

Zasady wykorzystania sztucznej inteligencji w psychologii.

Omówienie dokumentu Global Psychology Alliance: *Top 10 Principles for Understanding Artificial Intelligence in Psychology*

Beata Krzywosz-Rynkiewicz¹

*Uniwersytet Warmińsko-Mazurski w Olsztynie
Katedra Psychologii Klinicznej, Rozwoju i Edukacji
<https://orcid.org/0000-0003-4306-4875>*

Krzysztof Mariusz Ciepliński

*Uniwersytet Vizja w Warszawie
Wydział Nauk o Człowieku
<https://orcid.org/0000-0003-1100-6419>*

W ostatnich latach pojawiło się wiele publikacji naukowych, w których wskazuje się na istotną rolę sztucznej inteligencji w psychologii i obszarze zdrowia psychicznego. Technologie bazujące na AI są m.in. wykorzystywane w promocji zmian zachowań zdrowotnych (Aggarwal i in., 2023), w diagnostyce, monitorowaniu i interwencjach w obszarze zdrowia psychicznego (Cruz-Gonzalez i in., 2025; Kolding i in., 2024). Dakanalis i in. (2024) określają sztuczną inteligencję jako „game-changer” w opiece psychiatrycznej. Ponadto AI znajduje zastosowanie w edukacji psychologicznej i psychiatrycznej (Prégent i in., 2025), może także przyczyniać się do zmniejszenia obciążenia klinicystów poprzez automatyzację procesu przygotowywania dokumentacji medycznej (Olson i in., 2025; McCrudden i in., 2026). Przedmiotem analiz bywa również relacja między narzędziami diagnostycznymi generowanymi przez AI a tradycyjnymi standaryzowanymi testami psychologicznymi (Kirthivasan i Bhardwaj, 2026).

¹ Adres do korespondencji: beata.rynkiewicz@uwm.edu.pl.

Wykorzystanie narzędzi bazujących na AI w codziennej działalności psychologów praktyków czy badaczy to zatem nie „pieśń przyszłości”, lecz kwestia niezwykle aktualna. Rozwój sztucznej inteligencji stawia jednak przed psychologią ważne pytania dotyczące odpowiedzialności zawodowej, etyki, trafności oddziaływań, bezpieczeństwa klientów oraz roli psychologa w świecie coraz bardziej nasyconym technologiami cyfrowymi.

Odpowiedzią na te wyzwania jest dokument *10 Top Principles for Understanding Artificial Intelligence in Psychology 10* opracowany w ramach Global Psychology Alliance (GPA) (Global Psychology Alliance, 2026). GPA jest międzynarodową inicjatywą skupiającą organizacje psychologiczne i ekspertów z różnych regionów świata, powołaną przez Amerykańskie Towarzystwo Psychologiczne, by reagować na problemy przekraczające możliwości środowisk krajowych. Dokument dotyczący sztucznej inteligencji został opracowany przez zespół autorów reprezentujących różne tradycje zawodowe, instytucje i konteksty kulturowe. Wśród autorów znaleźli się m.in. przedstawiciele American Psychological Association, Pontificia Universidad Católica del Perú, Jamaican Psychological Society, University College Dublin, The Finnish Psychological Association oraz Counselor Council of India. Prace były koordynowane przez Marię Elenę Humphrey z Central American Union of Psychology Associations, a dokument został poddany recenzji przez ekspertów z Kanady, Kolumbii, Pakistanu, Węgier, Tajwanu i Filipin.

Z perspektywy polskiego środowiska psychologicznego szczególne znaczenie ma fakt, że Polskie Towarzystwo Psychologiczne, jako członek Global Psychology Alliance, włącza się w tę międzynarodową współpracę. Dokument GPA ma wartość nie tylko ze względu na treść, lecz także na sposób powstania: jest efektem międzynarodowej refleksji nad tym, jak chronić godność, bezpieczeństwo i prawa osób korzystających z wiedzy psychologicznej w warunkach transformacji cyfrowej.

Dokument nie jest kodeksem prawnym ani szczegółową procedurą stosowania AI. Ma charakter normatywno-koncepcyjny: porządkuje główne obszary ryzyka, wskazuje podstawowe zasady odpowiedzialnego korzystania z AI i tworzy ramę dla dalszych regulacji zawodowych, edukacyjnych i instytucjonalnych.

Zawartość dokumentu

Dokument przedstawia dziesięć zasad rozumienia i stosowania sztucznej inteligencji w psychologii. Zostały one uporządkowane w trzech głównych obszarach: (1) etyka, transparentność i odpowiedzialność; (2) kompetencje zawodowe i nadzór człowieka; (3) dostęp, sprawiedliwość i inkluzja. Taka struktura dobrze oddaje najważniejsze pytania, które powinien zadawać sobie psycholog korzystający z narzędzi AI, a mianowicie:

- 1) Czy ich użycie jest etycznie uzasadnione?
- 2) Czy psycholog rozumie nie tylko ich zalety, ale też i ograniczenia?
- 3) Czy zachowuje profesjonalny osąd?
- 4) Czy technologia nie pogłębia nierówności i nie zwiększa ryzyka szkody?

Pierwsza grupa zasad dotyczy etyki, transparentności, świadomej zgody oraz odpowiedzialności

Dokument podkreśla, że stosowanie AI w psychologii powinno być zakorzenione w uznanych zasadach etycznych, takich jak poszanowanie godności osoby, odpowiedzialność, sprawiedliwość, troska o dobrostan oraz unikanie szkody. Psychologowie powinni rozumieć nie tylko techniczne możliwości AI, ale również jej wpływ na ludzkie decyzje, zaufanie, zależność od technologii, relacje społeczne i poczucie sprawstwa. Sztuczna inteligencja nie może być traktowana jako neutralne narzędzie, ponieważ jej działanie zawsze zależy od przyjętych założeń, danych, sposobu projektowania i kontekstu użycia. Przykładem może być aplikacja wspierająca samoocenę nastroju, która na podstawie krótkich odpowiedzi użytkownika sugeruje „wysokie ryzyko depresji” i zachęca do skorzystania z płatnej konsultacji. Jeśli użytkownik nie wie, że rekomendacja została wygenerowana automatycznie, na podstawie ograniczonych danych i bez diagnozy klinicznej, może potraktować ją jako wiarygodną ocenę psychologiczną. W takim przypadku problemem nie jest sama technologia, lecz brak transparentności, świadomej zgody i odpowiedzialnego nadzoru ze strony specjalisty.

Tak więc bardzo ważna w korzystaniu z AI jest transparentność. Psychologowie powinni wybierać wyłącznie takie narzędzia AI, które dostarczają możliwie jasne informacje o źródłach danych, ograniczeniach oraz sposobach ich przetwarzania. Osoby korzystające z usług psychologicznych, uczestnicy badań, studenci lub inni odbiorcy działań psychologa powinni wiedzieć, kiedy w procesie wykorzystywany jest system AI. Dotyczy to zwłaszcza sytuacji, w których użytkownik może błędnie uznać, że komunikuje się z człowiekiem albo że wygenerowana treść jest efektem samodzielnego osądu specjalisty. Wymóg przejrzystości ma zatem chronić zaufanie do psychologa i do psychologii jako profesji.

Z transparentnością ściśle wiąże się zasada świadomej zgody. Dokument rozszerza klasyczne rozumienie zgody o nowe aspekty wynikające z użycia AI. Klient, pacjent, uczestnik badania lub osoba szkolona powinna otrzymać informację o tym, czy dane będą zbierane, analizowane, przechowywane lub przetwarzane przez narzędzie AI. Zgoda powinna obejmować wiedzę o potencjalnych korzyściach i ryzykach, ograniczeniach systemu oraz ewentualnym wykorzystaniu danych osobowych. Autorzy podkreślają również, że zgoda nie może być uzyskiwana za pomocą manipulacyjnych rozwiązań projektowych, które utrudniają użytkownikowi świadome podjęcie decyzji.

Kolejna zasada dotyczy odpowiedzialności i zarządzania. Użycie AI nie przenosi odpowiedzialności zawodowej z psychologa na technologię. Nawet jeśli system generuje analizę, odpowiedź, interpretację, diagnozę wspierającą lub rekomendację, psycholog pozostaje odpowiedzialny za sposób wykorzystania tych informacji. Ma to szczególne znaczenie w diagnozie, interwencji kryzysowej, psychoterapii, opiniowaniu, badaniach naukowych i edukacji. Przykładem może być użycie programu, który po analizie wypowiedzi klienta oznacza „wysokie ryzyko samobójcze”. Psycholog nie może automatycznie przyjąć ani zignorować takiej informacji. Powinien zweryfikować ją w rozmowie, odnieść do kontekstu klinicznego i zastosować właściwe procedury bezpieczeństwa. Jeśli system się myli,

odpowiedzialność za decyzję nie spoczywa na algorytmie, lecz na specjaliście i instytucji, które dopuściły go do użycia. Dokument przypomina więc, że AI działa w złożonym układzie wielu interesariuszy: twórców systemów, dostawców, instytucji, pracodawców, regulatorów i użytkowników. Wielopodmiotowość nie może jednak prowadzić do rozmycia odpowiedzialności. Psycholog powinien wiedzieć, kto kontroluje narzędzie, jakie dane są wykorzystywane, jakie są procedury nadzoru i jak można reagować na błędy lub szkody.

Druga grupa zasad odnosi się do kompetencji zawodowych, podstaw naukowych, nadzoru człowieka i zapobiegania szkodom

Dokument wskazuje, że psychologowie powinni posiadać wystarczające rozumienie narzędzi AI, których używają. Nie oznacza to konieczności bycia programistą ani specjalistą uczenia maszynowego. Oznacza natomiast zdolność do krytycznej oceny: do czego zostało zaprojektowane narzędzie, na jakich danych działa, jakie ma ograniczenia, czy zostało zwalidowane i czy nadaje się do konkretnego celu psychologicznego. Przykładem może być psycholog szkolny korzystający z aplikacji, która na podstawie krótkiego kwestionariusza wskazuje uczniów „zagrożonych zaburzeniami lękowymi”. Psycholog nie musi znać kodu programu, ale powinien wiedzieć, czy narzędzie było badane wśród dzieci i młodzieży, w jakim kraju powstało, jakie ma normy, jak często daje wyniki fałszywie dodatnie i czy może być stosowane w polskim kontekście szkolnym. Bez tej wiedzy wynik powinien być traktowany wyłącznie jako sygnał do dalszej oceny. Kompetencja cyfrowa i umiejętność krytycznej oceny AI stają się więc elementem współczesnej kompetencji zawodowej psychologa.

Bardzo ważna jest zasada ludzkiego nadzoru i profesjonalnego osądu. AI może wspierać pracę psychologa, ale nie powinna zastępować jego decyzji. Wyniki generowane przez system wymagają interpretacji w kontekście osoby, sytuacji, kultury, języka, historii życia, celu badania lub charakteru interwencji. Przykładem może być system, który po analizie wypowiedzi nastolatka kwalifikuje go do grupy „niskiego ryzyka” kryzysu psychicznego, ponieważ nie pojawiają się wprost słowa dotyczące samookaleczeń. Psycholog, znając kontekst rodzinny, wcześniejsze próby autoagresji i sposób komunikacji ucznia, może uznać wynik za mylący. Profesjonalny osąd wymaga wtedy odrzucenia rekomendacji AI oraz podjęcia dalszej rozmowy i działań ochronnych. Dokument zwraca uwagę na ryzyko tzw. automation bias, czyli nadmiernego zaufania do rekomendacji generowanych automatycznie. W psychologii jest to szczególnie niebezpieczne, ponieważ pozornie obiektywny wynik może przesłonić złożoność doświadczeń człowieka. Psycholog musi zachować prawo i obowiązek zakwestionowania, skorygowania lub odrzucenia wyniku generowanego przez AI.

Zasada bezpieczeństwa psychologicznego i zapobiegania szkodom wskazuje, że AI może generować nowe ryzyka: błędne rekomendacje, fałszywe informacje, nadmierną personalizację, wzmacnianie zależności emocjonalnej od systemu, stygmatyzację, niewłaściwe komunikaty kryzysowe albo błędne oznaczanie ryzyka samobójczego czy nawrotu choroby. W obszarze zdrowia psychicznego konsekwencje takich błędów mogą być poważne. Psychologowie powinni więc monitorować

skutki używania AI, przewidywać możliwe szkody i rezygnować z narzędzia, gdy ryzyko przewyższa korzyści.

Dokument mocno akcentuje również konieczność walidacji i oparcia narzędzi AI na dowodach naukowych. Narzędzia AI nie są samowalidujące. Ich użyteczność zależy od jakości danych, trafności modelu, rzetelności wyników, kontekstu użycia i populacji, dla której zostały sprawdzone. Dowody pochodzące z jednej kultury, grupy językowej lub systemu opieki nie muszą być automatycznie przenoszalne na inne warunki. Jeżeli dane narzędzie nie ma niezależnej walidacji, psycholog powinien traktować je najwyżej jako pomoc eksploracyjną lub organizacyjną, a nie jako podstawę decyzji wpływających na życie, zdrowie, praca lub możliwości osób.

Trzecia grupa zasad dotyczy stronniczości algorytmicznej, sprawiedliwości, dostępu i inkluzji

Autorzy przypominają, że systemy AI mogą powielać lub wzmacniać nierówności obecne w danych treningowych i praktykach społecznych. Bias algorytmiczny może dotyczyć płci, wieku, języka, pochodzenia etnicznego, statusu społeczno-ekonomicznego, niepełnosprawności, miejsca zamieszkania lub kontekstu kulturowego. W praktyce psychologicznej może to prowadzić do błędnej diagnozy, nieadekwatnych rekomendacji, nierównego dostępu do usług lub utrwalania stereotypów. Przykładem może być narzędzie AI wspierające diagnozę trudności emocjonalnych, trenowane głównie na danych dorosłych użytkowników anglojęzycznych z dużych miast. Zastosowanie wobec nastolatków z małych miejscowości lub osób posługujących się językiem mniejszości może błędnie interpretować ich sposób wyrażania emocji jako „niski wgląd” albo „opór”. W efekcie system może rekomendować mniej adekwatne formy pomocy. Psycholog powinien więc sprawdzić, czy narzędzie było walidowane dla danej grupy językowej, kulturowej i wiekowej. Z tego powodu psychologowie powinni pytać, czy dane narzędzie działa równie dobrze w różnych grupach i czy nie systematycznie dyskryminuje osoby już narażone na marginalizację.

Ostatnia zasada dotyczy dostępu, równości i inkluzji. AI bywa przedstawiana jako narzędzie demokratyzujące dostęp do pomocy psychologicznej, ponieważ może zwiększać dostępność usług i obniżać ich koszt. Dokument zachowuje jednak ostrożność wobec tej tezy. Dostęp do narzędzi AI zależy od infrastruktury cyfrowej, języka, kompetencji technologicznych, kosztów, jakości połączenia internetowego, dostępności dla osób z niepełnosprawnościami i dopasowania kulturowego. Bez refleksji nad tymi czynnikami AI może nie zmniejszać, lecz pogłębiać nierówności. Przykładem może być bezpłatny chatbot „pierwszego wsparcia psychologicznego”, dostępny wyłącznie przez aplikację mobilną, w języku angielskim i wymagający stałego dostępu do Internetu. Dla studentów dużych miast może być wygodnym uzupełnieniem pomocy, ale dla osób starszych, mieszkańców terenów słabiej skomunikowanych cyfrowo, osób z niepełnosprawnościami lub użytkowników nieznających angielskiego stanie się kolejną barierą. W takim przypadku AI formalnie zwiększa ofertę, ale realnie poszerza dostęp tylko dla wybranych grup. Dlatego psychologowie powinni oceniać, czy konkretne

rozwiązanie rzeczywiście poszerza dostęp do pomocy, czy raczej uprzywilejowuje osoby i grupy posiadające już zasoby technologiczne.

Dokument kończy się zestawieniem podstawowych pojęć, takich jak sztuczna inteligencja, generatywna AI, duży model językowy, dane treningowe, czarna skrzynka, stronniczość algorytmiczna, automation bias i halucynacje AI. Jest to ważny element praktyczny, ponieważ porządkuje język debaty i może ułatwiać korzystanie z dokumentu psychologom, którzy na co dzień nie zajmują się technologią.

Wartość, ograniczenia i potrzeba dalszych regulacji

Wartość dokumentu GPA polega przede wszystkim na tym, że wskazuje i omawia najważniejsze obszary odpowiedzialności psychologa wobec sztucznej inteligencji. Pokazuje, że AI nie jest jedynie narzędziem technicznym, lecz elementem relacji zawodowych, procesu decyzyjnego, organizacji usług, badań naukowych i edukacji. Dokument przypomina, że podstawowe wartości psychologii – szacunek dla godności osoby, jej dobrostan, odpowiedzialność, sprawiedliwość, trafność oddziaływań i ochrona przed szkodą – pozostają aktualne także w środowisku cyfrowym.

Istotną zaletą dokumentu jest jego międzynarodowy charakter. Włączenie perspektyw z różnych krajów pozwala uniknąć zawężenia refleksji do jednego systemu prawnego, jednej kultury lub jednego modelu świadczenia usług. Dokument podkreśla znaczenie globalnych nierówności w dostępie do technologii, różnic kulturowych oraz ograniczeń infrastrukturalnych. Jest to szczególnie ważne, ponieważ wiele narzędzi AI powstaje w warunkach dominacji języka angielskiego, dużych rynków technologicznych i danych pochodzących z wybranych populacji.

Ograniczeniem dokumentu jest natomiast jego ogólny charakter. Nie zawiera on szczegółowych procedur dotyczących polskich realiów prawnych, zawodowych i instytucjonalnych. Nie rozstrzyga, które narzędzia AI mogą być dopuszczalne w diagnozie i pomocy psychologicznej, psychoterapii, interwencji kryzysowej, opiniowaniu, edukacji lub badaniach naukowych. Nie wskazuje również minimalnych kryteriów walidacyjnych, standardów dokumentowania użycia AI, zasad pracy z danymi wrażliwymi ani procedur reagowania na błąd systemu. Z tego powodu dokument należy traktować jako ramę refleksji, a nie jako gotowy standard praktyki.

Dla polskich psychologów dokument GPA może być ważnym punktem wyjścia do dyskusji służących opracowaniu lokalnych przepisów. Potrzebne są szczegółowe rekomendacje dotyczące m.in.: ochrony danych, świadomej zgody, dokumentowania użycia AI, zakresu zastosowania i dopuszczalności narzędzi w diagnozie i terapii, zasad korzystania z AI w pracy z dziećmi i młodzieżą oraz odpowiedzialności zawodowej w sytuacjach wysokiego ryzyka. Przyszłe regulacje powinny łączyć perspektywę etyczną, prawną, naukową i praktyczną, tak aby sztuczna inteligencja mogła wspierać psychologię, nie naruszając jej podstawowych wartości.

Bibliografia

- Aggarwal, A., Tam, C. C., Wu, D., Li, X., Qiao, S. (2023). Artificial intelligence-based chatbots for promoting health behavioral changes: Systematic review. *Journal of Medical Internet Research*, 25, artykuł e40789. <https://doi.org/10.2196/40789>
- Cruz-Gonzalez, P., He, A. W.-J., Lam, E. P., Ng, I. M. C., Li, M. W., Hou, R., Chan, J. N.-M., Sahni, Y., Vinas Guasch, N., Miller, T., Lau, B. W.-M., Sánchez Vidaña, D. I. (2025). Artificial intelligence in mental health care: A systematic review of diagnosis, monitoring, and intervention applications. *Psychological Medicine*, 55, artykuł e18. <https://doi.org/10.1017/S0033291724003295>
- Dakanalis, A., Wiederhold, B. K., Riva, G. (2024). Artificial intelligence: A game-changer for mental health care. *Cyberpsychology, behavior and social networking*, 27(2), 100–104. <https://doi.org/10.1089/cyber.2023.0723>
- Global Psychology Alliance. (2026). *Principles for understanding artificial intelligence in psychology*. <https://www.apa.org/international/networks/global-psychology-alliance/principles-artificial-intelligence.pdf>
- Kirthivasan, M., Bhardwaj, A. B. (2026). *ChatGPT (AI) vs. standardized psychological testing*. W: R. Sharma, F. Rabby, R. Bansal, T. Sachdeva, J. Mrabet (red.), *Confronting mental health stigma with AI and machine learning* (s. 203–237). Wiley. <https://doi.org/10.1002/9781394347308.ch9>
- Kolding, S., Lundin, R. M., Hansen, L., Østergaard, S. D. (2024). Use of generative artificial intelligence (AI) in psychiatry and mental health care: A systematic review. *Acta Neuropsychiatrica*, 37, artykuł e37. <https://doi.org/10.1017/neu.2024.50>
- McCrudden, K. E., Swirbul, M. S., Peake, E. E., Rodio, M. J., Padmanabhan, A. (2026) AI-powered documentation for mental health providers: Retrospective observational mixed methods study. *JMIR Formative Research*, 10, artykuł e84628–e84628. <https://doi.org/10.2196/84628>
- Olson, K. D., Meeker, D., Troup, M., Barker, T. D., Nguyen, V. H., Manders, J. B., Stults, C. D., Jones, V. G., Shah, S. D., Shah, T., Schwamm, L. H. (2025). Use of ambient AI scribes to reduce administrative burden and professional burnout. *JAMA network open*, 8(10), artykuł e2534976. <https://doi.org/10.1001/jamanetworkopen.2025.34976>
- Prégent, J., Chung, V. H. A., El Adib, I., Désilets, M., Hudon, A. (2025). Applications of artificial intelligence in psychiatry and psychology education: Scoping review. *JMIR Medical Education*, 11, artykuł e75238. <https://doi.org/10.2196/75238>