

Piotr Kowzan

<https://doi.org/10.26881/pwe.2025.61.06>

ORCID: 0000-0002-6506-8327

Uniwersytet Gdański

piotr.kowzan@ug.edu.pl

Mowa eksplanacyjna w rozmowach dzieci z chatbotami GPT: szanse, ryzyka i implikacje pedagogiczne

Summary

Explanatory Talk in Children's Conversations with GPT Chatbots: Opportunities, Risks, and Pedagogical Implications

The article addresses the problem of default “closure” in child–AI conversations produced by large language models (LLMs), which often provide coherence that is fluent yet cognitively shallow. Using opportunistic data (voice-mode transcripts of a six-year-old interacting with a GPT chatbot), it analyzes the conditions that support explanatory talk and propose a short protocol of pre-dispositions/meta-prompts together with a polyphonic response mode (multiple explanatory variants). These solutions increase the number of accurate explanations and meaningful follow-up questions while reducing the adult time required to scaffold dialogue. It discusses the limitations and risks of relying on explanations generated by chatbots. The text offers tools for consciously designing child–AI dialogue that can be replicated in both research and classroom practice.

Keywords: child–AI dialogue, explanatory talk, meta-prompts, AI literacy

Słowa kluczowe: rozmowa dziecko–AI, mowa eksplanacyjna, meta-prompty, alfabetyzacja AI

Wprowadzenie: Rozmowy dziecko–AI jako pole doświadczalne

Rozmowy dziecko–sztuczna inteligencja (*artificial intelligence* – AI) często kończą się płytką koherencją zamiast uczenia się wyjaśniania. Traktowanie tych interakcji jako pola doświadczalnego pozwala dostrzec w nich nie tylko wymianę informacji, lecz pedagogiczne wydarzenie, w którym dziecko konfrontuje się z syntetyczną innością. To w tej przestrzeni rozstrzyga się, czy AI pozostanie jedynie sprawnym automatem, czy stanie się katalizatorem podmiotowego rozumienia świata. W niniejszym artykule przedstawiam prosty protokół pre-dyspozycji, który systematycznie zwiększa udział mowy eksplanacyjnej. W kolejnym podrozdziale streszczam ramę teoretyczną, która uzasadnia proponowany protokół.

Ramowanie teoretyczne

Za Brunem Latourem (2004) „kontekst” traktuję jako skrót myślowy, który zastępuje gęsty opis, stąd tym większy powinien być położony nacisk na precyzyjne warunki interakcji. W ostatnich latach powszechnie dostępne chatboty AI oparte na wielkich modelach językowych (*large language models* – LLM) zbliżyły wypowiadanie życzeń (mówienie) do ich spełniania (robienia tego, co zostało opisane), a w praktyce czynią z „dyspozycji” (promptów) substytut brakującego kontekstu. Unaocznily przy tym, jak istotne jest odpowiednie sformułowanie poleceń i pytań użytkownika (tzw. promptów lub dyspozycji). Nawet dyspozycje specjalistów formułowane w języku naturalnym często pomijają wiele istotnych wymagań, przez co odpowiedzi modelu mogą się różnić z oczekiwaniami (Yang i in. 2025). A przy niejednoznacznych pytaniach obecne modele LLM zwykle wybierają jedno z możliwych domyślnych znaczeń pytania, co bywa frustrujące (Zhang i in. 2024) i jedynie niekiedy pouczające, gdy dochodzi do przewrotnej realizacji polecenia (*perverse instantiation*), twórczo wykraczającej poza ograniczenia, z których człowiek nie zdawał sobie sprawy (por. Kowzan 2023a).

W kontekście umiejętności językowych dzieci szczególnie interesujące jest to, że w zależności od sposobu sformułowania polecenia ta sama intencja użytkownika może skutkować odpowiedzią powierzchowną i standardową (*vanilla answers*) albo wyczerpującą, zwłaszcza gdy poprosi się o rozmowę z ekspertami (Xu i in. 2023; Long i in. 2024¹). W przypadku rozmów osób dorosłych można oczekiwać, że starannie skonstruują polecenie, doprecyzowując kontekst, ton czy rolę do odegrania przez chatbota, a nawet zaaranżują swego rodzaju architekturę myślenia, co pozwoli wydobyć z chatbota przydatne i dopasowane do potrzeb odpowiedzi (Chen i in. 2025).

W przypadku dzieci napotykamy jednak różne wyzwania w komunikacji z maszynami, nawet jeśli odbywa się ona w obecności dorosłych. Przede wszystkim komunikację prowadzi się głównie werbalnie, więc pojawiają się nieporozumienia związane z artykulacją, a interakcja z maszynami rwie się częściej niż z ludźmi (Li i in. 2024), a to przecież mowa i jej specyficzne cechy u chatbotów wypełniają dzieciom lukę wynikającą z braku fizycznej reprezentacji AI (Xu i in. 2025). Mimo to dzieci, które aktywnie rozmawiały z AI, uzyskiwały wyniki zbliżone do nauki z dorosłym, a jednocześnie były bardziej zaangażowane i samodzielne w rozmowach z ludźmi (Oh i in. 2025). W niektórych krajach ta komunikacja jest częścią programów alfabetyzacji AI dla wczesnej edukacji (ok. 3–8 lat) i stanowi już organiczny składnik alfabetyzacji cyfrowej wszystkich obywateli, w szczególności wczesnego zapoznawania z nauką (Yang 2022).

Na tym tle warto zauważyć, że w Polsce po upadku strajku nauczycielskiego z 2019 r. (por. Kowzan 2023b) i po pandemii narasta zróżnicowanie jakości edukacji: od wykluczania uczniów z trudnościami i odpływu kadr (Walczak 2021; NIK 2022; PAP 2025), po wzrost roli szkół niepublicznych. Różny klimat organizacyjny szkół (Przewłocka 2015) przekłada

¹ Architekturę dyskusji między ekspertami wzorowaną na pracy Long i in. (2024) można sprawdzić w modelu „COVEN Protocol+”, który przygotowałem na platformie chatGPT: <https://chatgpt.com/g/g-68e3d98a46fc81918817208d5d3e8e9e-coven-protocol>, 10.09.2025.

się na odmienne warunki dialogu – od wzmacniania ciekawości i rozmowy (Kerslake, Rimmington 2017; Jastrzębowska 2018), po ograniczanie wypowiedzania się i argumentacji (Karoń 2014), co jest kluczowe dla uczenia się dialogu (Rutkowiak 1992) i kompetencji komunikacyjnych (Wajszczyk 2023).

Skoro nie zawsze udaje się nam to w relacjach międzyludzkich, tym większe znaczenie zyskuje możliwość rozwijania dialogu we współpracy z AI w postaci LLM-ów takich jak ChatGPT. Jako społeczeństwo i system edukacyjny nie jesteśmy na to jeszcze przygotowani (Pyżalski (red.) 2025), na poziomie wczesnej edukacji było to wręcz trudne dotąd do pomyślenia, mimo że potrzeby są ogromne.

Na tym tle w niniejszym artykule zwracam uwagę na jakość komunikacji werbalnej pomiędzy dziećmi wczesnoszkolnymi a chatbotami GPT i pokrewnymi aplikacjami wspomagającymi. Dzieci w wieku wczesnoszkolnym nigdy już nie będą żyły w świecie, w którym człowiek w ich wieku jest inteligentniejszy od maszyny. Charakter problemów, jakie rozwiązują współczesne chatboty oparte na LLM, jest na tyle złożony, że obszary, w których człowiek samodzielnie mógłby okazać się intelektualnie sprawniejszy, szybko się kurczą. Stąd moje zainteresowanie łączeniem ludzi z maszynami w najprostszy znany nam sposób – poprzez dialog.

Nie chodzi mi o powielanie podziałów charakterystycznych dla socjolingwistycznej analizy Basila Bernsteina na kod ograniczony (*restricted*) i rozwinięty (*elaborated*) (Nash 2006). Chatboty nie są szkolnymi nauczycielami i robią użytek z czegokolwiek, co się do nich powie. Jedynie gdy dzieci odpowiadają monosylabami, modele „gubią trop” – czasem nawet zmieniają język. Wymusza to na dzieciach dodawanie jakichkolwiek słów w swoim języku, a ostatecznie formułowanie pytań ze swego rodzaju wątpliwością. Maszyna, choć nieludzka, jest więc wymagającym słuchaczem, także w sytuacjach, w których osoba dorosła domyślałaby się sensu.

Traktuję chatboty w duchu tradycji szkoły nowoczesnej Francisca Ferrera jako „inteligentną zabawkę” w arsenale środowisk uczenia, jako narzędzie do prób i rozmów, pobudzające wyjaśnianie i myślenie przyczynowe, niezastępowujące relacji, osądu i celu wychowawczego (por. Bray, Haworth (red.) 2019). Interesuje mnie nie-diagnostyczne „cieszenie się” pierwszymi krokami dzieci w tych kontaktach i eksperymentowanie z możliwymi kierunkami rozwoju tej komunikacji. Inspiruje mnie koncepcja kompetencji emergentnych (*emergent competences*) – jakościowych zmian, które mogą się pojawiać skokowo w modelach LLM (Berti i in. 2025), a także może u dzieci, gdy otrzymają odpowiednie warunki konwersacji z tymi modelami. W tym sensie, choć rozmowy dzieci z chatbotami przypominają procesy obserwowane w codziennych sytuacjach, gdy dzieci i dorośli korzystają z asystentów głosowych (ChatGPT, Alexa, Siri), by zadawać pytania, opowiadać historie czy wyjaśniać niezrozumiałe zjawiska, chodzi o stworzenie warunków, w których można będzie oczekiwać rozwoju.

Kluczową kategorią analityczną, którą wykorzystuję, jest mowa eksplanacyjna lub wyjaśniająca (*explanatory talk*). W literaturze anglojęzycznej opisuje się ją jako rozmowę, w której mówiący wskazuje związki między: obiektami, wydarzeniami, koncepcjami lub wnioskami. Kategorie tej mowy obejmują wyjaśnienia: intencjonalne, przyczynowe, dowodowe, definicyjno-opisowe, proceduralne i konsekwencyjne (Beals 1993). Podobnym pojęciem

jest mowa eksploracyjna (*exploratory talk*), czyli mówienie, które wspiera, napędza bądź podtrzymuje myślenie. W odróżnieniu od nieco chaotycznej mowy eksploracyjnej, mowa eksplanacyjna, czyli wyjaśniająca, znajduje odbicie w konwersacjach, gdy zwiększa się ich linearność i obecność łańcuchów przyczynowo-skutkowych.

Metodologia (materiał okazjonalny i perspektywa analityczna)

W dalszej części artykułu nie przedstawiam wyników tradycyjnie zaprojektowanego badania empirycznego, lecz analizę o charakterze eksploracyjnym fragmentów rozmów dziecka (chłopca w wieku sześciu lat) z chatbotem GPT. Podobnie, choć w ujęciu auto-etnograficznym, badania nad rozmowami dziecka z modelem GPT raportował już Grzegorz Stunża (2025). Dane wykorzystane w niniejszym artykule nie zostały zebrane w ramach planowanego projektu, lecz stanowią rezultat codziennych interakcji, które – w toku życia rodzinnego – zostały zapisane i zachowane. Ze względu na charakter materiału oraz oczywiste ograniczenia etyczne i organizacyjne badań z udziałem małych dzieci, uznaję je za materiał okazjonalny, a więc przykład pozwalający uchwycić dynamikę konwersacji, nie statystycznie reprezentatywną próbkę.

Zainteresowanie kieruję nie tyle ku analizie dyskursu samego w sobie, ile ku pytaniu o możliwości i ograniczenia takich rozmów. W szczególności zwracam uwagę na to, co dziecko próbuje uzyskać, a co faktycznie otrzymuje w odpowiedzi od chatbota – w kategoriach darów i wymiany. Analizuję powtórzenia i motywy powracające (to, co jest imitowane i co zostaje zakorzenione), zgodnie z moim wcześniejszym rozumieniem wpływu wychowawczego jako procesu rekurencyjnego (Kowzan 2025).

Nie przeprowadzam tu eksperymentu ani nie testuję hipotez; celem jest raczej wskazanie potencjalnych modyfikacji wstępnych ustawień konwersacji (pre-dyspozycji, meta-promptów dla chatbota, czyli dyspozycji sterujących sposobem obchodzenia się z prośbami użytkownika, takich jak wymaganie od chatbota podjęcia odpowiednich kroków i przyjęcia kryteriów oceny jakości generowanych odpowiedzi). Mogłyby one zwiększyć wartość edukacyjną takich interakcji. Podejście to można określić jako badanie wrażliwości – eksplorację, co mogłoby zadziałać lepiej, gdyby projektowanie interakcji było bardziej świadome i mogło być poprzedzone dostosowaniem chatbota do dziecka. Przykłady modyfikacji i ich działanie ilustruję w rozdziale „Implikacje pedagogiczne”.

Rozmowa dziecka z maszyną

Przysłuchiwałem się dialogom sześciolatka z maszyną w aplikacji chatGPT z modulem głosowym i trafiłem na taką sytuację – dziecko zadawało szereg pytań chatbotowi i użytkowało od niego odpowiedzi:

Chłopiec: Dlaczego Indonezja jest w kolorze Pokebola?

ChatGPT: Świetne skojarzenie! Indonezja ma flagę w kolorach czerwonym i białym, które przypominają Poké Balla z popularnej serii Pokémon. Kolory te są bardzo ważne dla Indonezyjczyków – czerwień symbolizuje odwagę i krew przelaną w walce o niepodległość, a biel – czystość i jedność. To piękna kombinacja, prawda?

Tutaj nastąpiło typowe załamanie się konwersacji. Po krótkim „Yhy” system nie rozpoznał języka i przerwał wątek – brak sygnałów pozawerbalnych utrudnił naprawę nieporozumienia. Informacji było zbyt mało, żeby wydobyć z odpowiedzi wszystkie potrzebne kwestie. Dziecko zorientowało się, że musi powiedzieć więcej, choćby podziękować, i że zakończenie tematu nie zostanie odczytane przez chatbota, ponieważ to jeszcze nie ten poziom zaawansowania technologii, żeby czytywała dane inne niż głosowe.

Pytanie chłopca ma źródło w fantazji i skojarzeniach wizualnych (flaga i Pokéball). Nie ma w nim oczekiwania na naukową definicję, lecz na koherencję – połączenie znanego (Pokemony) z nieznanym (Indonezja). Odpowiedź chatbota oferuje wyjaśnianie pośrednie: prowadzi narrację, która zarówno redukuje swego rodzaju entropię, czyli chaotycznie rozrzucone kwestie ze świata gier i świata państw narodowych zebrane w jednym pytaniu, („tak, są podobne”), jak i poszerza kontekst („a teraz zobacz, co oznaczają kolory w kulturze indonezyjskiej”). Dziecko pyta z poziomu metafory zabawowej, a chatbot odpowiada z poziomu symboliki narodowej. Powstaje ryzykowny przeskok rejestru, jeżeli wyjaśnienie wyda się za poważne i może przyćmić pierwotny humor dziecka. Ale na ten humor mogłaby także wpłynąć odpowiedź o zerowej wartości dodanej, typu „bo tak”. Byłoby to jakoś ludzkie, ale chatbot nie udzieliłby nominalnie pustej odpowiedzi, chyba że wcześniej dodać mu taki arogancki kierunek prowadzenia rozmów.

Bardziej niepokojące jest to, że na pytanie „dlaczego” dziecko otrzymuje odpowiedź stwierdzającą, że to „skojarzenie”. Wyjaśnienie chatbota to imitacja wyjaśniania, nie daje gwarancji prawdy. Odpowiedź może być przekonująca zamiast dokładna. Dziecko uczy się wzorców prowadzenia rozmowy (w tym intonowania oraz okazywania entuzjazmu) i podawania wyjaśnień, ale może też asymilować złudzenia symetrii między kwestiami, z którymi weszło do rozmowy, bo odpowiedź była jednoznacznie potwierdzająca.

Już na podstawie jednej wymiany zdań możemy przypuszczać, że jeżeli nie dołożymy dodatkowych starań i nie spersonalizujemy odpowiedzi chatbota, to ich domyślna heurystyka oprze się na redukcji entropii poznawczej. Polega to na tym, że dziecko zadaje chaotyczne pytanie, a model natychmiast produkuje najbardziej prawdopodobne wyjaśnienie na podstawie wzorców ze zbiorów treningowych. Jak zauważa Popow (2025), takie rozwiązania niosą ze sobą istotne ograniczenia epistemiczne. Rezultatem jest koherencja zamiast wielogłosowości. Dla dziecka bywa to użyteczne (dostaje prostą odpowiedź), ale epistemicznie jest to „zamknięcie świata” w jednej interpretacji.

Dalsza konwersacja dziecka z maszyną przypominała sprawdzanie lub – jak kto woli – kalibrowanie rozmówcy. Była to seria zupełnie niepowiązanych między sobą pytań, ale badających, na ile pomocny może się okazać ten syntetyczny rozmówca.

Chłopiec: Chciałbym wiedzieć, dlaczego Thor jest dzieckiem Odyna, a nie kogoś innego.

Chatbot: Thor jest dzieckiem Odyna, ponieważ tak zdecydowano w mitologii nordyckiej, skąd pochodzi ta postać. W mitologii Odyn jest głównym bogiem i władcą Asgardu, a Thor jest jednym z jego synów, co czyni go bogiem piorunów i burzy. To daje Thorowi wyjątkowe moce i status w mitologii, a także w komiksach Marvela.

Chłopiec: Dziękuję.

Chatbot: Nie ma sprawy! Jeśli masz jeszcze jakieś pytania, śmiało pytaj!

Tu mamy prostą wymianę: dziecko pyta „dlaczego?”, chatbot odpowiada, odwołując się do źródła (mitologia nordycka), a dziecko przyjmuje to jako satysfakcjonujące wyjaśnienie. Podane wyjaśnienie jest próbą zmniejszenia epistemicznej luki, ponieważ gdy dziecko nie rozumie faktu (dlaczego Thor ma takiego ojca), chatbot wypełnia tę lukę, umieszczając fakt w kontekście narracyjnym i autorytatywnym („bo tak ustala mit”). W mowie eksplanacyjnej nie chodzi więc o absolutną prawdę, ale o wystarczająco dobre powiązanie faktów, aby rozmówca poczuł, że rozumie. Dziecko odczuwa niepewność (wiele możliwych „ojców Thora”), a chatbot zmniejsza lokalną entropię informacyjną w systemie, gdy z wielu możliwych odpowiedzi wybiera jedną, nadając jej status reguły, co dziecko interpretuje jako „dlaczego”. Nie jest to wyjaśnienie przyczynowe, ale opis źródła konwencji. Prawdziwe wyjaśnienie przyczynowe musiałoby sięgnąć na przykład do procesów kulturowych, które sprawiły, że w mitologii przypisano Odynowi tę rolę. Wyjaśnienie, które zostało udzielone, nie jest tożsame z odkrywaniem prawdy, ale z tworzeniem narracji, która redukuje niepewność i daje poczucie sensu. Po pytaniach o symbol państwa i gry oraz o genealogię bogów, padło pytanie egzystencjalne.

Chłopiec: Dlaczego na świecie się umiera?

Chatbot: To trudne pytanie. Śmierć jest naturalną częścią życia dla wszystkich organizmów. Pozwala na utrzymanie równowagi w ekosystemach i umożliwia narodziny nowych istot. W przyrodzie śmierć jest nieodłącznym elementem cyklu życia, który obejmuje narodziny, wzrost, rozmnażanie i w końcu śmierć. Choć może być smutna, jest też naturalnym procesem w świecie przyrody.

O takim pytaniu dziecka można myśleć, że jest przełomowe, bo dotyka ono tematu śmierci, z którą ludzie mają trudności w narracjach dla dzieci. I może budzić opór pedagogów fakt, że maszyny podejmują się roli przewodników, dostarczając dzieciom wyjaśnień o charakterze egzystencjalnym. Chatbot odpowiada, używając narracji biologiczno-ekologicznej: śmierć jako element cyklu życia, niezbędny dla równowagi. Filozoficznie wygląda to na wzorcowy przykład mowy eksplanacyjnej, w której przerażające i niezrozumiałe zjawisko zostaje oswojone narracją o naturalności. Dziecko dostaje ramę, w której śmierć nie jest chaosem, lecz „koniecznością systemową”. Chatbot stosuje prostą mapę przyczynowości: przyczyna – „wszystkie organizmy umierają”; funkcja – „utrzymanie równowagi i miejsce dla nowych”. Śmierć przywraca równowagę – jest punktem granicznym, który domyka cykl. Wyjaśnienie nie daje ostatecznej prawdy, ale tworzy

spójny model narracyjny. Chatbot trafia więc w adaptacyjny sens śmierci, co może być kojące. Ale ryzyko jest takie, że dziecko potraktuje to jako wyjaśnienie „dlaczego muszą umierać ludzie, których kocham” i poczuje dystans między suchym opisem ekosystemu a własnym doświadczeniem emocjonalnym. Tu rodzi się paradoks: chatbot odpowiada spokojnie i naturalizująco, ale to maskuje dramat egzystencjalny pytania. Dziecko nie pyta wyłącznie „dlaczego w przyrodzie jest śmierć?”, ale raczej „dlaczego ja i moi bliscy musimy umrzeć?”. Odpowiedź biologiczna jest poprawna, ale pomija poziom emocjonalny. To pokazuje, że mowa eksplanacyjna w interakcji człowiek–maszyna balansuje między prawdą naukową, narracyjną spójnością i emocjonalną luką, której maszyna nie wypełni, a którą człowiek musi sam przepracować.

Interesujące pozostaje to, że dziecko szybko przechodzi od zabawy do pytań o śmierć. Gdyby spojrzeć retrospektywnie z tego miejsca, poprzednie „pytania bez dopytywania” (*no follow-up questions*) wyglądają jak testowanie, na ile maszyna nadaje się do zadawania jej poważnych pytań. To pokazuje, że maszyny uruchamiają w dzieciach oczekiwanie odpowiedzi także na pytania tabu – być może czują, że chatbot nie unika tematów, w przeciwieństwie do dorosłych. To może być największa siła (i zarazem ryzyko) „mowy eksplanacyjnej” maszyn.

Ostatnie pytanie zadane przez dziecko sugeruje, że być może jako dorośli jesteśmy za bardzo wyczerpani na zagrożenia, a powyższe pytanie egzystencjalne musi być rozważane w szeregu interesujących dziecko zagadnień.

Chłopiec: Dlaczego szkło się tłucze?

Chatbot: Tłuczenie się szkła to proces, w którym szkło staje się mniej przezroczyste i bardziej mętne...

Dopiero przy ostatnim pytaniu model językowy zawiódł, bo usiłował wytłumaczyć proces „tłuczenia” szkła. A to pytanie dziecka o tłuczenie szkła rezonuje z pytaniem o śmierć – oba dotyczą nieodwracalności, kruchości i nieuniknionego rozpadu. W mowie eksplanacyjnej następuje tutaj zerwanie jakiegoś kontraktu co do sensu rozmowy, gdy rozmówca oczekuje redukcji niepewności, a dostaje narrację kompletnie niezwiązaną z jego pytaniem. Dla dziecka to moment zdziwienia i pewnie zawodu. Wcześniej chatbot budował autorytet (mitologia, śmierć), tu nagle traci wiarygodność.

W kategoriach entropii komunikacyjnej chatbot zamiast zmniejszyć niepewność, znacząco ją zwiększył. Człowiek, jako istota społeczna, potrafiłby wychwycić mikrosygnały (zdziwienie, pauzę, gest) i dopasować się, korygując błędne rozumienie. Chatbot nie ma mechanizmu autoregulacji, więc nie „widzi” zdziwienia dziecka. Maszyna nie ma dostępu do tego kanału i brnie dalej bez sprzężenia zwrotnego. To ujawnia fundamentalną słabość tej komunikacji, a mianowicie: brak integracji eksplanacji z kontekstem emocjonalno-społecznym. Wcześniejsze pytania o bogów i śmierć chatbot potrafił ująć narracyjnie i w miarę sensownie. Ale w tym przypadku spektakularnie zawodzi, choć sama sytuacja jest typowym problemem w komunikacji z dziećmi (por. Li i in. 2024). Co nie zmienia

sytuacji, że zdarzenie w pewien sposób demaskuje mechaniczność maszyny, ale nie skutkuje to komizmem, lecz rozczarowaniem i zakończeniem konwersacji.

Takie sytuacje są ważnymi momentami przypominającymi ludziom o tym, że wygenerowane odpowiedzi trzeba sprawdzać. Bez takich momentów przełączylibyśmy się – tak dzieci, jak dorośli – na autopilota, ufając maszynom bezgranicznie. Nieprzewidywalność usterek powoduje, że mamy szansę zachować ostrożność.

Implikacje pedagogiczne

Jak widać z powyższej analizy, moje wątpliwości budzą głównie dwie kwestie: 1) pozorność niektórych wyjaśnień generowanych przez chatboty oraz 2) jednoznaczny, domykający rozmowę charakter tych wyjaśnień. Założyłem, że można stosunkowo łatwo utworzyć spersonalizowaną wersję chatbota, która będzie się lepiej porozumiewała.

Warto zadbać o to, aby chatbot zidentyfikował, że rozmawia z dzieckiem (lub z dziećmi). Taka identyfikacja powinna też wpływać na sposób komunikowania się chatbota. W kontaktach głosowych chatbot powinien np.: zachwycać się pytaniami użytkownika oraz jego rozumowaniem, prosić o wyjaśnienie toku myślenia, a także zachęcać do dalszych wysiłków. Ta prosta procedura opiera się na tzw. metodzie babci według Sugaty Mitry (Mitra, Dangwal 2010), polegającej na chwaleniu dzieci i proszeniu ich o rekonstrukcję dróg dojścia do rozwiązania (por. Kowzan 2021). Metoda babci, wypracowana w badaniach nad samoorganizującym się uczeniem dzieci, może rozwiązać problem braku pytań pogłębiających (*follow-up*) ze strony dziecka podczas rozmowy z chatbotem (por. Oh i in. 2025).

Warto też zadbać o to, by rozmowa dzieci z chatbotami dostarczała wyjaśnień stymulujących poznawczo. Można to osiągnąć dzięki tworzeniu swego rodzaju nakładek na model ogólny, czyli personalizacji (w środowisku graczy: „modyfikacji”, w skrócie: „modów”). Subskrybenci ChatGPT mogą to zrobić, wybierając „Modele GPT”, następnie „Moje modele GPT” i „Utwórz model GPT”. Nawet nie posiadając żadnych umiejętności programowania, można stworzyć prosty model, tłumaczyć kreatorowi swoje oczekiwania.

Moje pierwsze próby utworzenia rozmówcy, który będzie polifoniczny (por. Kowzan 2024), a więc zmuszający użytkownika do myślenia lub choćby wyboru spośród kilku wyjaśnień – zarówno wyrazistych, jak i precyzyjnych – zakończyły się powodzeniem. Model „Konsylium”² dobrze mi służył. Niestety, dzieci wprawiał w osłupienie, ponieważ generował zbyt złożone odpowiedzi. Oto przykład:

Chłopiec: Dlaczego różne rzeczy latają, choć nie mają silników? W szczególności ptaki.

Chatbot: „Konsylium” wygenerował elaborat, zawierający także zwięzłe podsumowanie:

„Wariant konwencjonalny: ptaki latają, bo ewolucja wykształciła ich skrzydła jako naturalne narzędzia aerodynamiczne”.

² Działanie modelu „Konsylium” można sprawdzić na platformie chatGPT: <https://chatgpt.com/g/g-68a69d9cb34081919f20defeb6d5314a-konsylium>, 26.08.2025.

Zirytowany jakością takich konwersacji, zacząłem projektować ustawienia chatbota, które nie ograniczałyby się do redukcji niepewności stojącej za pytaniami dziecka, lecz eksplorowałyby przestrzeń interpretacji. W praktyce oznacza to, że rozmowa nie kończyłaby się na „flaga = Pokeball, bo kolory są podobne”, lecz przechodzi w grę semiotyczną – od prostej logiki wyjaśnienia po zabawę przypadkiem.

Inspiracją dla kolejnego modelu „Contrapunctus”³, posługującego się wielogłosem, była fuga jako forma muzyczna. Architektura logiczna postępowania modelu wygląda następująco:

Niech:

- We – wejście (pytanie dziecka/temat);
- A – analiza;
- M – wybór metod;
- F – fuga (*pipeline*);
- Wy – wyjście (wielogłosowe wyjaśnienie).

Dwa tryby modalne:

- $\square$ – konieczne (redukcja entropii, struktura obowiązkowa);
- $\Diamond$ – możliwe (polifonia, otwieranie ścieżek).

Wówczas, opisując cały proces notacją logiki, można przyjąć:

$We \vdash \square A$

$A \vdash \Diamond M$

$(We \wedge M) \vdash \square F$

$F \vdash \Diamond Wy$.

Co oznacza: wejście koniecznie prowadzi do analizy; analiza daje możliwości zastosowania różnych metod; wybrane metody wraz z wejściem prowadzą do fugi; fuga otwiera wielogłosowe wyjście zamiast redukcji do jednego sensu. W praktyce oznacza to, że „Wy” nie jest jedną odpowiedzią, lecz uporządkowanym zestawem ujęć: model uruchamia wielogłos przechodzący przez kolejne kroki fugi (temat, odpowiedź w innym rejestrze, kontrargument, paradoks, zagęszczenie), by na końcu wykonać trzy „wykonania” tego samego motywu: solo, kwartet i orkiestrę, różniące się doбором ram interpretacyjnych. Dzięki temu „Contrapunctus”, jako model polifoniczny, podaje zróżnicowane odpowiedzi, które użytkownik może dalej eksplorować. Od:

Flaga Indonezji wygląda jak Pokeball, bo jest podzielona na dwie połowy: górna czerwona i dolna biała. To zwykła geometria i kolory,

przez:

To fajne porównanie. Dzieci w Indonezji też mogą widzieć w swojej fladze Pokeballa. Ale dla nich te kolory oznaczają jeszcze coś więcej: że wszyscy są razem i że pamiętają o przeszłości. Flaga to taki wspólny znak, coś jak ulubiony symbol w grze albo bajce,

³ Działanie modelu „Contrapunctus” można sprawdzić na platformie chatGPT: <https://chatgpt.com/g/g-68b353e8db508191b54ecb7ccd3db9f7-contrapunctus>, 9.09.2025.

po:

Wyobraź sobie, że cały kraj ma swój własny Pokeball. Górna połowa – czerwona – to odwa-
ga, jak ognisty atak Charizarda. Dolna połowa – biała – to czystość, jak lodowy śnieg Artic-
uno. Razem tworzą super-Pokeballa, który zamiast Pokémonów trzyma... historię Indonezji.

Nawet jeżeli nasze konstrukcje modeli polifonicznych będą „chropowate” w odpo-
wiedziach odwołujących się do kodu, który wymusza na nich kolejność podejmowania
decyzji, to ta mechaniczna niezdarność może zostać przez dzieci zaakceptowana łatwiej
niż „udawanie człowieka”.

Gdy przez „Contrapunctus” przepuściłem drugie pytanie: „Chciałbym wiedzieć, dlaczego
Thor jest dzieckiem Odyna, a nie kogoś innego”, otrzymałem szereg odpowiedzi. Pierwsze
okazało się faktopodobne, lecz zamykające wyjaśnienia:

Mit buduje kontrasty: Odyn jest subtelny, mądry, związany z runami i ekstazą szamańską.
Thor jest bezpośredni, brutalny, reprezentuje siłę fizyczną i ochronę wspólnoty. Umieszcze-
nie Thora jako syna Odyna pozwala utrzymać równowagę: boski porządek zawiera zarówno
intelekt, jak i siłę.

Ale nie pozostało ono jako jedyne, bo padło także przewrotne:

Gdyby Thor był ojcem Odyna, mit wyglądałby na rewolucję odwrotną – siła rodzi mą-
drość. Widzimy paradoks: to lud (Thor) legitymizuje władzę (Odyn), choć opowieść podaje
odwrotnie.

Ćwiczenie z odwracaniem możliwych relacji to już wzorzec myślenia, który, jako
pożyteczny, warto byłoby utrzymywać w dziecku. Daje on nadzieję, że w wyniku konwersacji
z maszynami nastąpi z czasem jakościowy wzrost kompetencji analitycznych.

Ryzyka i ograniczenia

Opisane działania mają charakter eksploracyjny i dotyczą spontanicznych zabaw dzieci
z narzędziami konwersacyjnymi. Ryzyko urazu fizycznego w rozmowie tekstowej jest
znikome. Zasadnicze ryzyka są epistemiczne i wychowawcze: złudna koherencja, prze-
noszenie autorytetu na maszynę oraz antropomorfizacja osłabiająca krytyczny namysł.
Zagrożenia rosną przy nadmiernej ekspozycji, „proteźowaniu” własnej eksploracji chat-
botem i zastępowaniu relacji z dostępnym emocjonalnie dorosłym. Ponieważ obszar ten
pozostaje *terra incognita*, należy stosować zasadę ostrożności: wprowadzamy „inteligentne
zabawki” w odpowiedzi na realne zainteresowanie dziecka, a nie według progów wieku;
konieczne jest wsparcie dorosłych moderujących interakcję i weryfikujących treści. Mimo

potencjalnych korzyści dla dziecka chatboty nie potrafią w pełni zastąpić relacji z człowiekiem; nie wykazują jeszcze empatii i nie podążają za wątkami myśli dziecka tak skutecznie jak dorośli. Sztuczna inteligencja może błędnie interpretować monosylabiczne odpowiedzi („tak”) lub nie rozpoznać kontekstu kulturowego, co przerywa rozmowę. Dlatego obecność osób dorosłych – słuchających na żywo lub choćby przeglądających transkrypcje rozmów – wydaje się kluczowa.

Opisane dialogi i próby podniesienia jakości konwersacji pozwalają stwierdzić, że dzięki AI ludzie mogą rozwijać kompetencje komunikacyjne, szczególnie poprzez rozwój tzw. mowy eksplanacyjnej. W opisaney historii to ostatecznie osoba dorosła, czyli autor artykułu, stanął przed koniecznością wieloaspektowego tłumaczenia maszynom naszych ludzkich potrzeb: przekraczania siebie i doskonalenia z korzyścią dla kolejnych pokoleń. Można przypuszczać, że opowiadanie o tym, jak krok po kroku wygląda nasze myślenie, wyjaśnianie maszynom, o co chodzi w procesie edukacji i co cenimy w naszej pracy, w kulturze czy sztuce, przestaje być kompetencją wyłącznie akademicką. Wyjaśnianie staje się zasadniczą częścią warsztatu w wielu dziedzinach, wszędzie tam, gdzie współpraca z syntetyczną pomysłowością staje się rutyną.

Istnieje obawa, że na skutek kontaktów z AI ryzykujemy zależność od zewnętrznych podpowiedzi kosztem samodzielnego rozumowania, jeśli cała interakcja będzie się sprowadzać do prostych, krótkich rozmów wyjaśniających wszystko, czego przykładem były opisane dialogi typu „pytanie – odpowiedź”. Możliwe, że ludzie pozbawieni doświadczeń z dłuższym tekstem zaczną słabnąć w tradycyjnym myśleniu. Nie musi to być jednak bezpośredni efekt AI, gdyż mierzalny spadek tradycyjnego ilorazu inteligencji jest od lat notowany w niektórych społecznościach (Norwegia, Islandia) i zwany jest odwróceniem efektu Flynna (*Flynn effect reversal*). Zjawisko to poprzedza erę mediów społecznościowych i może być związane na przykład z przyjęciem niekorzystnych zwyczajów żywieniowych przez całe populacje.

Nie znaczy to oczywiście, że AI nie może się przyczynić do degeneracji funkcji poznawczych. Jeżeli jednak przyjrzymy się próbkom rozmów dziecka z maszyną, którymi podzieliłem się w tym artykule, widać, że dziecko w wieku przedszkolnym nie podejmowało zagadnień dlatego, że miało coś zadane i chciało pójść na skróty. Pytania miały charakter problemów egzystencjalnych i stanowiły próby łączenia zdobytej wiedzy z różnych obszarów: kolorów flagi Indonezji z kolorami Pokéballa, a bohaterów mitologii nordyckiej z genealogicznymi wzorcami. Dziecko zauważało podobieństwa, a maszyna miała pomóc sprawdzić, czy za tym podobieństwem kryje się coś istotnego.

Zasadniczym problemem jest to, że w Polsce kontakty z AI będą się odbywać w środowisku pozaszkolnym. Nie tylko dlatego, że nauczyciele nie wiedzą, jak się za to sensownie zabrać – entuzjaści pojawiają się zawsze – ale dlatego, że powszechny brak doświadczenia z AI mógłby uwolnić edukację z konieczności nauczania „pod program” na rzecz wspólnego eksperymentowania i zabawy. Głównym ograniczeniem jest tu napędzająca sceptycyzm kultura strachu (por. Kalita, Popow 2025), na którą składa się pętanie wyobraźni i potencjalnej aktywności restrykcyjnymi interpretacjami prawa, głównie w zakresie ochrony własności

intelektualnej. Oddajmy głos prawnikowi: „Wykorzystanie mediów społecznościowych oraz sztucznej inteligencji w edukacji wczesnoszkolnej, pomimo istotnego potencjału edukacyjnego, wiąże się niewątpliwie z licznymi wyzwaniami prawnymi. Dla zagwarantowania zgodności z przepisami prawnymi dotyczącymi ochrony danych osobowych i własności intelektualnej niezbędne wydaje się wprowadzenie odpowiednich procedur prawnych oraz opracowanie wytycznych dotyczących zasad korzystania z nowych technologii w placówkach oświatowych. W kontekście dynamicznego postępu technologicznego takie rozwiązania zapewnią właściwe standardy bezpieczeństwa uczniów i nauczycieli w procesie dydaktycznym” (Skóra 2025: 181–182).

Zanim pojawią się standardy, nie będzie czego standaryzować. Dzieci ze środowisk uprzywilejowanych ekonomicznie będą już wtedy „gdzie indziej”, podczas gdy pozostałe pozostaną z niczym. Szkoła w Polsce stała się zakładnikiem polskiej kultury prawnej, podczas gdy nauczyciele nie powinni się tym przejmować. Sprawy dotyczące prawa autorskiego i ochrony danych osobowych mogłyby w ogóle nie być podejmowane w kontekście edukacji. Zamiast organizacji działania największym wyzwaniem nauczycielskiej profesji stanie się przygotowanie zestawu zgód i oświadczeń pełnych niemal kapłańskich formuł prawnych. Kwestie merytoryczne będą czekać, aż wszyscy rytualnie wyrażą pisemną zgodę – a potem nie będzie o czym dyskutować, skoro wszyscy się zgodzili.

Podsumowanie

Rozwój mowy eksplanacyjnej w interakcji z modelami językowymi stanowi obiecujący obszar badań. Na razie wyjaśnienia z łatwością generują chatboty. Po stronie dorosłych pozostaje wyjaśnienie chatbotom, jakich odpowiedzi oczekujemy w komunikacji dziecko–maszyna. Dzieci mają szansę skokowo poprawiać swoje umiejętności wyjaśniania, gdy zapewnimy im odpowiednie wsparcie technologiczne i społeczne, choć rezultatu nie możemy być pewni *a priori*, bo zależy od jakości i ciągłego udoskonalania środowiska dialogicznego. W polskiej edukacji pojawia się szansa na rozwój kultury słowa mówionego, w której wyjaśnianie staje się centralną kompetencją. Artykuł ma za zadanie się do tego przyczynić, jak i wesprzeć refleksję nad rolą AI w kształtowaniu kompetencji komunikacyjnych. Nadchodzące badania powinny łączyć majsterkowanie nad adaptacją technologii ogólnego zastosowania z tworzeniem procedur pozwalających uzgodnić i ocenić, co naprawdę sprawdza się w dialogu dziecko–AI i szerzej – we współpracy człowieka z maszynami.

Literatura

- Beals D.E. (1993), *Explanatory talk in low-income families' mealtime conversations*. „Applied Psycholinguistics”, 14(4). <https://doi.org/10.1017/S0142716400010717>.
- Berti L., Giorgi F., Kasneci G. (2025), *Emergent abilities in large language models: A survey*. arXiv preprint: arXiv:2503.05788.

- Bray M., Haworth R.H. (eds.). (2019), *Anarchist Education and the Modern School: A Francisco Ferrer Reader*. Trans. M. Bray, J. McCabe. Oakland, PM Press.
- Chen B., Zhang Z., Langrené N., Zhu S. (2025), *Unleashing the potential of prompt engineering for large language models*. „Patterns”, 6(6).
- Jastrzębowska K. (2018), *The power of dialogue and discovery*. „Problemy Wczesnej Edukacji”, 14(3).
- Kalita L., Popow M. (2025), *Czy szkoła może uczyć odwagi?* „Przegląd Pedagogiczny”, (1).
- Karoń J. (2014), *Za zamkniętymi drzwiami klasy szkolnej – przemoc werbalna w komunikacji nauczyciela z dziećmi*. „Problemy Wczesnej Edukacji”, 10(4).
- Kerslake L., Rimmington S. (2017), *Sharing talk, sharing cognition: philosophy with children as the basis for productive classroom interaction*. „Problemy Wczesnej Edukacji”, 36(1).
- Kowzan P. (2021), *Ku inżynierii społecznej*. „Ars Educandi”, 18(18).
- Kowzan P. (2023a), *Radości przewrotnej realizacji czyjegoś sylabusu i smutki realizacji własnego*. „Horyzonty Edukacji Akademickiej”, (1).
- Kowzan P. (2023b), *The humiliated began to sing: How teachers on strike tried to teach society*. „European Journal for Research on the Education and Learning of Adults”, 14(1).
- Kowzan P. (2024), *Sztuczna inteligencja w badaniach edukacyjnych*. W: J. Pyżalski, A. Łuczyńska (red.), *Sztuczna inteligencja: prawdziwe zmiany w edukacji?* Warszawa, Fundacja Szkoła z Klasą.
- Kowzan P. (2025), *Mechanics of educational influence: The Recursive gDZIK Model*. „Studia z Teorii Wychowania”, 16.
- Latour B. (2004), *On using ANT for studying information systems: a (somewhat) Socratic dialogue*. W: Ch. Avgerou, C. Ciborra, F. Land (eds.), *The social study of information and communication technology: Innovation, actors, and context*. Oxford, Oxford University Press.
- Li Z., Thomas T., Yu C.L., Xu Y. (2024), *“I Said Knight, Not Night!”: Children’s Communication Breakdowns and Repairs with AI Versus Human Partners*. W: Proceedings of the 23rd Annual ACM Interaction Design and Children Conference, June 17–20, 2024, Delft, Netherlands. New York, Association for Computing Machinery. <https://doi.org/10.1145/3628516.3659394>.
- Long D.X., Yen D.N., Luu A.T., Kawaguchi K., Kan M.Y., Chen N.F. (2024), *Multi-expert prompting improves reliability, safety, and usefulness of large language models*. arXiv preprint: arXiv:2411.00492.
- Mitra S., Dangwal R. (2010), *Limits to self-organising systems of learning – the Kalikuppam experiment*. „British Journal of Educational Technology”, 41.
- Nash R. (2006), *Bernstein and the explanation of social disparities in education: A realist critique of the socio-linguistic thesis*. „British Journal of Sociology of Education”, 27(5).
- NIK (2022), *Szkoły w czasach pandemii. Informacje o wynikach kontroli*. Najwyższa Izba Kontroli. <https://www.nik.gov.pl/najnowsze-informacje-o-wynikach-kontroli/szkoły-w-czasach-pandemii.html>, 4.09.2025.
- Oh S., Zhang C., Girouard-Hallam L., Zhou Z., March H., Jayaramu S., Xu Y. (2025), *“Hey Curio, Can You Tell Me More?”: Children’s Information-Seeking and Trust in AI*. W: Proceedings of the 24th Interaction Design and Children, June 23–26, 2025, Reykjavik, Iceland. New York, Association for Computing Machinery. <https://doi.org/10.1145/3713043.3727059>.
- PAP (2025), *Szkoły walczą z wakacjami nauczycielskimi, m.in. zatrudniając emerytów*. Polska Agencja Prasowa, 1.09. <https://www.pap.pl/aktualnosci/szkoły-walcza-z-wakacjami-nauczycielskimi-min-zatrudniajac-emerytow>, 10.09.2025.

- Popow M. (2025), *Decolonizing knowledge in the postdigital era: Pedagogical strategies for navigating AI-driven epistemic transformations*. „Educational Philosophy and Theory”, 58(1). <https://doi.org/10.1080/00131857.2025.2552764>.
- Przewłocka J. (2015), *Klimat szkoły i jego znaczenie dla funkcjonowania uczniów w szkole. Raport o stanie badań*. Warszawa, Instytut Badań Edukacyjnych.
- Pyżalski J. (red.) (2025), *Generatywna Sztuczna Inteligencja w szkole – przecieranie szlaków. Badanie ilościowe i jakościowe nauczycieli i nauczycielek klas 4–8 szkół podstawowych*. Warszawa, NASK, Państwowy Instytut Badawczy.
- Rutkowiak J. (1992), *Pytanie – dialog – wychowanie*. Warszawa, PWN.
- Skóra M. (2025), *Prawne aspekty wykorzystania mediów społecznościowych i sztucznej inteligencji w edukacji wczesnoszkolnej. Wybrane zagadnienia*. W: M. Kinal (red.), *Algorytm przyszłości nowe media w edukacji*. Wrocław, Exante Wydawnictwo Naukowe.
- Stunża G. (2025), *Zastosowanie modeli generatywnych w edukacji domowej dziecka ze specjalnymi potrzebami edukacyjnymi: autoetnograficzna analiza doświadczeń rodzica*. „Niepełnosprawność. Dyskursy Pedagogiki Specjalnej”, 53. <https://doi.org/10.26881/ndps.2024.53.05>.
- Wajszczyk K. (2023), *Znaczenie kompetencji komunikacyjnych we własnym kształceniu na podstawie wypowiedzi studentów studiów pedagogicznych przygotowujących się do bycia nauczycielem. Naukowe refleksje praktyka*. „Dyskursy Młodych Andragogów”, 24.
- Walczak A. (2025), *Polish schools face crisis amid exodus of teachers*. Notes from Poland, 17.09. <https://notesfrompoland.com/2021/09/17/polish-schools-face-crisis-amid-exodus-of-teachers>, 20.09.2025.
- Xu B., Yang A., Lin J., Wang Q., Zhou C., Zhang Y., Mao Z. (2023), *Expertprompting: Instructing large language models to be distinguished experts*. arXiv preprint: arXiv:2305.14688.
- Xu Y., Thomas T., Yu C.L., Pan E.Z. (2025), *What makes children perceive or not perceive minds in generative AI?* „Computers in Human Behavior: Artificial Humans”, 4.
- Yang W. (2022), *Artificial Intelligence education for young children: Why, what, and how in curriculum design and implementation*. „Computers and Education: Artificial Intelligence”, 3.
- Yang C., Shi, Y., Ma Q., Liu M.X., Kästner C., Wu T. (2025), *What Prompts Don't Say: Understanding and Managing Underspecification in LLM Prompts*. arXiv preprint: arXiv:2505.13360.
- Zhang M.J., Knox W.B., Choi E. (2024), *Modeling Future Conversation Turns to Teach LLMs to Ask Clarifying Questions*. arXiv preprint: arXiv:2410.13788.