaczego komputer mówi „nie”: przejrzystość algorytmiczna wpływa na postrzeganą wiarygodność zautomatyzowanego podejmowania decyzji. *Public Administration Review* . <https://doi.org/10.1111/puar.13483>.
- Hasanzadeh, F., Josephson, C., Waters, G., Adedinsewo, D., Azizi, Z. i White, J. (2025). Strategie rozpoznawania i łagodzenia uprzedzeń w zastosowaniach sztucznej inteligencji w opiece zdrowotnej. *NPJ Digital Medicine* , 8. <https://doi.org/10.1038/s41746-025-01503-7>.
- Międzynarodowa Organizacja Normalizacyjna i Międzynarodowa Komisja Elektrotechniczna. (2022). *ISO/IEC 27001:2022 Bezpieczeństwo informacji, cyberbezpieczeństwo i ochrona prywatności — Systemy zarządzania bezpieczeństwem informacji — Wymagania* . <https://www.iso.org/standard/27001>
- Międzynarodowa Organizacja Normalizacyjna i Międzynarodowa Komisja Elektrotechniczna. (2023). *ISO/IEC 42001:2023 System zarządzania sztuczną inteligencją — Wymagania* . <https://www.iso.org/standard/42001>
- Johnson, M., Liu, X. i McCaffrey, D. (2022). Metody psychometryczne oceny pomiaru i błędów algorytmicznych w automatycznym punktowaniu. *Journal of Educational Measurement* . <https://doi.org/10.1111/jedm.12335>.
- Lazcoz, G. i Hert, P. (2023). Ludzie w zarządzaniu systemami zautomatyzowanymi i algorytmicznymi zgodnie z RODO i AIA. Niezbędne warunki wstępne zapobiegające zrzeczeniu się odpowiedzialności. *Comput. Law Secur. Rev.* , 50, 105833. <https://doi.org/10.1016/j.clsr.2023.105833>.
- Little, RJA i Rubin, DB (2019). *Analiza statystyczna z brakującymi danymi* (wyd. 3). Wiley. <https://doi.org/10.1002/9781119482260>
- Liu, D. i Bates, S. (2025). *Sztuczna inteligencja generatywna w szkolnictwie wyższym: obecne praktyki i kierunki rozwoju* .
- Luo, X., Wang, X. i Jiang, T. (2025). Zastosowanie technologii sztucznej inteligencji w ocenie i kontroli ryzyka audytu: na przykładzie audytu wewnętrznego instytucji szkolnictwa wyższego. *Journal of Infrastructure, Policy and Development* . <https://doi.org/10.24294/jipd10125>.
- Murdoch, B. (2021). Prywatność i sztuczna inteligencja: wyzwania dla ochrony informacji medycznych w nowej erze. *BMC Medical Ethics* , 22. <https://doi.org/10.1186/s12910-021-00687-3>.
- Narodowy Instytut Norm i Technologii [NIST]. (2023). *Ramy zarządzania ryzykiem sztucznej inteligencji (AI RMF 1.0) (NIST AI 100-1)*. <https://doi.org/10.6028/NIST.AI.100-1> (PDF: <https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf> )
- Salhab, W., Ameyed, D., Jaafar, F. i Mcheick, H. (2024). Systematyczny przegląd literatury na temat bezpieczeństwa sztucznej inteligencji: identyfikacja trendów, wyzwań i przyszłych kierunków rozwoju. *IEEE Access* , 12, 131762-131784. <https://doi.org/10.1109/ACCESS.2024.3440647>.
- Turahman, I. (2024). System wspomaganie decyzji dla stypendystów za osiągnięcia w szkołach zawodowych z wykorzystaniem metody analitycznego procesu hierarchicznego. *Engineering: Journal of Mechatronics and Education* . <https://doi.org/10.59923/mechatronics.v1i1.14>.
- UNESCO (2023a). *Wskazówki dotyczące generatywnej sztucznej inteligencji w edukacji i badaniach* (red. F. Miao i W. Holmes). <https://unesdoc.unesco.org/ark:/48223/pf0000386693> (strona docelowa z ostatnią aktualizacją: 14 kwietnia 2025 r.)
- Ocena wpływu etycznego UNESCO (2023b). Narzędzie do oceny wpływu etycznego systemów AI.
- Ocena gotowości UNESCO (2023c). Wytyczne dotyczące oceny gotowości krajowej w zakresie etyki w dziedzinie sztucznej inteligencji.





Rekomendacja UNESCO dotycząca etyki sztucznej inteligencji (2022). Wartości i zasady etycznej sztucznej inteligencji.

Amerykańska Komisja ds. Równych Szans Zatrudnienia (Equal Employment Opportunity Commission). (1978). *Jednolite wytyczne dotyczące procedur doboru pracowników* (29 CFR część 1607). <https://www.ecfr.gov/current/title-29/subtitle-B/chapter-XIV/part-1607> (alternatywny dostęp: <https://www.law.cornell.edu/cfr/text/29/part-1607> ).

Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G. i Rafaeli, A. (2023). Kontrola decyzji i wyjaśnienia we współpracy człowieka ze sztuczną inteligencją: poprawa percepcji i zgodności użytkowników. *Comput. Hum. Behav.* , 144, 107714. <https://doi.org/10.1016/j.chb.2023.107714>.

Yurdakul, B. i Naranjo, JD (2020). Właściwości statystyczne wskaźnika stabilności populacji. *Journal of Risk Model Validation*, 14 (3), 89–100. <https://doi.org/10.21314/JRMV.2020.227>



Co-funded by  
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the National Agency. Neither the European Union nor National Agency can be held responsible for them.  
Grant no:2024-1-LT01-KA220-HED-000251565



**Odwiedź stronę internetową projektu:**

[www.ai-erit.eu](http://www.ai-erit.eu)