

WARSAW PRINCIPLES

Zasady etycznego wykorzystania sztucznej inteligencji w public relations

Wstęp

Profesjonaliści public relations w agencjach PR coraz częściej wykorzystują technologie oparte na sztucznej inteligencji, w celu usprawnienia swojej pracy oraz osiągania znaczących wyników. Jednak, w miarę jak sztuczna inteligencja staje się coraz bardziej powszechna w branży PR, kluczowe znaczenie przy jej wykorzystaniu ma opracowanie wytycznych, które zapewnią, że będzie się to odbywało w sposób etyczny i odpowiedzialny. Dzięki tym wytycznym profesjonaliści PR mogą uporać się ze złożonością sztucznej inteligencji, a jednocześnie przestrzegać standardów etycznych, respektować prawo do prywatności i prawa autorskie oraz zachować poufność klientów.

Stosując się do tych zasad, agencje PR mogą skutecznie wykorzystywać sztuczną inteligencję do świadczenia strategicznych konsultacji, tworzenia różnorodnych treści, zarządzania komunikacją w sytuacjach kryzysowych i problemowych oraz budowania pozytywnych i opartych na zaufaniu relacji z dziennikarzami i innymi interesariuszami.

- Treści **wygenerowane przez AI (ang. AI-generated)** to treści lub rezultaty w całości wytworzone przez sztuczną inteligencję bez bezpośredniej ingerencji człowieka w ten proces.
- Treści **współtworzone przy pomocy AI (ang. AI-assisted)** oznaczają procesy, decyzje lub treści, w których sztuczna inteligencja odgrywa rolę pomocniczą (wspierającą lub wzmacniającą). Zazwyczaj oznacza to, że stanowi ona wsparcie dla człowieka, aby podejmował trafniejsze decyzje, ale nie wyręcza go z tego.
- Termin **„stworzony przy wykorzystaniu AI” (ang. AI-powered)** oznacza dowolne wykorzystanie sztucznej inteligencji, również generatywnej.

WARSAW PRINCIPLES

1

Transparentność, jawność i autentyczność:

Profesjonaliści public relations powinni zawsze otwarcie informować, gdy treści zostały przygotowane przy wykorzystaniu generatywnej sztucznej inteligencji, gdyż mogą one zniekształcić postrzeganie rzeczywistości przez opinię publiczną. Kluczowe jest każdorazowe informowanie o wykorzystaniu treści **wygenerowanych przez AI**, zwłaszcza dlatego, że żyjemy w czasach stosowania *deepfakes*.

Przykłady:

- Należy jednoznacznie informować, że do stworzenia danych treści została wykorzystana generatywna sztuczna inteligencja, aby nie wprowadzić odbiorców w błąd i zachować etyczną transparentność.
- Należy wyraźnie informować użytkowników, że wchodzi w interakcję z opartym na sztucznej inteligencji chatbotem, a nie człowiekiem.
- Należy informować klientów, kiedy i w jaki sposób sztuczna inteligencja jest wykorzystana w kampanii PR.
- Należy zrozumieć ograniczenia i stronniczość technologii oraz jej wpływ na podejmowanie decyzji.
- Należy współpracować ze stowarzyszeniami, organami zarządzającymi, organizacjami prowadzącymi badania nad sztuczną inteligencją lub ekspertami w tej dziedzinie, aby zapewnić, że technologia jest wykorzystywana w sposób zgodny ze standardami etycznymi i zawodowymi.

2

Rzetelność, Fact-Checking i zwalczanie dezinformacji:

Profesjonaliści public relations powinni zadbać o rzetelność, wiarygodność i weryfikację treści **stworzonych przy wykorzystaniu AI** przed ich rozpowszechnieniem. Zalecane jest stosowanie narzędzi wykorzystujących sztuczną inteligencję do sprawdzania faktów oraz demaskowania dezinformacji i błędnych informacji; należy jednak pamiętać, że wiele ogólnodostępnych narzędzi wykorzystujących generatywną sztuczną inteligencję często dostarcza niedokładne informacje, więc warto sprawdzać ich wiarygodność również w źródłach, które nie są oparte na tej technologii. Z uwagi na zdolność sztucznej inteligencji do szybkiego tworzenia i rozpowszechniania dezinformacji, profesjonaliści PR muszą zachować szczególną ostrożność, aby zapobiegać nieumyślnemu lub celowemu rozpowszechnianiu nieprawdziwych informacji.

Przykłady:

- Należy sprawdzić rzetelność informacji, danych i analiz przed zawarciem ich w raportach dla klientów lub materiałach dla mediów.
- Należy porównywać informacje z aktualnymi i wiarygodnymi źródłami, aby zweryfikować ich autentyczność i dokładność.
- Należy upewnić się, że narzędzia służące do tłumaczeń językowych oparte na sztucznej inteligencji wiernie oddają znaczenie oryginału. W tym celu należy poprosić o sprawdzenie przekładu rodzimych użytkowników danego języka lub osoby, które znają ten język na poziomie natywnym.
- Należy upewnić się, że narzędzia sztucznej inteligencji zaprojektowane w celu zwiększenia zasięgu treści nie wzmacniają nieumyślnie dezinformacji ani nie promują wprowadzających w błąd narracji.

3

Prywatność, ochrona danych i odpowiedzialne udostępnianie:

Profesjonaliści PR muszą traktować dane klientów i poufne informacje o klientach z najwyższą ostrożnością oraz przestrzegać przepisów o ochronie danych. Są oni odpowiedzialni za udostępnianie zweryfikowanych i niewprowadzających w błąd treści.

Przykłady:

- Należy upewnić się, że algorytmy wykorzystywane do monitorowania mediów są zgodne z prawem do prywatności i nie mają dostępu do danych osobowych ani ich nie przechowują.
- Należy anonimizować i agregować dane klientów, aby chronić ich prywatność, w trakcie generowania informacji wykorzystywanych do tworzenia kampanii PR-owych. Zabrania się wykorzystywać prywatnych, chronionych lub wrażliwych informacji o klientach do doskonalenia ogólnodostępnych modeli sztucznej inteligencji.
- Należy podjąć kroki mające na celu ochronę prywatności osób, których opinie są analizowane i zagwarantować, że podane przez nich dane nie pozwolą na ustalenie ich tożsamości.

4

Wykrywanie uprzedzeń, łagodzenie ich skutków i inkluzywność:

Profesjonaliści PR powinni aktywnie identyfikować i eliminować uprzedzenia, które mogą wynikać z treści **stworzonych przy wykorzystaniu AI**, a także szukać możliwości wykorzystania sztucznej inteligencji w taki sposób, aby kampanie były niewykluczające i dostosowane do szerszego grona odbiorców. Ważne jest również, aby korzystać z różnorodnych danych do uczenia modeli AI w celu zapewnienia inkluzywności.

Przykłady:

- Należy analizować materiały medialne **współtworzone przy pomocy AI**, aby zidentyfikować i skorygować wszelkiego rodzaju uprzedzenia w wyborze odbiorców lub języka.
- Należy regularnie monitorować algorytmy sztucznej inteligencji, aby sprawdzić, czy nie utrwalają stereotypów i uprzedzeń wynikających z pobudek politycznych lub ideologicznych w rekomendowanych treściach.
- Należy polegać na ocenie człowieka przy wyborze obrazów, słów i filmów z materiałów stworzonych przez sztuczną inteligencję.
- Należy stale monitorować i oceniać algorytmy sztucznej inteligencji, aby ograniczyć do minimum uprzedzenia w wyborze influencerów oraz zapewniać sprawiedliwą reprezentację i różnorodność.

5

Własność intelektualna, przestrzeganie praw autorskich i umiejętności korzystania z mediów:

Profesjonaliści PR muszą przestrzegać praw autorskich i praw własności intelektualnej za każdym razem, gdy korzystają ze sztucznej inteligencji, niezależnie od sposobu jej wykorzystania. Powinni promować umiejętności korzystania z mediów pod kątem identyfikowania treści **stworzonych przy wykorzystaniu AI** i wykrywania *deepfakes*.

Przykłady:

- Należy zapoznać się z warunkami korzystania z użytkowanych i subskrybowanych narzędzi AI, aby poznać materiały źródłowe, na podstawie których działają (czyli tzw. „karmienie”), a także zobowiązania dotyczące własności intelektualnej i praw autorskich oraz inne stosowane zasady.
- Należy upewnić się, że obrazy lub filmy **wygenerowane przez AI** nie naruszają praw autorskich i uzyskać stosowną zgodę na ich wykorzystanie.
- Należy weryfikować treści **stworzone przy wykorzystaniu AI**, aby zagwarantować, że nie są plagiatem ani nie naruszają praw autorskich.
- Należy zagwarantować, że algorytmy sztucznej inteligencji przeznaczone do przechowywania (ang. content curation) treści są zgodne z zasadami dozwolonego użytku i nie naruszają praw autorskich.

6

Nadzór człowieka, interwencje i współpraca:

Profesjonaliści PR powinni zapewnić nadzór i możliwość interwencji człowieka w proces **tworzenia treści przy wykorzystaniu AI** oraz współpracować ze specjalistami ds. sztucznej inteligencji, aby uzyskać wytyczne i wskazówki.

Przykłady:

- Należy angażować ekspertów do weryfikacji informacji o charakterze strategicznym, jak i komunikatów dotyczących reagowania na sytuacje kryzysowe przed ich opublikowaniem.
- Należy dzielić się spostrzeżeniami ekspertów i fachową wiedzą w materiałach opiniotwórczych, aby zagwarantować ich rzetelność i adekwatność.
- Należy zaangażować specjalistów do sprawdzenia i dostosowania współtworzonych przy pomocy AI komunikatów dla mediów, aby zapewnić ich dobrą jakość i adekwatność.

7

Zrozumienie kontekstu, adaptacja i personalizacja:

Profesjonaliści PR muszą przeanalizować kontekst i stosowność treści **współtworzonych przy pomocy AI** z myślą o różnych kanałach i grupach docelowych, a także zgłębić potencjał sztucznej inteligencji w zakresie dopasowania doświadczeń.

Przykłady:

- Należy dostosować treści **współtworzone przy pomocy AI** do zróżnicowanego tonu, stylu i wymagań poszczególnych platform mediów społecznościowych.
- Należy weryfikować i dostosowywać treści **współtworzone przy pomocy AI**, aby zaspokajały potrzeby i odpowiadały na preferencje zróżnicowanych grup odbiorców.
- Należy dostosować treści **współtworzone przy pomocy AI** skierowane na rynki międzynarodowe w taki sposób, aby uwzględniały one niuanse i wrażliwość różnych kultur.

8

Odpowiedzialna automatyzacja i wydajność:

Profesjonaliści PR powinni wykorzystywać sztuczną inteligencję do automatyzacji powtarzalnych zadań i zwiększania wydajności tak, by nie naruszać standardów etycznych i zawodowych.

Przykłady:

- Należy stosować narzędzia wykorzystujące sztuczną inteligencję do automatyzacji monitorowania i analizy mediów, co pozwoli profesjonalistom PR skoncentrować się na podejmowaniu decyzji o znaczeniu strategicznym.
- Należy używać algorytmów sztucznej inteligencji do tworzenia harmonogramów publikacji treści, co pozwoli na optymalizację pracy i odpowiednie wykorzystanie zasobów.
- Należy stosować sztuczną inteligencję do tworzenia różnych wariantów treści na podstawie zatwierdzonych materiałów źródłowych, o ile do narzędzi wykorzystujących sztuczną inteligencję nie są wprowadzane informacje poufne.

9

Ciągłe monitorowanie, ocena i informacje zwrotne:

Profesjoniści PR powinni regularnie monitorować i oceniać wydajność oraz wpływ **treści stworzonych przy wykorzystaniu AI** oraz umożliwić zainteresowanym stronom wyrażanie opinii i zgłaszanie wątpliwości.

Przykłady:

- Należy stale oceniać dokładność i skuteczność analizy sentymentów **wygenerowanej przez AI**, aby sprawdzić, czy jest zgodna z interpretacją dokonywaną przez człowieka.
- Należy monitorować zaangażowanie i reakcje na materiały medialne **współtworzone przy pomocy AI**, aby ocenić ich skuteczność i udoskonalać przyszłe strategie.
- Należy na bieżąco oceniać materiały medialne (wizualne i tekstowe) **współtworzone przy pomocy AI** uwzględniając informację zwrotną od odbiorców tych treści.

10

Etyczny rozwój zawodowy, edukacja i promowanie sztucznej inteligencji:

Profesjoniści PR powinni nieustannie się dokształcać i rozwijać zawodowo, aby być na bieżąco z postęпами w zakresie sztucznej inteligencji i względami etycznymi. Są oni odpowiedzialni za promowanie etycznego wykorzystania sztucznej inteligencji i dobrych praktyk w szeroko pojętym otoczeniu medialnym.

Przykłady:

- Należy stworzyć fora edukacyjne i dyskusyjne w swoich agencjach/zespołach, aby pracownicy na wszystkich etapach rozwoju zawodowego mogli ze zrozumieniem korzystać z rozwiązań opartych na sztucznej inteligencji.
- Należy uczestniczyć w programach szkoleniowych i warsztatach dotyczących etycznego wykorzystania sztucznej inteligencji, aby rozumieć jej tendencyjność oraz skuteczniej odpowiadać na wyzwania etyczne.
- Należy aktywnie uczestniczyć w dyskusjach, konferencjach i wydarzeniach branżowych w celu wymiany wiedzy i dobrych praktyk, aby promować etyczne wykorzystanie sztucznej inteligencji w branży PR.

Niniejsze wytyczne podkreślają znaczenie transparentności, dokładności, ochrony prywatności, wykrywania stronniczości i uprzedzeń, zgodności z prawem własności intelektualnej, nadzoru ludzkiego, zrozumienia kontekstu, odpowiedzialnej automatyzacji, ciągłego monitorowania i rozwoju zawodowego, zapewniając ramy do poruszania się po wyzwaniach i szansach związanych z technologią AI.

Przestrzeżenie tych zasad to gwarancja, że agencje PR będą skutecznie wykorzystywać sztuczną inteligencję bez wyrządzania szkód lub naruszania praw człowieka, praw autorskich, prawa do prywatności lub do znaków towarowych. Profesjonaliści PR muszą również walczyć z dezinformacją, aby dbać o spójność i wiarygodność komunikacji i PR w dobie cyfryzacji.

Sztuczna inteligencja, choć jest zaawansowanym narzędziem, powinna uzupełniać, a nie zastępować człowieka, czy też bezcenną wiedzę, umiejętność oceny sytuacji i kreatywność, które zapewniają profesjonaliści PR. Sztuczna inteligencja i konsekwencje jej wykorzystania będą się zmieniać, dlatego zdolność adaptacji i zachowanie czujności są niezbędne, aby skutecznie odpowiadać na wyzwania wynikające z szybko ewoluujących technologii AI.

„**Warsaw Principles**”, czyli zasady etycznego wykorzystania sztucznej inteligencji w public relations zostały przyjęte przez stowarzyszenia członkowskie ICCO, w tym przez Związek Firm Public Relations, podczas ICCO Global Summit, które odbyło się w październiku 2023 r. w Warszawie.

Autor: Christina Forsgård, Ethics Chair, ICCO

Bibliografia:

- *ICCO Global Code of Ethics for PR Professionals* (tj. Kodeks Etyki ICCO dla profesjonalistów PR) <https://iccopr.com/ethics/>
- *General Data Protection Regulation (GDPR) - European Commission* (tj. Ogólne rozporządzenie o ochronie danych osobowych (RODO) - Komisja Europejska) https://commission.europa.eu/law/law-topic/data-protection_en
- *Guidelines for AI Ethics in Public Relations - PR Council* (tj. Wytyczne dotyczące etycznego wykorzystania sztucznej inteligencji w public relations - PR Council) <https://prcouncil.net/wp-content/uploads/2023/04/PR-Council-Guidelines-on-Generative-AI-042423.pdf>
- *OpenAI*: <https://openai.com/blog/how-should-ai-systems-behave>
- *AI in the Communications Industry 2023* (tj. Sztuczna inteligencja w branży komunikacyjnej w 2023 roku): https://www.provokemedia.com/docs/default-source/default-document-library/sp-pr-ai-report_final.pdf
- *The Future of AI in PR* (tj. Przyszłość sztucznej inteligencji w PR): <https://www.cision.ca/resources/webinars-events/future-ai-pr/>

Uwaga: Aby zapewnić zgodność ze standardami etycznymi i wymogami prawnymi, należy koniecznie zapoznać się z aktualnymi i właściwymi dla danego regionu i branży materiałami źródłowymi i wytycznymi.