

Julia Mucha

Uniwersytet Komisji Edukacji Narodowej w Krakowie, Polska / University of the National Education
Commission, Krakow, Poland

Sztuczna inteligencja w pracy nauczyciela – szanse, ograniczenia i możliwości zastosowania w ocenianiu

Artificial intelligence in the work of teachers: opportunities, limitations and possibilities for use in assessment

Streszczenie: Artykuł podejmuje problem wykorzystania sztucznej inteligencji w edukacji z perspektywy pracy nauczyciela, łącząc przegląd literatury z próbą praktycznej weryfikacji możliwości zastosowania AI w ocenianiu wypowiedzi pisemnych. W pierwszej części omówiono główne obszary, w których technologie oparte na AI mogą wspierać nauczyciela, przede wszystkim personalizację nauczania, przygotowywanie materiałów dydaktycznych, analizę postępów uczniów oraz automatyzację wybranych zadań administracyjnych i organizacyjnych. Zwrócono także uwagę na strukturę czasu pracy nauczycieli, która pokazuje, że znaczna część ich obowiązków obejmuje czynności powtarzalne i czasochłonne, potencjalnie podatne na optymalizację. W drugiej części, empirycznej, przedstawiono eksperyment polegający na wykorzystaniu dużego modelu językowego do odtworzenia logiki oceniania prac studenckich ocenionych wcześniej przez nauczyciela akademickiego. Wyniki wskazują, że AI stosunkowo trafnie odwzorowuje proces oceniania, zwłaszcza w zadaniach o bardziej uporządkowanej strukturze. Jednak w przypadku oceniania tekstów wymagających pogłębionej refleksji, interpretacji i uchwycenia subtelnych różnic jakościowych ujawnia ograniczenia i błędy. Wnioski prowadzą do stwierdzenia, że sztuczna inteligencja może stanowić wartościowe narzędzie wspierające nauczyciela, lecz nie powinna zastępować jego osądu pedagogicznego, doświadczenia i kompetencji relacyjnych.

Abstract: The article examines the use of artificial intelligence in education from the perspective of teachers' work, combining a literature review with a practical attempt to verify the potential of AI in assessing written assignments. The first part discusses the main areas in which AI – based technologies can support teachers, particularly in personalizing instruction, preparing teaching materials, analyzing students' progress, and automating selected administrative and organizational tasks. Attention is also given to the structure of teachers' working time, which shows that a substantial part of their duties consists of repetitive and time-consuming activities that may be optimized. The empirical part presents an experiment involving the use of a large language model to reconstruct the grading logic applied by an academic teacher to students' written work. The results indicate that AI can reproduce the grading process with relatively high accuracy, especially in tasks with a more structured format. At the same time, the

study reveals limitations in assessing texts that require deeper reflection, interpretation, and sensitivity to subtle qualitative differences. The findings lead to the conclusion that artificial intelligence may serve as a valuable tool supporting teachers, but it should not replace their pedagogical judgment, professional experience, or relational competences.

Słowa kluczowe: automatyzacja; edukacja; nauczyciel; ocenianie; personalizacja nauczania; sztuczna inteligencja; technologie edukacyjne

Keywords: artificial intelligence; assessment; automation; education; educational technologies; personalized learning; teacher

Otrzymano: 26 kwietnia 2026

Received: 2 April 2026

Zaakceptowano: 8 czerwca 2026

Accepted: 8 June 2026

Sugerowana cytacja/Suggested citation:

Mucha, J. (2026). Sztuczna inteligencja w pracy nauczyciela – szanse, ograniczenia i możliwości zastosowania w ocenianiu. *Przedsiębiorczość – Edukacja [Entrepreneurship – Education]*, 22(1), 11–24. <https://doi.org/10.24917/20833296.221.1>

Wstęp

Dynamiczny rozwój nowych technologii, w szczególności sztucznej inteligencji, w ostatnich latach stał się istotnym wyzwaniem zarówno dla uczniów, jak i nauczycieli. Minister Barbara Nowacka zauważa, że „Internet i nowe technologie są dla młodych ludzi ich naturalnym środowiskiem, zarówno w kontaktach społecznych, jak i coraz częściej w nauce” (Ministerstwo Edukacji Narodowej, 2025; 2026, 14 czerwca). Uczniowie coraz częściej sięgają po nowoczesne rozwiązania technologiczne, aby nadążyć za trendami, zwiększyć wydajność i ułatwić sobie naukę. Również nauczyciele coraz częściej sięgają po rozwiązania oparte na generatywnej sztucznej inteligencji. Narzędzia oparte na dużych modelach językowych (LLM) umożliwiają automatyzację tworzenia materiałów, analizę danych uczniowskich i personalizację nauczania. Jednocześnie pojawiły się pytania o poziom gotowości nauczycieli do wdrażania tych technologii oraz o wpływ narzędzi AI na jakość procesu dydaktycznego (Walusiak, Ryłko, Nurtazina, 2023). W dobie tych zmian i ciągłego rozwoju technologii ich efektywne wykorzystanie przez nauczycieli pozostaje kluczem do skutecznej nauki. Mimo że rośnie liczba badań dotyczących wykorzystania sztucznej inteligencji w edukacji, to większość z nich koncentruje się na personalizacji nauczania, inteligentnych systemach wspomaganie uczenia się, generowaniu materiałów dydaktycznych oraz analizie postępów uczniów. Znacznie rzadziej podejmowany jest problem wykorzystania dużych modeli językowych (LLM) do wspierania procesu oceniania oraz odtwarzania sposobu podejmowania decyzji przez nauczycieli. W literaturze brakuje badań analizujących, w jakim stopniu sztuczna inteligencja jest zdolna do identyfikacji kryteriów oceniania stosowanych przez nauczyciela oraz reprodukcji przypisywanych przez niego ocen na podstawie wcześniej ocenionych prac. Niniejsze badanie ma na celu częściowe wypełnienie tej luki poprzez ocenę zdolności modelu językowego do rekonstrukcji logiki oceniania nauczyciela akademickiego oraz określenie ograniczeń tego rozwiązania w kontekście prac wymagających refleksji, interpretacji i argumentacji. Szczególną uwagę poświęcono potencjałowi zastosowania narzędzi opartych na dużych modelach językowych w procesie oceniania prac pisemnych

oraz próbie określenia, w jakim stopniu są one w stanie odtworzyć logikę oceniania stosowaną przez nauczyciela. Hipoteza badawcza zakłada, że sztuczna inteligencja oparta na dużych modelach językowych jest zdolna do odtworzenia logiki oceniania stosowanej przez nauczyciela akademickiego oraz przypisywania ocen z wysokim poziomem zgodności względem ocen rzeczywistych. Artykuł ma układ problemowo-empiryczny. W części teoretycznej zastosowano analizę źródeł dotyczących wykorzystania AI w edukacji, ze szczególnym uwzględnieniem korzyści, ograniczeń oraz ryzyka związanych z wdrażaniem tych technologii w pracy nauczyciela. W części empirycznej przeprowadzono eksperyment praktyczny polegający na wykorzystaniu dużego modelu językowego do analizy i oceny prac pisemnych uprzednio ocenionych przez nauczyciela akademickiego. Zastosowane podejście obejmowało analizę jakościową treści, rekonstrukcję kryteriów oceniania, budowę heurystycznego modelu decyzyjnego oraz analizę porównawczą ocen rzeczywistych i ocen wygenerowanych przez model AI.

Przegląd literatury

Wdrażanie technologii opartych na sztucznej inteligencji w procesie nauczania i uczenia się otwiera nowe horyzonty i niesie ze sobą zarówno obiecujące szanse, jak i poważne zagrożenia (Więckiewicz-Modrzewska, 2024). Nowoczesne technologie, w tym sztuczna inteligencja, nie są nowym zjawiskiem w organizacjach zajmujących się edukacją. Grand View Research oszacowało, że wartość globalnego rynku sztucznej inteligencji w edukacji w 2024 r. wynosiła 5,88 mld dolarów, a wzrośnie do 47,7 mld dolarów w 2030 (Grand View Research, 2024; 2026, 14 czerwca). Wraz z rosnącą obecnością AI w różnych procesach i zadaniach coraz większego znaczenia nabierają kompetencje umożliwiające jej świadome i efektywne wykorzystywanie w codziennej pracy (Wieteska, 2024; 2026, 30 marca). Takie narzędzia jak Khan Academy, Coursera, DreamBox Learning czy Knewton umożliwiają dostosowanie treści, tempa pracy i poziomu trudności do indywidualnych potrzeb uczących się. Dzięki temu nauczyciel może łatwiej różnicować proces dydaktyczny i lepiej odpowiadać na zróżnicowane możliwości uczniów. Istotne miejsce zajmują także inteligentne systemy nauczania, takie jak ALEKS, Carnegie Learning czy Knewton. Systemy te analizują odpowiedzi uczniów, identyfikują luki w wiedzy i proponują kolejne zadania adekwatne do poziomu opanowania materiału. W praktyce wspierają nauczyciela w diagnozowaniu trudności edukacyjnych oraz w planowaniu dalszych działań dydaktycznych (Zalewska-Bochenko, 2024). Ważną grupę stanowią narzędzia wykorzystujące przetwarzanie języka naturalnego. Przykładami są Grammarly, QuillBot, Rosetta Stone, Google Cloud Natural Language, IBM Watson oraz Microsoft Azure Cognitive Services. Technologie te wspomagają analizę tekstu, korektę językową, parafrazowanie, tłumaczenie oraz ocenę poprawności wypowiedzi (Owan i in., 2023). Jednak w przypadku automatycznych systemów oceniania wypowiedzi pisemnych (AWE) ich ocena bywa oparta na cechach powierzchownych, np. długości zdań, która sama w sobie nie przesądza o wartości merytorycznej tekstu. Systemy te mają również trudność z uchwyceniem kreatywności i oryginalności wypowiedzi (UNESCO, 2021).

W szerszym ujęciu zagrożeniem związanym z automatyzacją oceniania jest także ograniczenie roli nauczyciela w procesie rozpoznawania strategii uczenia się oraz możliwości poszczególnych uczniów. Nadmierne przekazanie oceny systemom AI może prowadzić do osłabienia znaczenia profesjonalnego osądu pedagogicznego, który uwzględni nie

tylko wynik pracy ucznia, lecz także jego potrzeby społeczne, emocjonalne i rozwojowe. Tym samym sztuczna inteligencja, choć może wspierać proces dydaktyczny, nie powinna zastępować nauczyciela w obszarach wymagających szerszego kontekstu wychowawczego i indywidualnego podejścia. Liczba narzędzi AI i ich zdolności pokazują, jak wielki jest wybór, który może znaleźć zastosowanie w kluczowych obszarach czy procesach edukacyjnych. Stosowanie AI w organizacjach edukacyjnych zależy od kilku czynników, które tworzą etapowość procesu wdrażania sztucznej inteligencji: dostępu do tego rodzaju technologii, odpowiedniej infrastruktury w szkołach, wiedzy i umiejętności doboru odpowiednich narzędzi AI, wiedzy i umiejętności posługiwania się narzędziami AI, w tym umiejętności zachowania krytycznego myślenia, a także świadomości zagrożeń i ograniczeń wynikających ze stosowania tego rodzaju narzędzi (Berkowicz, 2025). Transformatywność sztucznej inteligencji wywołuje również kontrowersje, zwłaszcza w odniesieniu do uczniów. Jan Fazlagić stwierdza, że system edukacji ma za zadanie kształtować ludzką inteligencję. W związku z tym zadaje pytanie, czy „odciążenie” sztucznią inteligencją uczniów w szkole nie będzie kontrproduktywne (Fazlagić, 2022).

Współczesny zawód nauczyciela charakteryzuje się wysokim poziomem obciążenia czasowego, wykraczającym poza formalnie określone pensum dydaktyczne. Liczne obowiązki pozalekcyjne, takie jak przygotowanie zajęć, sprawdzanie prac czy zadania administracyjne, powodują, że nauczyciel często dysponuje ograniczonym zasobem czasu na efektywną organizację pracy. Badanie czasu i warunków pracy nauczycieli, przeprowadzone przez Instytut Badań Edukacyjnych w latach 2011–2012, wskazuje na istotną rozbieżność między deklarowanym a rzeczywistym czasem pracy. Nauczyciele szacują, że pięć podstawowych czynności wykonywanych codziennie zajmuje im średnio 34 godziny i 35 minut tygodniowo, jednak analiza metodą DAR pokazuje, że łączny czas pracy wynosi aż 46 godzin i 40 minut. Oznacza to, że rzeczywiste obciążenia znacząco przekraczają formalne pensum dydaktyczne, przy czym liczba lekcji nie wpływa istotnie na czas przeznaczony na obowiązki pozalekcyjne. Badanie oparto na podejściu mieszanym, obejmującym 4762 nauczycieli z 921 szkół w badaniu ilościowym CAWI, 2617 nauczycieli z 477 szkół w badaniu DAR realizowanym metodą CAPI oraz 24 zogniskowane wywiady grupowe (FGI) (*Nauczyciele – menedżerowie czasu na etacie*, 2023; 2026, 14 czerwca). Jak zostało ukazane w tabeli 1, przygotowanie przed zajęciami oraz sprawdzanie prac po nich zajmują nauczycielom praktycznie tyle samo czasu co prowadzenie zajęć. W rezultacie pojawiają się problem przeciążenia oraz trudności w zachowaniu równowagi między różnymi obszarami aktywności zawodowej.

Tabela 1. Podstawowe zadania nauczycieli z przeznaczonym na nie czasem

Nazwa czynności	Średni czas wykonywania w tygodniu (w minutach)	Udział procentowy w czasie pracy
Prowadzenie lekcji	930	33%
Prowadzenie innych zajęć z uczniami	130	5%
Suma	1060	38%
Przygotowywanie lekcji	425	15%
Przygotowywanie innych zajęć	180	6%

Przygotowanie i sprawdzanie prac sprawdzających (ocenie)	440	16%
Suma	1045	37%
Pozostałe	775	25%

Źródło: opracowanie własne na podstawie: Nauczyciele – menedżerowie czasu na etacie (2023; 2026, 14 czerwca).

Struktura czasu pracy nauczycieli wskazuje na wyraźną dominację dwóch obszarów: prowadzenia zajęć oraz przygotowań, które łącznie pochłaniają aż 75% tygodniowego czasu pracy. Samo prowadzenie zajęć stanowi 38% czasu (w tym 33% to lekcje), a przygotowania – obejmujące przygotowanie materiałów oraz ocenianie – odpowiadają za kolejne 37%. Szczególnie czasochłonne okazuje się sprawdzanie prac i ocenianie (16%), co potwierdza, że znacząca część obowiązków realizowana jest poza bezpośrednim kontaktem z uczniami. Pozostałe czynności zajmują 25% czasu pracy, co wskazuje na istotne rozproszenie zadań i potencjał do ich optymalizacji, zwłaszcza w obszarach administracyjnych i organizacyjnych. Struktura czasu pracy nauczyciela pokazuje, że kluczowym i jednocześnie najmniej podatnym na optymalizację obszarem jest bezpośrednia praca z uczniem. Prowadzenie zajęć dydaktycznych ma charakter interakcyjny i jakościowy, dlatego jego skracanie mogłoby negatywnie wpłynąć na efekty kształcenia (Zarzecki, 2012). Inaczej wygląda sytuacja w obszarze przygotowań, który obejmuje czynności powtarzalne i często czasochłonne, takie jak opracowywanie materiałów czy organizacja treści dydaktycznych. W tym zakresie istnieje realny potencjał optymalizacji, szczególnie w odniesieniu do zadań niewymagających pracy manualnej, w których przypadku możliwe jest wykorzystanie narzędzi wspierających, automatyzacji lub standaryzacji procesów, bez pogorszenia jakości nauczania.

W tym kontekście zasadne staje się poszukiwanie możliwości optymalizacji czasu pracy, zarówno poprzez lepsze zarządzanie zadaniami, jak i potencjalne wsparcie narzędziowe oraz organizacyjne, które mogłoby zwiększyć efektywność wykonywanych obowiązków bez obniżenia jakości procesu dydaktycznego.

Wprowadzenie nowoczesnych programów automatyzujących procesy może w znacznym stopniu zwiększyć wydajność nauczycieli. Przykłady takich optymalizacji są szeroko stosowane w wielu firmach z bardzo pozytywnym rezultatem. Na podstawie ubiegłorocznego raportu firmy Gartner można wnioskować, że automatyzacja rutynowych procesów księgowych za pomocą specjalistycznego oprogramowania pozwala na skrócenie czasu ich realizacji nawet o 30% (Forbes, 2024; 2026, 14 czerwca). Innym przykładem działań ukierunkowanych na optymalizację czasu pracy w obszarze zadań rutynowych jest wdrażanie chatbotów oraz inteligentnych asystentów. W ramach partnerstwa z Microsoftem Żabka Polska jako pierwsza firma w Polsce udostępniła 2,5 tys. pracownikom narzędzie Copilot dla Microsoft 365. System ten, zintegrowany z takimi aplikacjami jak Outlook, Excel, Word, PowerPoint i Teams, wspiera realizację powtarzalnych i czasochłonnych czynności, przez co przyczynia się do zwiększenia efektywności pracy oraz uwolnienia zasobów czasowych, które mogą zostać przeznaczone na działania bardziej kreatywne i rozwojowe. Do funkcji tego rozwiązania należą m.in. podsumowywanie spotkań, wspomaganie redagowania wiadomości e-mail, przygotowywanie prezentacji oraz wsparcie w planowaniu tygodnia pracy (Żabka, 2024; 2026, 23 kwietnia).

Jak wskazują opracowania praktyczne, AI może wspierać nauczyciela w tworzeniu konspektów lekcji, przygotowywaniu quizów czy dostosowywaniu materiałów do poziomu uczniów, znacząco skracając czas realizacji tych zadań. Szczególnie istotne jest to, że narzędzia te przejmują wykonywanie czynności powtarzalnych i schematycznych, przez co pozwalają na ograniczenie nakładu pracy niewymagającej bezpośredniego zaangażowania pedagogicznego (WSIP, 2025; 2026, 14 czerwca). Sztuczna inteligencja może wspierać nauczyciela nie tylko w przygotowywaniu materiałów dydaktycznych, personalizacji nauczania czy ocenianiu prac, ale także w szerszym organizowaniu procesu edukacyjnego. Jednym z istotnych obszarów jest wspomaganie komunikacji, np. poprzez przygotowywanie jasnych komunikatów, instrukcji do zadań, informacji zwrotnych lub podsumowań postępów ucznia. AI może również pomagać w planowaniu pracy dydaktycznej, tworzeniu harmonogramów, rozkładów materiału, scenariuszy zajęć czy propozycji aktywności dostosowanych do wieku i poziomu uczniów (Więckiewicz-Modrzejewska, 2026). Ważnym zastosowaniem jest także wsparcie w pracy z uczniami o zróżnicowanych potrzebach edukacyjnych, m.in. poprzez upraszczanie tekstów, generowanie materiałów o różnym poziomie trudności, przygotowywanie ćwiczeń utrwalających lub alternatywnych form wyjaśnienia danego zagadnienia. Należy również podkreślić, że nauczyciele mogą wykorzystywać AI do generowania treści pytań zamkniętych i otwartych zgodnych z celami nauczania zawartymi w podstawie programowej, do wykorzystania na sprawdzianach (Baidoo-Anu, Ansah, 2023). Kolejnym obszarem jest wspieranie refleksji nauczyciela nad własną praktyką. Narzędzia AI mogą pomagać w analizie przebiegu zajęć, porządkowaniu wniosków, przygotowywaniu samooceny pracy czy planowaniu dalszego rozwoju zawodowego. Mogą również wspierać nauczyciela w wyszukiwaniu inspiracji dydaktycznych, nowych metod pracy oraz przykładów dobrych praktyk.

Jednocześnie podkreśla się, że sztuczna inteligencja nie stanowi substytutu nauczyciela, lecz pełni funkcję wspierającą. Nie zastępuje ona kompetencji związanych z budowaniem relacji, rozumieniem emocji ucznia czy interpretacją kontekstu edukacyjnego, które to kompetencje pozostają domeną człowieka. W konsekwencji implementacja narzędzi AI może prowadzić do bardziej efektywnego wykorzystania czasu pracy nauczyciela przez przesunięcie akcentu z czynności technicznych na działania wymagające kompetencji interpersonalnych i dydaktycznych.

Wykorzystanie sztucznej inteligencji w celu sprawdzania zadań pisemnych – eksperyment praktyczny

Sztuczna inteligencja to system, który pozwala na wykonywanie zadań wymagających procesu uczenia się i uwzględniania nowych okoliczności w toku rozwiązywania danego problemu i który może w różnym stopniu, w zależności od konfiguracji, działać autonomicznie oraz wchodzić w interakcję z otoczeniem (Zalewski, 2020). Obecnie dużą popularnością cieszą się duże modele językowe, takie jak ChatGPT, Gemini czy Copilot. Badania potwierdzają, że społeczeństwo coraz częściej korzysta z narzędzi opartych na sztucznej inteligencji. Badania, przeprowadzone m.in. przez zespół naukowców z University of Melbourne we współpracy z firmą KPMG, objęły próbą ponad 48 tys. respondentów z 47 krajów, w tym 1082 osoby z Polski. Według badania niemal 70% badanych Polaków regularnie korzysta ze sztucznej inteligencji (2025; 2026, 26 kwietnia).

Dane

W badaniu wykorzystano dwa zadania pisemne przygotowane przez studentki czwartego roku jednolitych studiów magisterskich na kierunku pedagogika przedszkolna i wczesnoszkolna w roku akademickim 2025–2026. Prace oraz oceny zostały udostępnione dobrowolnie przez autorki, które uprzednio poinformowano o celu badania, jego przebiegu oraz zapewnieniu pełnej anonimowości. Udostępnienie pracy było równoznaczne z wyrażeniem zgody na jej wykorzystanie w analizie oraz poddanie ocenie. Wszystkie opracowania zostały początkowo ocenione przez nauczyciela akademickiego ze stopniem doktora. Pierwszym zadaniem było napisanie listu otwartego do ministra edukacji. W liście należało się odnieść do współczesnych wyzwań edukacyjnych:

1. Dlaczego zmiana w edukacji jest konieczna?
2. Jakie bariery (systemowe, społeczne, kulturowe, organizacyjne) mogą tę zmianę utrudniać?
3. Jaką rolę w procesie zmiany odgrywa współczesny nauczyciel jako profesjonalista, refleksyjny praktyk i uczestnik przemian cywilizacyjnych?

Sumarycznie w analizie zgromadzono 38 prac, które podzielono na zbiory w proporcji 80% : 20%, przy czym 80% materiału wykorzystano do uczenia modelu, a pozostałe 20% do jego walidacji.

Drugie zadanie polegało na opracowaniu koncepcji szkoły roku 2050. W ramach zadania studentki miały wyobrazić sobie swój udział w zespole projektowym oraz przedstawić wizję przyszłego systemu edukacji, uwzględniając jego organizację, metody nauczania oraz rolę nauczyciela i ucznia w zmieniającym się kontekście społecznym i technologicznym – „Opisz jeden kluczowy fundament tej szkoły (np. aksjologiczny, społeczny, organizacyjny, technologiczny lub obywatelski) i uzasadnij, dlaczego właśnie on powinien stanowić podstawę funkcjonowania edukacji przyszłości”. Wymagania:

1. Wyjaśnij, na czym polega wybrany fundament.
2. Odnieś go do wyzwań przyszłości (np. niepewność, zmiany technologiczne, kryzysy globalne, rynek pracy).
3. Podaj co najmniej jeden przykład, jak ten fundament mógłby funkcjonować w praktyce szkolnej. wypowiedź powinna mieć charakter krytyczno-refleksyjny, uwzględnić aktualne orientacje w edukacji oraz odwoływać się do wiedzy teoretycznej i doświadczeń praktycznych.

W badaniu uzyskano łącznie 32 wyniki, które podzielono na zbiory w proporcji 24:8 (75% : 25%), przy czym 75% danych wykorzystano do uczenia modelu, a pozostałe 25% przeznaczono do jego walidacji.

Skala ocen stosowana w odniesieniu do obu zadań obejmowała przedział od 3,0 do 5,0, przy czym dopuszczono stopniowanie co 0,5 punktu. Oznacza to, że możliwe było przypisanie następujących ocen: 3,0; 3,5; 4,0; 4,5 i 5,0.

Zastosowane metody

W badaniu zastosowano podejście oparte na wykorzystaniu modelu językowego klasy LLM (Large Language Model), pełniącego funkcję głównego narzędzia analitycznego. W szczególności wykorzystano model ChatGPT (GPT-5), który umożliwił przeprowadzenie wieloetapowej analizy treści prac uczniów oraz rekonstrukcji mechanizmów oceniania.

Model odpowiadał za analizę semantyczną tekstów, identyfikację wzorców oceniania, odtworzenie kryteriów stosowanych przez nauczyciela (*reverse engineering*), budowę modelu oceny i symulację procesu decyzyjnego. Analiza została osadzona w paradygmacie przetwarzania języka naturalnego (NLP), obejmującym zarówno analizę semantyczną, jak i analizę dyskursu. Pozwoliło to na ocenę jakości argumentacji, spójności wypowiedzi oraz poziomu refleksyjności tekstów. Dodatkowo przeprowadzono ekstrakcję cech jakościowych, takich jak obecność argumentów, odniesienia do teorii, stopień konkretności czy struktura wypowiedzi. Mechanizmy te stanowiły integralną część modelu językowego i nie wymagały zastosowania odrębnych narzędzi. W celu przeprowadzenia badania modelowi podano następujący prompt: „W celu przeprowadzenia badania dotyczącego możliwości odtworzenia sposobu oceniania prac pisemnych przez nauczycieli zostaniesz poproszony o przeprowadzenie analiz prac już ocenionych oraz na tej podstawie ocenę innych prac. W pierwszym etapie przekażę szczegółowe polecenie zadania. W ramach tego zadania możliwy był wybór jednego z dwóch tematów. Następnie dostarczę zestaw prac już ocenionych przez nauczyciela. Każda praca będzie oznaczona numerem tekstu oraz przypisaną oceną. Skala ocen obejmuje przedział od 2,0 do 5,0, ze stopniowaniem co 0,5. Celem analizy będzie zrekonstruowanie sposobu oceniania zastosowanego przez nauczycielkę. Należy przeanalizować, jakie kryteria mogły wpływać na ocenę, jakie zależności występują między cechami tekstu a przyznaną oceną oraz jakie wzorce decyzyjne powtarzają się w ocenionych pracach. Szczególną uwagę należy zwrócić na zgodność pracy z poleceniem, jakość argumentacji, poziom refleksyjności, wykorzystanie wiedzy teoretycznej, konkretność przykładów, spójność wypowiedzi oraz sposób rozwinięcia tematu. Analizie powinna podlegać wyłącznie warstwa merytoryczna tekstów. Nie należy uwzględniać formatowania, układu strony, czcionki, odstępów ani ogólnego wyglądu dokumentu. Po przeprowadzeniu analizy przekażę plik zawierający prace bez ocen. Zadaniem będzie ocenienie ich zgodnie z odtworzonym wzorcem oceniania, tak aby zaproponowane oceny były możliwie najbardziej zbliżone do ocen, które mogłaby wystawić nauczycielka. Przy każdej pracy należy podać proponowaną ocenę, krótkie uzasadnienie oraz wskazać najważniejsze cechy tekstu, które wpłynęły na wynik”.

Kluczowym etapem badania było opracowanie heurystycznego modelu oceniania opartego na analizie przykładów (*case-based reasoning*). Na podstawie porównania prac ocenionych przez nauczyciela zrekonstruowano zestaw reguł decyzyjnych odzwierciedlających jego sposób oceniania. Model ten przyjął postać systemu reguł heurystycznych, w którym określone konfiguracje cech tekstu (np. obecność bibliografii, poziom refleksji, stopień konkretności argumentacji) odpowiadały określonym progom ocen. Szczegółowe kryteria dla pierwszego zadania zawiera tabela 2.

Tabela 2. Kryteria oceny dla zadania „List otwarty do ministra edukacji”

Kryterium	Waga	5,0	4,0–4,5	3,5	3,0
Realizacja polecenia	30%	wszystkie trzy elementy (zmiana, bariery, rola nauczyciela) rozwinięte i powiązane	wszystkie elementy obecne, ale mniej pogłębione	brakuje 1 elementu lub jest potraktowany marginalnie	braki w 2 elementach lub bardzo powierzchowne ujęcie

Krytyczno-refleksyjna argumentacja	25%	analiza przyczyn, konsekwencji, napięć; własny zamysł	opis problemów + częściowa refleksja	głównie opis, niewielka refleksja, brak głębi	hasłowość, brak analizy, ogólniki
Wiedza teoretyczna	20%	literatura użyta do argumentacji, bibliografia	ogólne odniesienia do teorii	pojedyncze, powierzchowne odniesienia	brak teorii lub całkowicie deklaratoryjna
Konkret i wykonalność	15%	konkretne rozwiązania i przykłady	pojawiają się propozycje, ale ogólne	sporadyczne lub bardzo ogólne przykłady	brak przykładów, same postulaty
Spójność i kompozycja	10%	spójny, logiczny wywód	drobne niespójności	widoczne przeskokki, słabsza logika	niespójny, fragmentaryczny tekst

Źródło: ChatGPT.

Dobór przyjętych założeń oceniania należy uznać za metodycznie trafny i spójny z charakterem zadania, którego istotą jest formułowanie wypowiedzi o charakterze krytyczno-refleksyjnym. Zastosowane kryteria obejmują kluczowe komponenty odpowiadające wymaganiom stawianym tego typu pracom, takie jak realizacja polecenia, jakość argumentacji, wykorzystanie wiedzy teoretycznej, poziom operacjonalizacji wniosków oraz spójność wypowiedzi. Taki zestaw kryteriów jest zgodny z przyjętymi metodami oceniania prac, zapewnia pełne pokrycie ocenianej kompetencji, łącząc elementy wiedzy, rozumienia i analizy z umiejętnością ich praktycznego zastosowania (Merta, Pacewicz, 2010). Istotnym atutem przyjętych założeń jest także odpowiednia hierarchia wag, w której największe znaczenie przypisano realizacji polecenia oraz jakości argumentacji. Jednocześnie uwzględnienie kryterium wiedzy teoretycznej oraz konkretności i wykonalności pozwala na ocenę stopnia integracji refleksji z praktyką edukacyjną. Należy jednak wskazać pewne ograniczenia przyjętego modelu. Część kryteriów, w szczególności odnoszących się do poziomu refleksyjności i pogłębienia argumentacji, ma charakter interpretacyjny, co może wprowadzać element subiektywności do procesu oceniania. Jednak, jak zauważa L. Zacher (2011), ponowoczesny człowiek ma być refleksyjny i kreatywny. A edukacja ma zapobiec pasywności i zmarginalizowaniu człowieka. Człowiek działający schematycznie i rutynowo „zostanie w tyle”. Jednak przez sztuczną inteligencję działanie pozaschematyczne zostanie uznane za błąd. Ponadto wyraźne premiowanie odniesień teoretycznych i obecności literatury może prowadzić do uprzywilejowania stylu akademickiego kosztem prac bardziej praktycznych, choć merytorycznie trafnych. Wątpliwości może budzić także stosunkowo niska waga przypisana spójności wypowiedzi, która w praktyce często istotnie wpływa na odbiór i wartość tekstu.

Pomimo wskazanych ograniczeń przyjęte założenia oceniania można uznać za dobrze zaprojektowane i adekwatne do celu badania. Przyjęte założenia zostały pozytywnie ocenione i na ich podstawie dopuszczono model do stworzenia analizy.

Podobny zestaw kryteriów został opracowany przez model dla zadania drugiego (tabela 3).

Tabela 3. Kryteria oceny dla zadania „Szkoły 2050”

Kryterium	Waga	5,0	4,0–4,5	3,5	3,0
Fundament i jego konceptualizacja	30%	precyzyjny, dobrze zdefiniowany	sensowny, ale ogólny	fundament nie w pełni doprecyzowany	fundament niejasny lub bardzo ogólny
Odniesienie do wyzwań przyszłości	25%	analiza zależności i skutków	wyzwania opisane, ale mniej pogłębione	wymienione bez rozwinięcia	brak realnego powiązania
Przykład działania	20%	konkretna, realistyczna praktyka	poprawny, ale opisowy przykład	bardzo ogólny przykład	brak przykładu lub czysto teoretyczny
Teoria i orientacje	15%	wyraźne osadzenie w literaturze	odniesienia obecne, ale słabsze	sporadyczne lub powierzchowne	brak teorii
Krytyczno-refleksyjna spójność	10%	refleksja + świadomość ograniczeń	częściowa refleksja	minimalna refleksja	brak refleksji

Źródło: ChatGPT.

Dobór założeń przyjętych w rubryce oceny zadania dotyczącego projektowania szkoły 2050 r., podobnie jak w poprzednim zadaniu, należy uznać za spójny z jego koncepcyjnym i problemowym charakterem. Konstrukcja kryteriów odzwierciedla specyfikę wypowiedzi opartej na myśleniu projektowym, wymagającej nie tylko opisu idei, lecz także jej uzasadnienia, osadzenia w kontekście współczesnych wyzwań oraz przełożenia na praktykę edukacyjną. Uwzględnienie takich obszarów jak konceptualizacja fundamentu, odniesienie do wyzwań przyszłości, przykład działania, wykorzystanie teorii oraz spójność refleksyjna pozwala na kompleksową ocenę zarówno poziomu abstrakcji, jak i zdolności operacjonalizacji idei. Szczególnie istotne jest przypisanie najwyższej wagi kryterium dotyczącemu fundamentu i jego konceptualizacji.

Kryteria obu zadań są spójne, dobrze zbudowane oraz stanowią ważny zarys sposobu oceniania. Do potencjalnych ograniczeń należy zaliczyć częściową trudność w jednoznacznej ocenie stopnia precyzyjności fundamentu oraz poziomu refleksji, które mają charakter interpretacyjny. Ponadto w obu kryteriach widoczne jest uprzywilejowanie stylu akademickiego poprzez premiowanie odniesień teoretycznych. Mimo to przyjęte założenia należy ocenić jako adekwatne i funkcjonalne, ponieważ skutecznie różnicują poziomy jakości wypowiedzi oraz promują myślenie systemowe, refleksyjne i osadzone w realiach współczesnej edukacji.

Uzupełniająco zastosowano elementy analizy ilościowej, inspirowane metodami stosowanymi w środowisku Python (np. biblioteka Scikit-learn), w szczególności analizę zależności między wybranymi cechami tekstu a oceną. Pozwoliło to na weryfikację, czy obserwowane zależności jakościowe znajdują potwierdzenie w prostych relacjach korelacyjnych (np. długość tekstu, obecność bibliografii, poziom szczegółowości). Istotnym komponentem badania była również analiza porównawcza, oparta na rozpoznawaniu wzorców (*pattern recognition*). Model identyfikował powtarzalne schematy oceniania poprzez zestawienie prac o różnych poziomach ocen (np. 4,0; 4,5; 5,0), co pozwoliło na wyodrębnienie progów

jakościowych i typowych konfiguracji cech dla poszczególnych ocen. Efektem końcowym było opracowanie modelu decyzyjnego o charakterze regułowym (*rule-based scoring model*), który przyjmuje tekst jako dane wejściowe i generuje ocenę wraz z uzasadnieniem. Model ten stanowi symulację procesu myślowego nauczyciela, odwzorowuje jego sposób podejmowania decyzji o ocenach na podstawie zidentyfikowanych kryteriów i zależności.

Wyniki

W celu oceny trafności modelu sztucznej inteligencji przeprowadzono analizę porównawczą ocen przypisanych przez nauczyciela oraz ocen wygenerowanych przez model AI dla dwóch typów zadań: listu otwartego do ministra edukacji oraz pracy koncepcyjnej „Szkoła 2050” (tabele 4 i 5). Dodatkowo obliczono względny błąd oceny, umożliwiającą ilościową ocenę stopnia rozbieżności.

Tabela 4. List otwarty – porównanie ocen

Ocena rzeczywista	Ocena AI
4,0	4,5
5,0	4,0
4,0	4,5
4,5	4,0
4,5	4,0
4,5	4,0
4,5	4,0
4,0	4,5

Źródło: opracowanie własne.

W przypadku listu otwartego obserwuje się umiarkowany poziom zgodności ocen. Wartości błędu, z wyjątkiem jednej pracy, mieszczą się w przedziale jednego poziomu skali (0,5 punktu). Widoczna jest systematyczna tendencja modelu do „spłaszczania” ocen – AI ma skłonność do obniżania ocen najwyższych oraz podwyższania ocen niższych. Średni błąd względny wynosi ok. 0,125, co wskazuje na relatywnie stabilne, choć nie w pełni precyzyjne odwzorowanie ocen.

Tabela 5. Szkoła 2050 – porównanie ocen

Ocena rzeczywista	Ocena AI
4,0	4,0
4,5	4,0
4,0	4,0
4,5	4,5
3,0	3,0
5,0	4,5
4,5	4,5
4,5	4,5

Źródło: opracowanie własne.

W przypadku zadania dotyczącego szkoły przyszłości poziom zgodności ocen jest wyraźnie wyższy. W większości przypadków model AI przypisał ocenę identyczną jak nauczyciel, co znajduje odzwierciedlenie w zerowej wartości błędu. Średni błąd dla tego zadania jest istotnie niższy niż w przypadku listu otwartego, co wskazuje na większą trafność modelu w ocenie prac o charakterze koncepcyjnym.

Przeprowadzona analiza wskazuje, że model sztucznej inteligencji w wysokim stopniu odtwarza logikę oceniania stosowaną przez nauczyciela, przy czym skuteczność ta jest zróżnicowana w zależności od typu zadania. W przypadku tekstów wymagających pogłębionej argumentacji i refleksji (list otwarty) model wykazuje tendencję do uśredniania ocen, co może wynikać z trudności w uchwyceniu subtelnych różnic jakościowych. Natomiast w zadaniach o bardziej jednoznacznej strukturze (projekt szkoły 2050) trafność ocen jest wyższa, a rozbieżności marginalne.

Uzyskane wyniki potwierdzają, że sztuczna inteligencja może stanowić efektywne narzędzie wspierające proces oceniania poprzez zwiększanie jego spójności i powtarzalności. Jednocześnie wskazują na konieczność zachowania roli nauczyciela jako ostatecznego decydenta, szczególnie w przypadku ocen wymagających pogłębionej interpretacji i oceny jakości argumentacji.

Podsumowanie

Rozwój nowych technologii, w tym sztucznej inteligencji, istotnie zmienia sposoby funkcjonowania szkoły oraz pracy nauczyciela. Współczesne środowisko edukacyjne jest coraz silniej osadzone w rzeczywistości cyfrowej, co sprawia, że efektywne wykorzystanie narzędzi technologicznych staje się ważnym elementem kompetencji nauczyciela. Jednocześnie analiza struktury czasu pracy wskazuje na znaczące obciążenie obowiązkami pozalekcyjnymi, które w dużej mierze mają charakter powtarzalny i czasochłonny. Największy potencjał optymalizacji dotyczy obszaru przygotowań dydaktycznych oraz zadań administracyjnych, które w przeciwieństwie do bezpośredniej pracy z uczniem mogą być wspierane przez narzędzia cyfrowe, automatyzację i standaryzację procesów. Doświadczenia z innych sektorów gospodarki oraz przykłady wdrożeń pokazują, że zastosowanie nowoczesnych rozwiązań może znacząco zwiększyć efektywność pracy bez pogorszenia jej jakości.

Hipoteza badawcza została częściowo potwierdzona. Przeprowadzony eksperyment potwierdza, że sztuczna inteligencja jest w stanie w dużym stopniu odtworzyć logikę oceniania stosowaną przez nauczyciela, szczególnie w przypadku zadań o bardziej uporządkowanej strukturze. Jednocześnie ujawniają się ograniczenia modeli AI w ocenie wypowiedzi wymagających pogłębionej refleksji i interpretacji, co prowadzi do tendencji do uśredniania ocen. W konsekwencji sztuczna inteligencja powinna być traktowana jako narzędzie wspierające, a nie zastępujące nauczyciela. Należy również zauważyć ograniczenia przeprowadzonej analizy, m.in. mały zbiór uczący, punkty oceniania sformułowane ogólnikowo oraz brak odpowiedzi zwrotnej dla sztucznej inteligencji. Przeprowadzona analiza wykazała potencjał tego rozwiązania, jednak w dalszych analizach należy poszerzyć zbiór badanych przedmiotów, a także rozszerzyć badanie o dodatkowy etap nauki z użyciem informacji zwrotnej od nauczycieli. Należy również zauważyć, że sztuczna inteligencja nie musi w całości przejmować zadań oceniania prac, a może być jedynie pomocą sugerującą oceny na podstawie danych historycznych. Największą wartością AI w pracy nauczyciela pozostanie więc przejmowanie zadań rutynowych, co pozwala nauczycielowi skoncentrować

się na działaniach wymagających kompetencji pedagogicznych, relacyjnych i interpretacyjnych. Takie podejście będzie sprzyjać bardziej efektywnemu wykorzystaniu czasu pracy oraz podniesieniu jakości procesu dydaktycznego.

Literatura

References

- Berkowicz, A. (2025). Sztuczna inteligencja w organizacjach edukacyjnych: narzędzia, obszary wdrażania oraz korzyści i ryzyka stosowania AI. W: M. Wójcik, A. Berkowicz (red.), *AI w organizacjach publicznych. Rozwój, zastosowanie, trendy i wyzwania. Perspektywa interdyscyplinarna*, Kraków: Uniwersytet Jagielloński, 54–66.
- Baidoo-Anu, D., Ansah, L.O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52–62. DOI: <https://doi.org/10.2139/ssrn.4337484>.
- Fazlagić, J. (2022). The development of artificial intelligence as a challenge for the education system. W: J. Fazlagić (red.), *Artificial Intelligence (AI) as a Megatrend Shaping Education: How to Prepare for the Socio-Economic Opportunities and Challenges Presented by Artificial Intelligence*, Warsaw: IBE, 24–37.
- Forbes. (2024; 2026, 14 czerwca). *Zbyt duże koszty w firmie? Postaw na automatyzację*. Pozyskano z: <https://www.forbes.pl/biznes/chmura-obliczeniowa-przyspiesza-procesy-i-zmniejsza-koszty/87mwj4t>.
- Grand View Research. (2024; 2026, 14 czerwca). *AI in education market (2026–2033): Size, Share & Trends Analysis Report By Component (Solution, Services), By Deployment (Cloud, On-premises), By Technology (NLP, ML), By Application (Learning Platform & Virtual Facilitators), By End Use, By Region, And Segment Forecasts*. Pozyskano z: <https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-education-market-report>.
- KPMG. (2025; 2026, 26 kwietnia). *Sztuczna inteligencja w Polsce – krajobraz pełen paradoksów*. Pozyskano z: <https://assets.kpmg.com/content/dam/kpmgsites/pl/pdf/2025/07/pl-Raport-KPMG-w-Polsce-KPMG-AI-Trust-2025-web.pdf.coredownload.inline.pdf>.
- Ministerstwo Edukacji Narodowej. (2025; 2026, 14 czerwca). *Barbara Nowacka: „Cyfrowa rewolucja w polskich szkołach już się dzieje”*. Pozyskano z: <https://www.gov.pl/web/edukacja/barbara-nowacka-cyfrowa-rewolucja-w-polskich-szkolach-juz-sie-dzieje>.
- Merta, T., Pacewicz, A. (2010). *Kształcenie obywatelskie w szkołach samorządowych. Jak oceniać uczniów? Wskazówki dla nauczycieli*. Warszawa: Centrum Edukacji Obywatelskiej.
- Nauczyciele – menedżerowie czasu na etacie*. (2023; 2026, 14 czerwca). Pozyskano z: <https://www.ibe.edu.pl/images/prasa/informacja-prasowa-czas-pracy-nauczycieli.pdf>.
- Owan, V.J., Kingsley B.A., Idika D.O., Etta, E.O., Bassey B.A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education*, 19(8), 1–15.
- UNESCO. (2021). *AI and education. Guidance for policy-makers*. Paris: UNESCO.
- Walusiak, Ł., Ryłko, N., Nurtazina, K. (2023). Przemiany cyfrowe w edukacji. *Edukacja Biologiczna i Środowiskowa*, 2(80), 172–181.
- Wieteska, M. (2024; 2026, 30 marca). *Era sztucznej inteligencji. Kompetencje i umiejętności przyszłości – jak zmieni się rynek i edukacja w obliczu rewolucji AI?* Pozyskano z: <https://sklep.infor.pl/artykul-era-sztucznej-inteligencji-kompetencje-i-umiejtnosci-przyszlosci-jak-zmieni-sie-rynek-i-edukacja-w-obliczu-rewolucji-ai>.
- Więckiewicz-Modrzewska, J. (2024). Sztuczna inteligencja w edukacji – szanse i zagrożenia. *Szkoła Specjalna*, 85(2), 128–136.

- Więckiewicz-Modrzewska, J. (2026). ChatGPT w pracy nauczyciela. O możliwościach. *Annales Universitatis Mariae Curie-Skłodowska. Sectio J, Paedagogia-Psychologia*, 39(1), 187–197. DOI: <https://doi.org/10.17951/j.2026.39.1.187-197>
- WSIP. (2025; 2026, 14 czerwca). *Narzędzia AI w codziennej pracy nauczyciela*. Pozyskano z: <https://wsip.pl/blog/narzedzia-ai-w-codziennej-pracy-nauczyciela/>.
- Zacher, L. (2011). Konteksty rozwoju edukacji. W: A. Karpińska, K. Borawska-Kalbarczyk (red.), *Horyzonty rozwoju edukacji w dialogu i perspektywie*. Warszawa: Difin, 27–61.
- Zalewska-Bochenko, A. (2024). Sztuczna inteligencja w procesie edukacji. *Optimum. Economic Studies*, 2, 194–210.
- Zalewski, T. (2020). Definicja sztucznej inteligencji. W: L. Lai, M. Świerczyński (red.), *Prawo sztucznej inteligencji*. Warszawa: C.H. Beck, 1–14.
- Zarzecki, L. (2012). *Teoretyczne podstawy wychowania. Teoria i praktyka w zarysie*. Jelenia Góra: Karkonoska Państwowa Szkoła Wyższa.
- Żabka. (2024; 2026, 23 kwietnia). *Żabka podpisała strategiczne partnerstwo z Microsoft w obszarze AI*. Pozyskano z: https://cdn.zabka.pl/wp-content/uploads/2024/07/11155453/Komunikat_Zabka-podpisała-strategiczne-partnerstwo-z-Microsoft-w-obszarze-AI_11-07-2024.pdf.

Julia Mucha, studentka czwartego roku jednolitych studiów magisterskich na kierunku pedagogika przedszkolna i wczesnoszkolna. Jej główne zainteresowania naukowo-badawcze koncentrują się wokół wykorzystania nowych technologii w edukacji, innowacji edukacyjnych i sztucznej inteligencji. Zdobyte doświadczenie i umiejętności wykorzystuje również w życiu zawodowym, łącząc studiowanie teoretycznych aspektów pedagogiki z realną pracą z uczniem.

Julia Mucha is a fourth-year student of a long-cycle Master's program in Preschool and Early School Education. Her main academic and research interests focus on the use of new technologies in education, educational innovations and artificial intelligence. She also applies her acquired experience and skills in her professional life, combining the study of theoretical aspects of pedagogy with practical work with students.

ORCID: <https://orcid.org/0009-0006-8689-5686>

Adres/Address:

Uniwersytet Komisji Edukacji Narodowej w Krakowie
Wydział Pedagogiki i Psychologii
ul. Romana Ingardena 4
30-060 Kraków
e-mail: juliamucha_2003@o2.pl