

piątek, 31.01.2025

Sztuczna inteligencja (AI) – szansa i zagrożenie

Nie należy traktować sztucznej inteligencji jako osoby, nie należy jej ubóstwiać ani pozwalać, by zastąpiła relacje międzyludzkie, lecz powinna być wykorzystywana „jedynie jako narzędzie uzupełniające ludzką inteligencję”. Ostrzeżenia Papieża dotyczące sztucznej inteligencji z ostatnich lat stanowią ośnowę dokumentu „*Antiqua et Nova*” (odwołanie do „mądrości starej i nowej”), noty dotyczącej relacji między sztuczną inteligencją a ludzką inteligencją. Jest to owoc wspólnej refleksji Dykasterii Nauki Wiary oraz Dykasterii ds. Kultury i Edukacji. Dokument jest skierowany do rodziców, nauczycieli, kapłanów, biskupów i wszystkich powołanych do wychowania i przekazywania wiary, a także do tych, którzy uznają konieczność rozwoju naukowego i technologicznego „w służbie osoby i dobra wspólnego” [5]. Opublikowana 28 stycznia *Nota* została zatwierdzona przez Papieża. W 117 paragrafach *Antiqua et nova* podkreśla wyzwania i możliwości rozwoju sztucznej inteligencji (AI) w dziedzinie edukacji, ekonomii, pracy, zdrowia, stosunków międzynarodowych i międzyludzkich oraz w kontekście działań wojennych. W tym ostatnim przypadku – jak ostrzega *Nota* – możliwości AI mogłyby zwiększyć zasoby militarne „daleko poza zasięgiem kontroli człowieka”, przyspieszając „destabilizujący wyścig zbrojeń, prowadzący do katastrofalnych konsekwencji dla praw człowieka” [99].

Niebezpieczeństwa i korzyści Dokument szczegółowo i w wyważony sposób wymienia zarówno zagrożenia związane ze sztuczną inteligencją, jak i korzyści z niej płynące, które wręcz zachęca do rozwijania jako „część współpracy” człowieka z Bogiem „w doprowadzaniu widzialnego stworzenia do doskonałości” [2]. Mimo to niepokój jest duży i dotyczy wszystkich innowacji, których skutki pozostają wciąż nieprzewidywalne – nawet tych, które obecnie wydają się nieszkodliwe, jak generowanie tekstów i obrazów.

Rozróżnienie między sztuczną inteligencją a inteligencją ludzką Refleksje dwóch Dykasterii koncentrują się na etycznych i antropologicznych aspektach, które znajdują się w centrum dokumentu. Wiele paragrafów *Noty* poświęcono „decydującemu” rozróżnieniu między sztuczną inteligencją (AI) a inteligencją ludzką. Inteligencja ludzka „urzeczywistnia się w relacjach” [18], jest kształtowana przez Boga i „plastyczna dzięki niezliczonym doświadczeniom przeżywanym w cielesności”. Sztucznej inteligencji „brakuje zdolności do ewolucji w tym znaczeniu” [31]. Jej perspektywa jest „czysto funkcjonalistyczna”, oceniająca ludzi wyłącznie przez pryzmat ich pracy i osiągnięć, podczas gdy ludzka godność jest niezbywalna i zawsze nienaruszalna. Dotyczy to także „dziecka nienarodzonego”, „osoby w stanie nieświadomości” czy „cierpiącego seniora” [34]. „Mylące” jest zatem używanie słowa „inteligencja” w odniesieniu do sztucznej inteligencji: nie jest ona „*sztuczna forma? inteligencji*”, ale „jednym z jej wytworów” [35].

Władza w ręku nielicznych Jak każdy wytwór ludzkiego geniuszu, sztuczna inteligencja może być skierowana zarówno ku „celom pozytywnym, jak i negatywnym” – podkreśla *Antiqua et Nova*. Dokument nie zaprzecza, że sztuczna inteligencja może wprowadzić „ważne innowacje” w różnych dziedzinach [48], może ona również pogłębić zjawiska marginalizacji, dyskryminacji, ubóstwa, „wykluczenia cyfrowego” i nierówności społecznych [52]. W szczególności „obawy etyczne” budzi fakt, że „większość władzy nad głównymi zastosowaniami sztucznej inteligencji jest zgromadzona w rękach kilku potężnych firm” [53], co powoduje, że technologią tą można manipulować dla osiągnięcia „osobistych lub korporacyjnych korzyści” albo do „kierowania opinią publiczną w interesie określonego sektora” [53].

Wojna Następnie *Nota* analizuje różne aspekty życia w odniesieniu do sztucznej inteligencji. Nie mogło zabraknąć odniesienia do wojny. „Zdolności analityczne” sztucznej inteligencji mogłyby zostać wykorzystane, aby pomóc narodom w dążeniu do pokoju i

bezpieczeństwa, ale „poważne powody do niepokoju etycznego” budzą systemy autonomicznej i śmiertelnej broni, zdolne do „identyfikowania i atakowania celów bez bezpośredniej interwencji człowieka” [100]. Papież wezwał do pilnego zakazania ich stosowania, jak powiedział na szczycie G7 w Apulii: „Żadna maszyna nie powinna nigdy decydować, czy odebrać życie istocie ludzkiej”. Maszyny zdolne do autonomicznego precyzyjnego zabijania oraz inne technologie masowego niszczenia stanowią realne zagrożenie dla „przetrwania ludzkości lub całych regionów” [101]. Te technologie „nadają wojnie niekontrolowaną niszczycielską moc, która dotyka wielu niewinnych cywilów, nie oszczędzając nawet dzieci” – ostrzega *Antiqua et Nova*. Aby zapobiec wciągnięciu ludzkości w „spirale samozniszczenia”, konieczne jest „zajęcie jednoznacznego stanowiska przeciwko wszelkim zastosowaniom technologii, które zagrażają w sposób fundamentalny życiu i godności osoby ludzkiej”.

Relacje ludzkie Jeśli chodzi o relacje międzyludzkie, w dokumencie zauważono, że sztuczna inteligencja może „sprzyjać nawiązywaniu kontaktów”, ale jednocześnie prowadzić do „szkodliwej izolacji” [58]. „Antropomorfizacja sztucznej inteligencji” stwarza również szczególne trudności dla rozwoju dzieci, które są zachęcane do rozumienia „relacji międzyludzkich w sposób utylitarny”, jak ma to miejsce w przypadku *chatbotów* [60]. Za „błędne” uznaje się przedstawianie sztucznej inteligencji jako osoby, a „poważnym naruszeniem etycznym” jest wykorzystywanie tego w nieuczciwych celach. Podobnie, „używanie AI do wprowadzania w błąd w innych kontekstach – takich jak edukacja czy relacje międzyludzkie, w tym w sferze seksualności – jest głęboko niemoralne i wymaga uważnego nadzoru” [62].

Ekonomia i praca Podobna czujność jest wymagana w kontekście gospodarki i finansów. W szczególności w obszarze pracy zauważa się, że choć sztuczna inteligencja ma „potencjał” do zwiększenia kompetencji, produktywności czy tworzenia nowych miejsc pracy, to jednocześnie może „obniżać kwalifikacje pracowników, poddawać ich automatycznemu nadzorowi oraz sprowadzać ich rolę do sztywnych i powtarzalnych funkcji” [67]. Taki proces może wręcz „dusić” wszelką zdolność do innowacji. *Nota* podkreśla: „Nie możemy dążyć do coraz większego zastępowania pracy ludzkiej postępem technologicznym – postępując w ten sposób, ludzkość szkodzi samej sobie” [70].

Zdrowie Znaczną uwagę poświęca się tematowi ochrony zdrowia. Chociaż sztuczna inteligencja ma ogromny potencjał w różnych zastosowaniach medycznych (na przykład wspieranie diagnostyki), istnieje ryzyko, że zastąpienie relacji między lekarzem a pacjentem interakcją wyłącznie z maszynami mogłoby „pogłębić samotność, która często towarzyszy chorobie”. Optymalizacja zasobów również nie powinna „krzywdzić najsłabszych” ani prowadzić do „uprzedzeń i dyskryminacji” [75], które mogłyby umocnić „medycynę dla bogatych”. W takim systemie osoby posiadające środki finansowe korzystają z zaawansowanych narzędzi profilaktyki i spersonalizowanych informacji medycznych, podczas gdy inni mają trudności z uzyskaniem dostępu nawet do podstawowych usług.

Edukacja W obszarze edukacji również wskazuje się na potencjalne zagrożenia. Jeśli sztuczna inteligencja będzie wykorzystywana w sposób rozsądny, może poprawić dostęp do nauki i zapewnić uczniom „natychmiastową informację zwrotną” [80]. Problem polega jednak na tym, że wiele programów „ogranicza się do udzielania gotowych odpowiedzi zamiast zachęcać uczniów do samodzielnego ich poszukiwania lub tworzenia własnych tekstów”. Prowadzi to do utraty umiejętności gromadzenia informacji oraz rozwijania krytycznego myślenia [82]. Należy również pamiętać o ryzyku generowania przez niektóre programy „zniekształconych lub fałszywych informacji” bądź „nieścisłych treści”, co może przyczyniać się do legitymizowania *fake newsów* [84].

Fake News i Deepfake W odniesieniu do fake news dokument zwraca uwagę na poważne ryzyko, że sztuczna inteligencja może „generować zmanipulowane treści i fałszywe informacje” [85], tworząc tym samym „halucynacje” AI, czyli nieprawdziwe treści, które

wydają się realne. Jeszcze bardziej niepokojące jest celowe wykorzystywanie takich fikcyjnych treści do manipulacji. Na przykład, gdy celowo rozpowszechnia się obrazy, filmy czy nagrania dźwiękowe w formie deepfake (przerobione lub wygenerowane przez algorytmy), aby „oszukiwać lub szkodzić” [87]. W związku z tym apeluje się o zawsze staranne „sprawdzanie prawdziwości” udostępnianych treści oraz o unikanie w każdym przypadku „udostępniania słów i obrazów uwłaczających godności człowieka”. Podkreśla się również konieczność wykluczania wszystkiego, „co podsycia nienawiść i nietolerancję, poniża piękno i intymność ludzkiej seksualności, a także wykorzystuje słabszych i bezbronnych” [89].

Prywatność i kontrola W kwestii prywatności i kontroli dokument podkreśla, że niektóre rodzaje danych mogą dotyczyć najgłębszej intymności człowieka, „a może nawet jego sumienia” [90]. Dane pozyskiwane są obecnie przy użyciu minimalnej ilości informacji, co rodzi ryzyko, że wszystko stanie się „swego rodzaju spektaklem, który można szpiegować i monitorować” [92]. Zwraca się także uwagę, że „cyfrowa inwigilacja może być wykorzystywana do kontrolowania życia wierzących oraz sposobu wyrażania ich wiary” [90].

Wspólny dom W kontekście troski o stworzenie za „obiecujące” uznaje się zastosowania sztucznej inteligencji, które mogą poprawić relację z naszym wspólnym środowiskiem. Można tu wspomnieć o modelach przewidujących ekstremalne zjawiska klimatyczne, wspierających działania ratunkowe czy promujących zrównoważone rolnictwo [95]. Jednocześnie jednak obecne modele AI i wspierająca je infrastruktura sprzętowa wymagają „ogromnych ilości energii i wody, przyczyniają się znacząco do emisji CO₂ oraz intensywnie zużywają zasoby”. W związku z tym obciążenie dla środowiska jest „ciężkie”: „Rozwój zrównoważonych rozwiązań ma kluczowe znaczenie dla zmniejszenia ich wpływu na nasz wspólny dom”.

Odniesienie do Boga „Przekonanie, że można zastąpić Boga dziełem własnych rąk, jest bałwochwalstwem” – dokument cytuje Pismo Święte, aby przestrzec przed tym, że sztuczna inteligencja może być „bardziej kusząca niż tradycyjne bożki” [105]. Przypomina, że AI jest jedynie „bladym odbiciem” ludzkości: „Sztuczna inteligencja nie powinna być ubóstwiana i czczona, ponieważ w ten sposób człowiek sam staje się niewolnikiem swojego własnego dzieła”. Stąd też końcowe zalecenie: „Sztuczna inteligencja powinna być wykorzystywana jedynie jako narzędzie uzupełniające ludzką inteligencję, a nie zastępujące jej bogactwo” [112].