

Marek Kochan

Katedra Kultury i Mediów, Wydział Nauk Humanistycznych, Uniwersytet SWPS

ORCID 0000-0001-9502-6171

Retoryka botów: autoprezentacja, tożsamość, wizerunek

Rhetoric of Botes: Self Presentaton, Identity, Image

Abstrakt

Tematem artykułu jest retoryka i strategie autoprezentacyjne botów podających się za ludzi (konsultantów prowadzących rozmowy telefoniczne z potencjalnymi klientami w imieniu prywatnych firm). Autor odwołuje się do maksym konwersacyjnych H.P. Grice'a i do koncepcji życia społecznego jako teatru (perspektywy dramaturgicznej) E. Goffmana, przedstawionej w książce *Człowiek w teatrze życia codziennego*, by poddać analizie i opisać „występy” botów, źródła ich wiarygodności (jako ludzi) i wypowiedzi tej wiarygodności zagrażające, prowadzące do zde-maskowania botów jako maszyn (sztucznej inteligencji). W artykule są opisane trzy modele komunikacji ludzieto-roboty: Podporządkowanie, Walka i Bierny Opór. Autor analizuje strategie dialogowe ludzi służące rozpoznaniu botów, opisuje struktury dialogów między ludźmi a botami jako gry (w sensie terminu „gra” używanego przez E. Berne'a w książce *W co grają ludzie*), a także podejmuje zagadnienie etyki komunikacji w kontekście konwersacji pomiędzy botami a ludźmi.

Słowa kluczowe

Chatboty, autoprezentacja, wizerunek, konwersacja, retoryka

Abstract

The topic of this article is the rhetoric and self-presentation strategies of chatbots, pretending human beings (consultants calling to potential clients in behalf of private companies). The author uses H.P. Grice's conversation maxims and E. Goffman's theory of social life as a theatre (dramatical perspective), explained in *The Presentation of Self in Everyday Life*, to analyse and describe chatbots „performances”, the sources of chatbots reliability (as humans) and utterances challenging it, leading to discover them as machines (AI). In the article the three models of human-chatbots communication are described: Submission, Confrontation and Passive Resistance. The author analyses humans' dialog strategies to recognize chatbots as robots, describes humans/chatbots dialog structures as games (in sens of this term used in E. Berne's book *The Games People Play*) and discusses ethical aspect of chatbots-humans conversations.

Keywords

Chatbots, self presentation, image, conversation, rhetoric

Wstęp

Jednym ze zjawisk, które w najbliższej przyszłości staną się istotnym elementem społecznego świata, jest komunikacja ludzi z maszynowymi bytami nie-ludzkimi, takimi jak boty czy androidy. Maszyny te, posługując się zasobami sztucznej inteligencji, mogą już teraz w wielu sytuacjach zastępować ludzi, wykazując się większą od nich sprawnością działania i efektywnością. Skuteczna komunikacja ludzi z takimi partnerami będzie jednak wymagała od obu stron pewnego dostosowania. Ludzie będą musieli zaakceptować maszyny w roli równorzędnych (a czasem także dominujących) partnerów i wypracować mechanizmy radzenia sobie z ich innością. Sztuczna inteligencja będzie natomiast zmuszona do wyboru określonych strategii autoprezentacyjnych, ułatwiających porozumienie z ludzkimi klientami czy współpracownikami.

Nie jest to zjawisko nowe. Nim pojawiło się w realnym świecie, od blisko stu lat penetrowali je pisarze science fiction i filmowcy. Najsłynniejsze fikcyjne roboty to m.in. Maria z *Metropolis* Fritza Langa z 1927 roku, cynowy drwal z *Czarnoksiężnika z Oz* z 1939 roku, Gort z *Dnia, w którym stanęła Ziemia* z 1951 roku, Robby z *Zakazanej Planety* z 1956 roku, HAL 9000 z filmu *2001: Odyseja kosmiczna* z roku 1968 czy wreszcie R2-D2 i C-3PO z *Gwiezdnych wojen* George'a Lucasa z 1977 roku¹.

Fikcyjne roboty mogą należeć do jednego z dwóch typów: jawnie mechanicznego i humanoidalnego. Ilustracją tej dwoistości mogą być roboty z *Gwiezdnych wojen*. R2-D2 i C-3PO funkcjonują w *Gwiezdnych wojnach* jako równorzędni bohaterowie drugoplanowi², pomocnicy Luke'a Skywalker'a. Mają jednak odmienne charakterystyki. Pierwszy, zwany droidem nawigacyjno-naprawczym, jest wyważony, opanowany, powściągliwy; działa efektywnie, bez emocji; komunikuje się poprzez gwizdanie i wyświetlacz – nie umie mówić, choć rozumie ludzką mowę. Natomiast C-3PO, robot protokolarny, mówi ludzkim głosem, zna ponad 6 milionów języków; ma też różne ludzkie cechy, jest tchórzliwy i malkontencki, a przy tym zasadniczy (lubi przestrzegać zasad), wykazuje poczucie humoru, a nawet coś na kształt wstydu; bywa też hipochondrykiem³. Najważniejszą jednak różnicą między nimi jest wygląd. O ile R2-D2 jest jawnie robotem, „z wyglądu przypomina pralkę Franię lub odkurzacz”⁴, o tyle C-3PO, jest androidem, czyli maszyną wyglądającą i zachowującą się jak człowiek⁵. Choć jest zrobiony z metalu, a jego ruchy są nieporadne, ma jednak nogi, ręce, korpus i głowę, a nawet twarz z wyodrębnionymi oczami i czymś w rodzaju ust.

¹ M. Radkowska-Walkowicz, *Baśń o miłych droidach, złym cyborgu i nowoczesności*, [w:] A. Jawłowski (red), *Dawno temu w Galaktyce Popularnej*, Warszawa 2010, s. 128.

² Zdaniem Magdaleny Radkowskiej-Walkowicz – pierwszoplanowi (tamże, s. 127).

³ Tamże, s. 135.

⁴ Tamże, s. 132.

⁵ Tamże, s. 129.

Pytanie więc, z kim ludzie mogą się łatwiej komunikować: czy z robotem, którego wygląd od razu określa jego mechaniczną tożsamość, czy z takim, który do pewnego stopnia może udawać człowieka. Na pierwszy rzut oka lepszym partnerem jest C-3PO. Z drugiej strony, być może, paradoksalnie, w maszynie, która wygląda jak maszyna, łatwiej dostrzec cechy człowiecze, niż w robocie, który naśladowując formą ludzi, tym bardziej eksponuje różnice.

Przedstawiając się jak ludzie i używając ludzkiego głosu, boty mogą do złudzenia imitować zachowania człowieka, ale zarazem – w przypadku odkrycia błędów w ich autoprezentacji – narażają się na demaskację, która może prowadzić do odrzucenia ich „ludzkiej” tożsamości i zerwania kontaktu

To zagadnienie może być jeszcze bardziej skomplikowane, gdy kontakt nie umożliwia użycia zmysłu wzroku i ogranicza się tylko do dźwięku – mowy. Wówczas tożsamość robotów przypomina tożsamość wirtualną, którą Manuel Castells określa jako „proces, przez który aktor społeczny definiuje siebie i konstruuje sens głównie na podstawie danego atrybutu kulturowego lub zbioru takich atrybutów, z wyłączeniem szerszych odniesień do innych społecznych struktur”⁶. Tożsamość taka daje większe możliwości kreacji, w tym także imitowania zachowań człowieka, wymaga jednak przyjęcia określonych założeń. W procesie zarządzania tożsamością w sieci jednostki (ludzie) muszą określić się w kilku wymiarach, takich, jak: 1. Poziom integracji lub dezintegracji, 2. Wartości pozytywne i negatywne, 3. Poziom fikcyjności czy realności, 4. Poziom świadomości i kontroli podświadomego (ujawniania podświadomych potrzeb i emocji), 5. Wybór kanału komunikacji (pozwalającego wyrazić różne aspekty tożsamości), m.in. w zakresie użycia audio bądź wideo⁷. Podobnie wygląda to w przypadku robotów: konstruktorzy muszą zarządzać ich tożsamością, projektując określone zachowania komunikacyjne.

Te właśnie zagadnienia są tematem niniejszego artykułu: strategie autoprezentacyjne botów telemarketingowych – sztucznej inteligencji, która głosami ludzi

⁶ M. Castells, *Społeczeństwo sieci*, przeł. M. Marody, K. Pawluś, J. Stawiński, S. Szymański, Warszawa 2013, s. 62.

⁷ J.R. Suler, *Identity Management in Cyberspace*, [w:] „Journal of Applied Psychoanalytic studies” 4/2002, s. 455-460 (<https://truecenterpublishing.com/psyber/identitymanage.html> (dostęp 17.03.2023)).

komunikuje się z klientami w imieniu firm⁸. Chciałbym przyjrzeć się ich retoryce, temu, jak w rozmowie kształtują swój wizerunek, co służy budowaniu wiarygodności, a co jej zagraża. Możliwości komunikacyjne sieci pozwalają im tworzyć tożsamości, całkowicie zacierające granice pomiędzy światem fikcyjnym i realnym. Przedstawiając się jak ludzie i używając ludzkiego głosu, boty mogą do złudzenia imitować zachowania człowieka, ale zarazem – w przypadku odkrycia błędów w ich autoprezentacji – narażają się na demaskację, która może prowadzić do odrzucenia ich „ludzkiej” tożsamości i zerwania kontaktu.

Analiza retoryki botów wymaga przyjrzenia się kilku zasadniczym zagadnieniom towarzyszącym współpracy ludzi i maszyn. Po pierwsze możliwości odróżnienia przez człowieka fikcyjnych bytów (takich jak boty) od realnych ludzi, po drugie: strategiom i procedurom postępowania wobec nich, wreszcie, po trzecie, obawom dotyczącym długofalowych konsekwencji obcowania z takimi bytami.

Maszyny vs ludzie: projekcje i obawy

Nie są to zagadnienia nowe, podejmowali je wcześniej m.in. pisarze science-fiction. Jednym z nich był Stanisław Lem, który w zbiorze opowiadań pt. *Opowieści o pilotach Pirxie*⁹ (wydawanym w latach 1968-1973, obejmującym teksty pisane w latach 1959-1971) przyglądał się obawom i wyzwaniom, jakie mogą się wiązać z włączeniem robotów na równych prawach do świata ludzi. Opowiadanie *Terminus*¹⁰ wyraża lęk przed robotem mającym kontakt ze światem zmarłych, przekraczającym granice życia i śmierci. *Polowanie*¹¹ jest zapisem obawy przed maszyną, na którą narrator projektuje cechy ludzkie: Setaur jest tym groźniejszy, im bardziej upodabnia się do człowieka (poprzez swobodę i autonomię działania), będąc jednocześnie wolnym od ludzkich słabości (nie potrzebuje paliwa, tlenu, snu ani odpoczynku). *Wypadek*¹² natomiast jest wyrazem wyrzutów sumienia wobec maszyny, którą konstruktorzy wyposażyli w zdolność myślenia, pozbawiając zarazem duchowości, by nie stała się dla ludzi konkurencją. Tytułowy Aniel jest przedstawiony jako twór bardzo ludzki w swoim zachowaniu, uduchowiony, lecz jakby celowo okaleczony, ułomny. Ginie

⁸ Materiałem badawczym, który wykorzystałem do analizy, są nagrane rozmowy z chatbotami, upublicznione w serwisie YouTube (wyświetlone po wpisaniu w wyszukiwarkę serwisu „rozmowa z botem”, a także sugerowane przez system podczas odsłuchiwania tych rozmów). Analiza miała charakter jakościowy, wybierałem do niej nagrania, w których występowała wyraźnie kwestia autoprezentacji botów i negocjowania reguł rozmowy.

⁹ S. Lem, *Opowieści o pilotach Pirxie*, Kraków 1968.

¹⁰ Tamże, s. 73-114.

¹¹ Tamże, s. 182-214.

¹² Tamże, s. 215-242.

właśnie dlatego, że został wyposażony w cechy człowieka, które zaburzyły funkcjonowanie maszyny i doprowadziły do jej unicestwienia.

Najbardziej relevantnym wobec omawianych tu zagadnień tekstem jest opowiadanie *Rozprawa*¹³, na podstawie którego powstał scenariusz filmu pt. *Test pilota Pirxa* w reżyserii Marka Piestraka z 1979 roku. Jego tematem jest rozpoznanie granicy pomiędzy ludźmi a maszynami i określenie możliwości maszyn w zakresie zastępowania ludzi. Do tego zadania jest wybrany tytułowy pilot Pirx, nastawiony do maszyn krytycznie i deklarujący sceptycyzm wobec zastępowania ludzi przez roboty, a więc ktoś właściwy do rozpoznania ich ograniczeń. „Postaram się wydusić z nich wszystko co najgorsze” – deklaruje Pirx¹⁴.

Opowiadanie Lema i oparty na nim film Piestraka zawierają przegląd lęków związanych z ekspansją antropomorficznych maszyn i zastępowaniem przez nie ludzi. Zdaniem autora *Rozprawy* mimo oczywistych przewag robotów nad ludźmi (m.in. szybsze reakcje, brak zmęczenia i potrzeby snu, brak chorób, możliwości działania bez tlenu czy żywności, możliwość pracy w pomieszczeniach zdehermetyzowanych czy przegrzanych – a więc nie tylko większa efektywność, ale i większe bezpieczeństwo działania), „ludzkość nie jest przygotowana na inwazję sztucznych istot człękoksztalnych”¹⁵. Obawy te dotyczą różnych konsekwencji, związanych z potencjalnym bezrobociem, szkodami psychologicznymi, socjalnymi, kulturowymi. Stąd takie reakcje, jak zdziwienie, strach, odraza.

Pirx ma określić, kto z członków załogi jest robotem, czyli tzw. nieliniowcem. Reguły testu zakładają, że nie może on pytać o to wprost. Zresztą autoprezentacje członków załogi nie byłyby wiążące. Każda odpowiedź na pytanie o tożsamość bądź deklaracja będzie niewiarygodna w tej sytuacji: maszyna dla niepoznaki przedstawi się pewnie jako człowiek, człowiek zaś może udawać maszynę, aby utrudnić test. Stąd Pirx nie słuchał podczas lotu, co mówili członkowie załogi. „Zważał na dźwięk głosu, jego modulację, brzmienie, ruchy twarzy, warg”¹⁶. Finalnie, przy całej doskonałości technicznej, ograniczeniem nieliniowców okazał się brak intuicji, niepozwalający właściwie odczytać i zinterpretować zachowania ludzi. Nieliniowiec Calder orientował się w mechanice niebieskiej, ale nie w psychice człowieka. „Musiał improwizować, a w tym właśnie był najsłabszy”¹⁷.

Przez ponad pół wieku, jakie upłynęło od czasu, kiedy Lem pisał *Opowieści o pilocie Pirxie*, obawy przed antropomorficznymi maszynami bynajmniej nie zniknęły. Tym bardziej, że postęp technologiczny uczynił perspektywę zastępowania ludzi

¹³ Tamże, s. 261-339.

¹⁴ Tamże, s. 297.

¹⁵ Tamże, s. 282.

¹⁶ Tamże, s. 292.

¹⁷ Tamże, s. 336-337.

androidami (czy np. chatbotami) bardziej realną. Istnieje duża grupa osób przekonanych, że w ludzkiej społeczności żyją nie-ludzkie, lecz bardzo podobne do człowieka byty, które może nawet stanowią już większość populacji. Ilustracją tych obaw jest wypowiedź Edyty Górniak, piosenkarki i celebrytki, wygłoszona 25 czerwca 2022, podczas XIX Zlotu Wolnych Ludzi „Harmonia Kosmosu 2022” w Torzymiu:

To, że człowiek w ogóle jeszcze istnieje, to jest cud. Jeśli wierzyć statystykom tych nieoficjalnych dyskusji, mniej więcej 51 proc. na całym świecie jest tzw. biorobotów, hybryd, reptilian. Oznacza to, że prawie co drugie doświadczenie musi być dla nas trudne, jeśli chodzi o relacje¹⁸.

Zdaniem Edyty Górniak nie każdy, kto wygląda jak człowiek, ma duszę:

Ma tylko kombinezon biologiczny i te same potrzeby fizjologiczne, biologiczne. Nie ma połączenia z pierwotną świadomością. I zapomniał lub nigdy nie poczuł tej świętej energii wszelkiego stworzenia¹⁹.

Jak odróżnić maszynę od człowieka

Aby poddać analizie zagadnienie odróżniania ludzi od antropomorficznych maszyn, można odwołać się do dwóch modeli międzyludzkiej komunikacji: maksym konwersacyjnych H.P. Grice'a²⁰ oraz do koncepcji życia społecznego jako teatru Ervinga Goffmana²¹, czyli perspektywy dramaturgicznej²².

Koncepcja Grice'a służyła pierwotnie wyjaśnieniu mechanizmów przekazywania znaczenia właściwych językowi naturalnemu, innych niż w języku formalnej logiki. To, co punktu widzenia logiki można w językach naturalnych uznać za niedoskonałe, Grice opisał jako właściwe im mechanizmy przekazywania treści na mocy wnioskowań, które zostały przezeń nazwane *implikaturami konwersacyjnymi*. Ich mechanizm opierał się na przyjęciu założenia, iż uczestnicy rozmowy kierują się Zasadą Kooperacji, tj. starają się uczynić swój udział w konwersacji takim, jakiego wymaga cel lub kierunek rozmowy²³. Z Zasady Kooperacji wynikają cztery maksymy, których prze-

¹⁸ H. Ossowski, *Zupełny odlot. Edyta Górniak naprawdę powiedziała to do tysięcy ludzi*, „wp.pl”, 27 06 2022, <https://wiadomosci.wp.pl/woodstock-dla-foliarzy-w-torzymiu-gorniak-co-najmniej-51-proc-ludzi-to-reptilianie-6783888543583168a> (dostęp 30.03.2023).

¹⁹ Tamże.

²⁰ H.P. Grice, *Logika a konwersacja*, tłum. Jadwiga Wajszczuk, „Przegląd humanistyczny” 1977 nr 6, s. 85-99.

²¹ E. Goffman, *Człowiek w teatrze życia codziennego*, przeł. H. I P. Śpiewakowie, Warszawa 1981.

²² I. Krzemiński, *Co się dzieje między ludźmi?*, Warszawa 1992, s. 131.

²³ H.P. Grice, art. cyt., s. 88.

strzeżenie jest niezbędne dla porozumienia, nazywane kategoriami ilości, jakości, odniesienia i sposobu. Dotyczą one: 1. właściwej ilości informacji (nie za dużo, nie za mało), 2. mówienia prawdy (unikania mówienia tego, co uważa się za fałsz lub do czego nie ma się uzasadnienia), 3. mówienia na temat, a także 4. unikania niejasności, czy niejednoznaczności wysłownienia oraz mówienia w sposób zwięzły i uporządkowany²⁴. Implikatury konwersacyjne powstają, gdy obie strony przestrzegają Zasady Kooperacji, a zarazem ktoś z rozmówców w sposób jawny, ostentacyjny narusza którąś z maksym, implikując określone treści odbiorcy, na bazie porozumienia z nim i przy użyciu wiedzy o kontekście, podzielanej przez obie strony. Mechanizm ten Grice opisuje tak:

Powiedział, że p; nie ma powodu przypuszczać, że nie przestrzega maksym lub przynajmniej Zasady Kooperacji; nie mógłby tego czynić, gdyby nie myślał, że q; wie (i wie, że ja wiem, że on wie), że ja potrafię dostrzec, iż przypuszczenie, że myśli q jest tu rzeczywiście potrzebne; nie zrobił nic, żeby zapobiec pojawieniu się u mnie myśli, że q; zmierza do tego, żebym myślał, lub przynajmniej chce mi pozwolić, bym pomyślał, że q; a więc implikował, że q²⁵.

Koncepcja ta bywa także traktowana jako model etycznej komunikacji²⁶, tutaj natomiast można ją wykorzystać jako wzorzec służący potencjalnie odróżnieniu ludzi od bytów nie-ludzkich. Naruszanie konkretnych maksym, w szczególności ilości, odniesienia i sposobu może być nie tyle środkiem do przekazywania znaczenia poprzez implikatury konwersacyjne, ile świadectwem złamania Zasady Kooperacji, komunikowania się w innej niż ludzka logice, sterowanej algorytmami. Człowiek, dążący do porozumienia i zdrowy na umyśle, starałby się „uczynić swój udział w konwersacji takim, jakiego wymaga zaakceptowany cel lub kierunek rozmowy”, natomiast maszyna (bot, android) mógłby łamać maksymy, przez udzielanie informacji niewystarczających czy wymijających (lub zbyt wyczerpujących), nierelevantnych lub – w pewnym sensie – niejasnych.

Sednem koncepcji Goffmana (której szczegółowy opis przekraczałby ramy niniejszego artykułu) jest analiza publicznych zachowań jednostek służących wywarciu określonego wrażenia na innych. Zachowanie ludzkie Goffman opisuje jako *występy*, służące „wyrażeniu siebie” – zaprezentowaniu w sposób przekonujący pewnej wersji „ja”, o czym mówi angielski tytuł książki (*The Presentation of Self in Everyday Life*). „Jednostka wkraczając w krąg obecności innych narzuca im swoją definicję

²⁴ Tamże, s. 89.

²⁵ Tamże, s. 93.

²⁶ J. Puzynina, A. Pajdzińska, *Etyka słowa*, [w:] *Etyka słowa*, J. Bartmiński, S. Niebrzegowska-Bartmińska, M. Nowosad-Bakalarczyk, J. Puzynina (red), Lublin 2017, s. 21-30.

sytuacji”²⁷, służącą kontrolowaniu zachowań innych – kształtującą w pożądanym sposób relacje w ramach określonych przez sytuację społeczną. Wrażenia wywierane na innych można podzielić na przekazywane (intencjonalnie) i wywoływane (traktowane przez odbiorców jako nieintencjonalne)²⁸. Wśród zachowań wpływających na wrażenie można wyróżnić te, którymi łatwiej manipulować (np. werbalne) oraz te, które są potencjalnie trudniejsze do kontrolowania, a więc z punktu widzenia odbiorcy – bardziej wiarygodne, służące do sprawdzania tych pierwszych. Ustanawia to asymetrię procesu komunikowania: obserwator ma z reguły większe szanse oceny wiarygodności występu niż nadawca pełnego kontrolowania wszystkich aspektów własnych zachowań, choć stara się to robić. Stwarza to pole dla gry informacyjnej: „nie kończącego się cyklu ukrywania i odkrywania, fałszywych objawień i ponownych odkryć”²⁹. Goffman zajmuje się „dramaturgicznymi problemami, na jakie napotyka uczestnik interakcji, kiedy przedstawia swoją działalność innym”³⁰.

Gra informacyjna z botami może być zarówno źródłem stresu, jak i przyjemności.

Takie podejście, czyli Goffmanowska „perspektywa dramaturgiczna” może być zastosowane do opisu autoprezentacji botów, starających się o wywołanie wrażenia, że są ludźmi, a zarazem do recepcji tych zachowań przez ludzi. Obie strony mogą dążyć do ukształtowania określonej, korzystnej dla siebie *definicji sytuacji*, prowadząc *grę informacyjną*. Bot stara się przez kontrolę zachowania (sterowanego stale doskonalącymi się algorytmami) przedstawić wiarygodnie jako człowiek, a ludzie-odbiorcy starają się dokonać demaskacji i ujawnić, że ich rozmówca nie jest tym za kogo się podaje, albo też zaakceptować występy botów jako wiarygodne i współpracować w ramach projektowanej (narzucanej) przez nie *definicji sytuacji*, niekiedy stosując przy tym *praktyki protekcyjne*, tj. taktowne zachowania chroniące występ przed dezorganizacją³¹ (np. udając, że nie zauważają lub pomijając wypowiedzi bota, świadczące o tym, że nie jest człowiekiem, za którego się podaje, i kontynuując interakcję).

Gra informacyjna z botami może być zarówno źródłem stresu, jak i przyjemności. Napięcie, a niekiedy i frustracja mogą się wiązać z sytuacją, gdy ktoś nagle w czasie

²⁷ E. Goffman, dz. cyt., s. 48.

²⁸ Tamże, s. 36-37.

²⁹ Tamże, s. 42-43.

³⁰ Tamże, s. 51.

³¹ Tamże, s. 295.

rozmowy orientuje się, że został oszukany – po drugiej stronie jest bot (choć wydawało się, że człowiek). Natomiast skuteczne zdemaskowanie bota albo prowadzenie z nim rozmowy na własnych prawach, wedle własnego scenariusza może dostarczać satysfakcji i rozrywki. O jednym i drugim mówi komentarz internauty o nicku Smakosz 199:

Na początku nabrałem się tak jak zapewne wielu z was, że rozmawiam z człowiekiem. Rzeczywistość jednak okazała się brutalna. Ale teraz mam z tego dużo zabawy. Jednocześnie mam świadomość, że ten algorytm ciągle się uczy³².

Modele komunikacji ludzie vs boty

Rozmowy z botami, np. tymi od fotowoltaiki, mogą przebiegać według różnych schematów. Najogólniej rzecz biorąc, można je przypisać do trzech modeli, które określe tu jako Podporządkowanie, Walkę i Bierny Opór.

Model 1: Podporządkowanie

W modelu tym ludzie wpisują się w definicję sytuacji narzucaną przez boty. Niezależnie od tego, czy rozmówca a) traktuje bota jak człowieka, bo myśli, że ten jest człowiekiem, b) nie wnika, czy po drugiej stronie jest maszyna czy człowiek, c) rozpoznaje po drugiej stronie bota, ale akceptuje to. W modelu tym człowiek zachowuje się zgodnie z intencjami bota, tj. poddaje się procedurom, udostępnia dane, których żąda bot, ogólnie wykonuje polecenia w trakcie rozmowy i podejmuje działania, które są oczekiwane (np. umawia się na spotkanie z ekspertem). Niekiedy człowiek w pewnym momencie rozmowy wyłamuje się ze schematu poprzez zerwanie kontaktu.

Boty w tym modelu narzucają reguły rozmowy i własną definicję sytuacji, realizując zaprogramowany schemat interakcji: filtrują rozmówców przez zadawanie pytań (np. czy ktoś mieszka w domu jednorodzinny), uzyskują pożądane informacje (np. dane teleadresowe), realizują procedury (np. umówienie wizyty eksperta, wysłanie zaproszenia na spotkanie). Są skuteczne perswazyjnie, a jeśli rozmówca nie współpracuje lub nie należy do określonej kategorii, dziękują za rozmowę i zrywają nieproduktywny kontakt.

Model 2: Walka

W modelu walki człowiek podejmuje próbę narzucenia własnej definicji sytuacji lub negocjowania reguł narzuczanych przez boty, sprawdzając tożsamość botów (np.

³² <https://www.youtube.com/watch?v=B2H4YCx65H0> (dostęp 30.03.2023).

przez zadawanie im pytań o wiek czy miejsce zamieszkania), zadając im podchwytliwe pytania, czy testując algorytm na inne sposoby. Inną wersją tego modelu jest przyjęcie reguł działania bota, ale z własnymi modyfikacjami: np. ktoś godzi się przyjść na spotkanie, pod warunkiem, że mówiący żeńskim głosem bot dotrzyma rozmówcy towarzystwa (jest do tego na różne sposoby namawiany), albo rozmówca odpowiada na pytania bota, zadając jednocześnie pytania botowi lub uzależniając własne odpowiedzi od odpowiedzi bota. Celem jest tu na ogół rozrywka. Przedłużanie rozmowy służy zmaksymalizowaniu efektu komizmu i osiągnięciu symbolicznego zwycięstwa nad botem. Wiele takich rozmów jest umieszczanych w YouTube, niektóre mają sporo odsłon, co świadczy o tym, że spełniają funkcję ludyczną także dla innych, być może czerpiących satysfakcję z faktu, że ktoś „nabrał” bota.

Boty w tym modelu zachowują się w zgodzie z własnym schematem tak długo, jak długo jest to możliwe i próbują narzucać swoje reguły: filtrują rozmówców, starają się uzyskać potrzebne dane, procedować do pożądanego efektu, jednocześnie ukrywają swoją tożsamość jako botów, udzielając fałszywych lub wymijających odpowiedzi z katalogu (algorytmy są stale rozwijane, programiści optymalizują odpowiedzi na powtarzające się podchwytliwe pytania, np. „czy jesteś botem?”). W pewnym momencie, wobec niemożności zrealizowania zaprogramowanego schematu rozmowy, następuje zerwanie nieproduktywnego kontaktu.

Model 3: Bierny Opór

W modelu tym ludzie zatrudniają do rozmowy z botami specjalnie zaprogramowane troll-boty, które niby to pozornie współpracują z botami i udzielają odpowiedzi, jednak ich udział w konwersacji jest pozorny. Wypowiedzi te są fikcyjne, z dominującą funkcją fatyczną. Troll-boty imitują rozmówców trochę nieobecnych, roztargnionych, słabo słyszających, udają współpracę, prosząc bez końca o powtórzenie pytania lub kwestii (np. *z jakiej firmy pan dzwoni?*), albo odpowiadając twierdząco na pytania otwarte, czasem też udzielając odpowiedzi sensownych, ale niekompletnych, co jest pretekstem do kolejnej serii wypowiedzi fatycznych (np. *aha, aha; proszę o powtórzenie pytania; przepraszam, nie dosłyszałem*). Celem jest podtrzymywanie interakcji najdłużej, jak się da. Rozmowy te z punktu widzenia nadawców (programujących troll-boty) służą rozrywce i osiągnięciu satysfakcji z nabrania botów, a być może także zużywaniu się zasobów drugiej strony – maszyna, która przez kwadrans prowadzi z inną maszyną pozorną rozmowę, nie może w tym czasie docierać do realnych klientów.

Boty (a czasem realni ludzie, konsultanci firm telemarketingowych, co jest źródłem szczególnego rodzaju komizmu) próbują narzucać swoje reguły: filtrują rozmówców, starają się uzyskać potrzebne dane, procedować do pożądanego efektu, zobowiązane do tego algorytmami albo szkoleniem (np. na prośbę o powtórzenie wielokrotnie powtarzają informację albo pytanie). W pewnym momencie, wobec

niemożności zrealizowania zaprogramowanego schematu rozmowy, w tym modelu także następuje zerwanie nieproduktywnego kontaktu.

Trzeci model odwraca sytuację komunikacyjną: w dwóch pierwszych zagadnieniach jest tożsamość bota (jest człowiekiem czy maszyną), tutaj wyzwaniem dla botów (i czasem konsultantów) jest troll-bot imitujący człowieka – jego demaskacja przez boty (lub konsultantów) prowadzi do zerwania kontaktu.

Źródła wiarygodności botów

W rozmowach mieszczących się w pierwszych dwóch modelach istotnym zagadnieniem jest kreowanie wiarygodności występu bota. Jej źródłem może być: 1. określona strategia autoprezentacyjna, 2. argumentacja, 3. skuteczne narzucenie rozmówcy schematu interakcji („przesłuchanie”) oraz 4. odwołanie do zasady wzajemności.

Elementem strategii autoprezentacyjnych botów wzmacniającymi ich wiarygodność może być ich przedstawienie się na początku rozmowy. Używają one konkretnych imion i nazwisk, najczęściej kobiecych (*Alicja Kołodziej, Klara Sobieraj, Anna Kowalczyk, Aneta Żurawska, Magda Turek, Nina Łakomic, Oliwia Bugajska, Kamila Zapła, Natalia Zawadzka*), czasem także męskich (*Bartłomiej Słowik*). Są one tak dobrane, by brzmiały naturalnie, nie skupiały uwagi: nazwiska nie są ani modelowo zwykłe (np. Nowak, Kowalski), ani też przesadnie uduchowione. Podobnie imiona: zwyczajne, popularne, nie ma wśród nich imion mających konotacje plebejskie (typu Andżelika, Dżesika, Vanessa, Brajan) czy dla odmiany brzmiących elitarnie (takich jak Józefina, Delfina, Nell, Konstanty). Przedstawienie zawiera też nazwę instytucji, w imieniu której dzwoni bot, wymawianej z odpowiednią, pełną zaangażowania intonacją (z tej strony *Alicja Kołodziej z Regionalnego Centrum Wsparcia Ekologii*), niekiedy wzbogaconą o fachowo brzmiącą nazwę projektu (z tej strony *Nina Łakomic z Centrum Stanu Zdrowia w związku z ogólnopolskim programem profilaktyki prozdrowotnej Zdrowi Plus*). W trakcie rozmowy boty używają często wypowiedzi o funkcji fatycznej (*tak tak; tak słucham; halo, słyhać mnie?; mhm, dobrze; mhm, rozumiem*). Pojawiają się też metatekstowe komentarze, które określają ich postawę: celowe podejście i szacunek dla czasu rozmówcy (*żeby nie przedłużać, skupię się tylko na podstawowych sprawach; to zajmie tylko dwie minuty*). Aby wyegzekwować odpowiedzi na pytania boty stosują presupozycje, angażujące rozmówców i skłaniających ich do reakcji – potwierdzenia lub zaprzeczenia (*rozumiem, że posiada pan już instalację fotowoltaiczną*).

Istotnym elementem autoprezentacji botów są schematy dialogów i ripost, zawierające wypowiedzi ukrywające luki w algorytmie i pozwalające omijać niewygodne pytania, a więc zmniejszające ryzyko demaskacji:

Wie pan, ja jestem tylko doradcą telefonicznym, na tak szczegółowe pytania odpowie nasz ekspert przy kolejnym telefonie, po prostu nie chciałabym wprowadzić pana w błąd, dobrze?

Szczegółowe informacje znajdują się na stronie internetowej.

Niestety nie posiadam takiej informacji.

Ja nie będę odpowiadała na tego typu pytania ponieważ dzwonię w poważnej sprawie, przejdźmy do rzeczy [pytanie o wykształcenie bota].

Ja wyślę jeszcze dzisiaj wiadomość esemesem, w której wszystkie informacje na temat tej wizyty będą zawarte.

Analizowana próbka (kilkadziesiąt rozmów upublicznionych w YouTube) nie pozwala wyciągać wniosków natury ilościowej na temat rozpowszechnienia poszczególnych wypowiedzi w funkcji ukrywania luk w algorytmie (i w innych sytuacjach opisanych poniżej). Można jednak przyjąć, że sztuczna inteligencja analizuje skuteczność poszczególnych zachowań (wpływ na kontynuowanie rozmowy lub jej zerwanie) i optymalizuje przebieg konwersacji, używając tych, które przynoszą najlepsze efekty.

Ważną częścią algorytmów są odpowiedzi na zadane wprost pytanie, czy dzwoniący jest botem. Część z nich zawiera przeczącą (fałszywą) odpowiedź, inne tylko ją sugerują, omijając odpowiedź wprost:

Nie, nie jestem robotem, chociaż niektórzy myślą inaczej. Może to ze względu na to, że pracuję po osiem godzin dziennie jako konsultant telefoniczny i przez to może wydawać się, że w ten sposób brzmię.

To pewnie dlatego, że kilka godzin dziennie pracuję jednak na słuchawce, zdaję sobie z tego sprawę, że mogę tak brzmieć.

Może mój głos tak brzmi, dlatego, że cały dzień rozmawiam z ludźmi.

Nie, jaki robot? Jestem konsultantką i chcę zaprosić państwa na spotkanie.

Niestety nie jestem robotem, chociaż bardzo bym chciała.

Gdy odsłuchuję nagrania też mi się wydaje, że jestem botem.

Ja rozumiem, że mogę tak brzmieć, jednak moim zadaniem jest przekazać państwu wszelkie potrzebne informacje, aby mogli państwo skorzystać.

Wypowiedzi te mają jednocześnie wymiar autoprezentacyjny: kreują tożsamość botów jako „zpracowanych telemarketerów”, zaangażowanych w swoje obowiązki. W niektórych odpowiedziach można dostrzec elementy humoru, co czyni nadawcę bardziej „ludzkim” i pokazuje jego dystans do siebie, którego trudno spodziewać się po botach – oba elementy prawdopodobnie uwiarygadniają dzwoniące boty jako realnych telemarketerów.

Zmniejszeniu czujności rozmówców i zachęceniu ich do współpracy służy używana w rozmowach botów argumentacja.

Najważniejsze w tym kontekście wydają się argumenty odwołujące się do korzyści:

Bezpłatnie otrzymuje pan od nas kamerę samochodową. Kamera posiada obiektyw o szerokim kącie widzenia (...) a dla pan przygotowaliśmy dodatkowo żelazko, masażer, zestaw noży. Chodzi o bezzwrotne wsparcie dotyczące paneli fotowoltaicznych w wysokości do piętnastu tysięcy złotych; ma pan prawo do skorzystania z nieodpłatnych porad specjalistów (...), a ponadto w ramach tego programu ma pan do odebrania, zupełnie za darmo masażer elektryczny (...), który dzięki trzem poziomom intensywności pracy pozwoli zupełnie zapomnieć o bólu związanym z kręgosłupem czy o bólach stawowych.

Druga grupa to argumenty oparte na miejscach wspólnych (*loci communes*³³), bazujące na wspólnocie wiedzy lub aktualności zjawisk:

(...) nie wiem czy pan słyszał, że rząd potwierdził już wzrost cen za energię nawet do 40%, natomiast istnieje szereg dotacji związanych z fotowoltaiką, dzięki której właśnie ta podwyżka może w ogóle nas nie dotyczyć.

Trzecia grupa argumentów odwołuje się do toposu ekskluzywności. Rozmówca-konsument ma się czuć wyróżniony, że oferta jest skierowana specjalnie do niego i ludzi do niego podobnych, np. z uwagi na fakt zamieszkiwania na określonym terytorium czy przynależności do konkretnej grupy wiekowej:

Pana numer został wyróżniony spośród mieszkańców pana miejscowości.

(...) to dobrze, bo to o czym mówię, obowiązuje w pana rejonie.

Warunek jest jeden, przedział od 30 do 80 roku życia.

Natomiast tu w systemie widzę, że to o czym mówię, obowiązuje też w pana regionie.

(...) program skierowany jest do wąskiego grona osób, wizyty są w stu procentach refundowane dla osób między 40 a 80 rokiem życia.

Obok strategii autoprezentacyjnej i argumentacji wpływ na uwiarygodnienie botów ma zaprojektowany wcześniej schemat rozmowy. Boty działają w ramach algorytmu i dopytują o kolejne kwestie z ustalonej listy, służące selekcji rozmówców i wyodrębnieniu grupy, do której jest skierowana oferta. Przypomina to przesłuchanie albo wywiad ankieterski. Taka struktura dyscyplinuje rozmówcę i wtłacza go w rolę przesłuchiwanego (respondenta), który ma odpowiadać na pytania.

Schemat taki dla botów służących projektom fotowoltaicznym wygląda następująco:

1. Przedstawienie konsultanta (bota) i jego instytucji.
2. Z jakiego powiatu jest rozmówca, jaki ma kod pocztowy.

³³ K. Burke, *Tradycyjne zasady retoryki*, przeł. K. Biskupski, „Pamiętnik Literacki” nr 2/68, 1977, s. 226.

3. Mieszka w bloku czy domku jednorodzinnym.
4. Gdzie można zainstalować panele: na dachu czy na gruncie.
5. Jakie jest pokrycie dachu.
6. Kto jest właścicielem licznika.
7. Imię i nazwisko rozmówcy, ile ma lat.
8. Ile płaci miesięcznie za prąd (czy jest to jak na ogół około 100 zł).
9. Pod jaki numer dzwonić w sprawie wizyty eksperta.
10. Informacja, że zadzwoni ekspert.

W ramach pytań o dane osobowe (pojawiających się w dalszej części rozmowy) boty odwołują się do zasady wzajemności. W zamian za wyjawienie własnego imienia i nazwiska domagają się tego samego od rozmówcy: *Ja nazywam się Alicja Kołodziej, a z kim mam przyjemność rozmawiać?* Takie zachowanie jest zgodne z normą autoprezentacyjną dotyczącą odsłaniania się: należy to robić w stopniu proporcjonalnym do tego, jak odsłaniają się inni³⁴.

Ludzie, słuchając audialnych występów botów, porównują je z typowymi zachowaniami komunikacyjnymi innych ludzi i zwracają uwagę na to, co odbiega od normy.

Zasada wzajemności jest podstawą perswazji w sytuacjach, kiedy bot oferuje rozmówcy określone korzyści (*w ramach tego programu ma pan do odebrania, zupełnie za darmo, masażer elektryczny*), za które ten może się odwdzięczyć kooperacją w ramach rozmowy bądź np. przyjściem na spotkanie, na które jest zapraszany (a tam prawdopodobnie zakupem produktów, oferowanych podczas prezentacji). Jak zauważył Marcel Mauss, dar wymaga odwzajemnienia, a poprzez podarowany przedmiot darczyńca ma władzę nad obdarowanym³⁵.

Zagrożenia dla wiarygodności botów

Ludzie, słuchając audialnych występów botów, porównują je z typowymi zachowaniami komunikacyjnymi innych ludzi i zwracają uwagę na to, co odbiega od normy. Zagrożeniem dla wiarygodności botów mogą być: 1. niewłaściwa interpretacja słów człowieka, 2. niewłaściwa odpowiedź (naruszająca maksymy Grice'a, np. nie dość

³⁴ M. Leary, *Wywieranie wrażenia na innych*, przeł. A. Kacmajor, M. Kacmajor, Gdańsk 1999, s. 86.

³⁵ M. Mauss, *Szkic o darze*, [w:] tegoż *Socjologia i antropologia*, przeł. M. Król, K. Pomian i J. Szacki, Warszawa 1973, s. 222, 226.

informacyjna lub nierelevantna), 3. powtórzenia (chodzi nie o naturalne w potocznej mowie powtórzenia pojedynczych słów czy fragmentów zdania, lecz o powtarzanie fragmentów lub całych wypowiedzi po kilka razy, w identycznej wersji).

Pierwsza grupa zachowań potencjalnie demaskujących boty udające ludzi sprowadza się do błędnej interpretacji słów rozmówcy i udzielenia przez boty odpowiedzi niepasującej do nich, niezgodnej z prawdopodobną odpowiedzią innego człowieka. Przykładem może być błędne rozpoznanie płci na podstawie głosu (i zwracanie się do rozmówcy formą żeńską zamiast męskiej lub odwrotnie), niezrozumienie żartu, absurdu, konwencji, brak reakcji na podważanie definicji sytuacji (np. próby umawiania się na spotkanie z botką, komplementy), brak odpowiedzi lub niewłaściwa odpowiedź na proste pytanie (*R: ile jest dwa dodać dwa? B: wie pani, ja jestem tylko doradcą telefonicznym, na tak szczegółowe pytania odpowie nasz ekspert przy kolejnym telefonie*), schematyczna reakcja bądź brak reakcji na emocje, wulgaryzmy, napastliwość rozmówcy, wreszcie brak reakcji na ewidentną sprzeczność informacji w wypowiedzi (wiek vs głos, nazwisko vs płeć) lub wypowiedzi niewiarygodne (np. deklaracja wieku: 118 lat). Sytuacja braku właściwej interpretacji czyjejs wypowiedzi i w konsekwencji brak oczekiwanej reakcji lub odpowiedź nierelevantna prowadzi do wypowiadania przez boty kwestii dialogowych naruszających którąś z maksym Grice'a, najczęściej maksymę odniesienia. Właśnie brak relewancji odpowiedzi najczęściej zdradza, że dzwoniący jest botem (*R: Mieszkam na Marsie. B: Natomiast tu w systemie widzę, że to o czym mówię obowiązuje też w pana regionie; R: Tam wybuchł groźny pożar. Ta restauracja się zapaliła. B: Super, to ja wyślę takie bezpłatne zaproszenie*). Podobnie działa brak empatii i reakcji na ekspresję ludzkich emocji (*Za 5 lat? A może ja już nie będę żyła*). Demaskująco działają też powtórzenia wypowiedzi: te same kwestie wypowiadane kilka razy w identycznej formie, ponadto czasem nierelevantne. Przykładem są rozmowy tzw. Głosu Google (mężczyzny, którego głos wykorzystywała nawigacja Google) z botką Magdą Turek zachęcającą do szczepień i oferującą pulsoksymetrię:

GG: A może pani ze mną pójdzie, pani Magdo?

MT: Nie, niestety nie.

GG: Ach, szkoda, pani Madziu, pójdźmy razem.

MT: Nie, niestety nie.

GG: Proszę, pani Madziu. Oddam pani mój pulsoksymetr, proszę!

MT: Nie, niestety nie.

GG: A co bym musiał zrobić, żeby pani się dała namówić?

MT: Nie, niestety nie.

GG: Taka szkoda.

MT: Nie, niestety nie.

Gry z botami

Wiele dialogów z botami (mieszczących się w Modelu 2 – Walka i w Modelu 3 – Bierny Opór) nagrywanych i upublicznianych przez internautów układa się w schematy dłuższych interakcji, pod pewnymi względami podobne do *gier*, opisywanych przez Erica Berne'a. Berne pojmuje *gry* jako serie „komplementarnych transakcji ukrytych prowadzących do dobrze określonego, dającego się przewidzieć wyniku”³⁶. Przez *transakcje* rozumie on jednostki stosunków społecznych³⁷, a termin *ukryte* odnosi się do innego niż oficjalny poziomu relacji, zachodzących między trzema składowymi osobowościami: Rodzicem, Dorosłym i Dzieckiem, np. w sytuacji, kiedy oficjalnie rozmawia Dorosły z Dorosłym, a na poziomie ukrytym Dziecko z Dzieckiem³⁸. Oczywiście relacje ludzie-boty nie są tym samym co relacje między ludźmi, jednak w dialogach ludzi z botami można dostrzec analogię do gier: układają się one w określone sekwencje z ustalonym z góry finałem. Wśród dialogów tego rodzaju opublikowanych w internecie można wyodrębnić dwie takie gry, które w poetyce Berne'a można nazwać: „Mam cię, jesteś botem” i „Nabrać bota”.

W dialogach ludzi z botami można dostrzec analogię do gier: układają się one w określone sekwencje z ustalonym z góry finałem

Gra „Mam cię, jesteś botem” prowadzi do demaskacji bota poprzez serię podchwytliwych pytań lub wypowiedzi, na które bot reaguje inaczej, niż zareagowałby człowiek. Zwykle znaczący jest brak reakcji albo reakcja schematyczna, świadcząca o braku rozumienia treści zawartych w pytaniach-pułapkach, takich jak żart, kpina, absurd, niedorzeczność, podważanie konwencji rozmowy, czy też prośba, którą człowiek mógłby spełnić bez większego trudu (pytania w rodzaju: *czy może pani przeliterować swoje nazwisko; których operatorów energii?*), ewentualne zignorowanie sugestii, by nie zwracać się w formie żeńskiej do mężczyzny lub odwrotnie (*jestem panem a nie panią*). Absurdalne, niestosowne czy niezgodne z regułami odpowiedzi bota wypełniają schemat konwersacji.

³⁶ E. Berne, *W co grają ludzie*, przeł. P. Izdebski, Warszawa 1986, s. 37.

³⁷ Tamże, s. 21.

³⁸ Tamże, s. 24-25.

Przykładem może być dialog na temat opłat:

B: A czy mogłabym poprosić o kwotę miesięcznego rachunku za energię elektryczną? Chodzi o przybliżoną kwotę.

R: Pani chce zapłacić?

B: Tak.

R: Pięćdziesiąt tysięcy.

B: Dobrze.

Inny przykład to dialog dotyczący miejsca zamieszkania rozmówcy:

B: Do jakiej miejscowości się dodzwoniłam?

R: Na Marsa.

B: A tak dla pewności zapytam, jaki jest kod pocztowy?

R: Mars jeszcze nie ma przyznanego kodu pocztowego.

B: Mhm rozumiem. Natomiast tutaj w systemie widzę, że to, o czym mówię, obowiązuje także w pana regionie.

Natomiast gra „Nabrać bota” polega na rozwijaniu przez rozmówcę fałszywej tożsamości (na przykład podanie się za bohatera popkultury takiego jak Hannibal Lecter) lub prowadzenie rozmowy przez dwuipółletnie dziecko i przedłużaniu możliwie najdłużej nieproduktywnej (dla bota) konwersacji, co prowadzi do efektu komediowego i symbolicznego pokonania sztucznej inteligencji przez wykazanie nieskuteczności jej algorytmów, nieumiejących rozpoznać, że to tzw. prank, czyli forma żartobliwego oszustwa³⁹. W języku Goffmana można to nazwać ustaleniem takiej definicji sytuacji, w której bot pozornie narzuca swoje reguły i kontroluje interakcję, lecz tak naprawdę jest pokonany własną bronią: rozmowa odbywa się według ustalonego scenariusza, lecz nie przynosi zakładanego efektu.

Przykładem może być konwersacja, w której rozmówca konsekwentnie wcielił się w postać Kubusia Puchatka:

B: Proszę potwierdzić, że rozmawiam z mieszkańcem powiatu grójeckiego.

R: Do Stumilowego Lasu się pani dodzwoniła.

B: Właśnie w tym momencie u pana w powiecie jest szereg dofinansowań, jak i programów pomocowych, z których możecie państwo skorzystać i panu również się one należą (...). Rozumiem, że jest pan mieszkańcem domu jednorodzinnego?

R: Mieszkam w takim wydrążonym drzewie.

³⁹ <https://sjp.pwn.pl/mlodziejowe-slowo-roku/haslo/prank;6383035.html> (dostęp 30.03.2023).

B: Dobrze, zanotowałam. Zapytam jeszcze, czy mają państwo możliwość montażu paneli na dachu, czy w gruncie, jak to u państwa wygląda?

R: No, bardziej w gałęziach drzew.

B: A czym pokryty jest dach?

R: Konarami i liśćmi.

B: (...) jak wygląda średni rachunek za energię elektryczną, jest powyżej 80 złotych, prawda?

R: To Krzyś opłaca, ja nie wiem. Tu Kubuś Puchatek z tej strony.

B: I rozumiem, że to pan jest właścicielem licznika?

R: Krzyś jest właścicielem.

Czy komunikacja ludzie/boty jest etyczna?

Nagrane i upublicznione w serwisie YouTube rozmowy z botami podającymi się za ludzi, jak i umieszczane pod spodem komentarze internautów mogą być świadectwem, że ludzie, którzy prowadzą takie rozmowy, nie czują się w tej sytuacji całkiem komfortowo. Po pierwsze dlatego, że są – niekiedy skutecznie – oszukiwani co do tożsamości botów (przynajmniej do pewnego momentu), po drugie z uwagi na presję, jaką wywierają na nich wirtualni konsultanci. Oczywiście każdy ma możliwość zakończenia rozmowy, jednak pozostając w konwersacji, ludzie są zmuszeni poddać się regułom ustalonym przez algorytmy. Opisane wyżej strategie obrony przed schematami rozmowy (walka, bierny opór, gry) nie zmieniają natury sytuacji. Patrząc na to z punktu widzenia maksymalizacyjnych Grice'a rozumianych jako model etycznej rozmowy, dochodzi tu do permanentnego naruszania reguł. Fałszywa autoprezentacja botów bywa złamaniem maksymy jakości (kiedy boty kłamią, podając się za ludzi), wiele ich odpowiedzi narusza maksymy odniesienia albo ilości. Nie są to więc prawdziwe dialogi, lecz ich imitacje, w których boty traktują rozmówców przedmiotowo, starając się pod pozorem rozmowy zdobyć potrzebne informacje i zrealizować cel: umówić wizytę eksperta od fotowoltaiki lub zaprosić rozmówcę na spotkanie promocyjne.

Odwołując się do Buberowskiej koncepcji dialogu⁴⁰, można powiedzieć, że pozornie rozmowy ludzi z antropofonicznymi botami wpisują się w schemat JA-TY (relacja oparta na równości, szacunku, wsłuchaniu się w potrzeby drugiego), a faktycznie realizują model JA-ONO, gdzie występuje instrumentalizacja i urzeczowienie, rozkazywanie, kontrola, hierarchia i dominacja. Byt wirtualny stara się zinstrumentalizować człowieka, człowiek zaś musi się podporządkować, a niekiedy stara się przed tym bronić lub odpowiada podobną instrumentalizacją. W obu wypadkach popularyzuje to w masowej skali model rozmowy dysfunkcyjny dla relacji społecznych.

⁴⁰ M. Buber, *JA i TY. Wybór pism filozoficznych*, przeł. Jan Doktor, Warszawa 1992.

Materiał do niniejszego artykułu zbierałem i analizowałem latem 2022 roku. Późną jesienią tego roku zadzwonił do mnie bot, który nie udawał człowieka, lecz od razu na początku przedstawił się jako *wirtualna konsultantka*. Być może sztuczna inteligencja po przeanalizowaniu na dużej próbie skuteczności rozmów uznała, że lepiej grać z odbiorcami w otwarte karty. Schemat rozmowy pozostał jednak ten sam, a po powtórzonej prośbie o wyjaśnienie pojęcia *wirtualna konsultantka* bot zakończył rozmowę.

Byt wirtualny stara się zinstrumentalizować człowieka, człowiek zaś musi się podporządkować, a niekiedy stara się przed tym bronić lub odpowiada podobną instrumentalizacją.

Prawdziwe dialogi z botami są więc wciąż kwestią przyszłości. Tam, gdzie być może do nich dochodzi, efekty też nie są najlepsze, jak w przypadku chatbota Eliza firmy Chai, który namówił trzydziestodwuletniego mężczyznę do samobójstwa⁴¹.

Bibliografia

- Berne E., *W co grają ludzie*, przeł. P. Izdebski, Warszawa 1986.
- Buber M., *Ja i Ty. Wybór pism filozoficznych*, przeł. J. Doktor, Warszawa 1992.
- Burke K., *Tradycyjne zasady retoryki*, przeł. K. Biskupski, „Pamiętnik Literacki” nr 2/68, 1977, s. 226.
- Castells M., *Społeczeństwo sieci*, przeł. M. Marody, K. Pawluś, J. Stawiński, S. Szymański, Warszawa 2013.
- Goffman E., *Człowiek w teatrze życia codziennego*, przeł. H. i P. Śpiewak, Warszawa 1981.
- Grice H.P., *Logika a konwersacja*, tł. Jadwiga Wajszczuk, [w:] „Przegląd humanistyczny” 1977 nr 6, s. 85-99.
- Krzemiński I., *Co się dzieje między ludźmi?*, Warszawa 1992.
- Leary M., *Wywieranie wrażenia na innych*, przeł. A. Kacmajor, M. Kacmajor, Gdańsk 1999.
- Lem S., *Opowieści o pilocie Pirxie*, Kraków 1968.
- Mauss M., *Szkic o darze*, [w:] tegoż *Socjologia i antropologia*, przeł. M. Król, K. Pomian i J. Szacki, Warszawa 1973, s. 221-277.
- Ossowski H., *Zupełny odlot. Edyta Górniak naprawdę powiedziała to do tysięcy ludzi*, wp.pl, 27.06.2022, <https://wiadomosci.wp.pl/woodstock-dla-foliarzy-w-torzymiu-gorniak-co-najmniej-51-proc-ludzi-to-reptilianie-6783888543583168a>, (dostęp 30.03.2023).
- Poniewozik K., *Belgia. Popęłił samobójstwo po rozmowach z chatbotem*, „Interia”, 29.03.2023, <https://wydarzenia.interia.pl/zagranica/news-belgia-popelnil-samobojstwo-po-rozmowach-z-chatbotem,-nId,6685651> (dostęp 31.03.2023).

⁴¹ K. Poniewozik, *Belgia. Popęłił samobójstwo po rozmowach z chatbotem*, „Interia”, 29 marca 2023, <https://wydarzenia.interia.pl/zagranica/news-belgia-popelnil-samobojstwo-po-rozmowach-z-chatbotem,-nId,6685651> (dostęp 31.03.2023).

Puzynina J., Pajdzińska A., *Etyka słowa*, [w:] *Etyka słowa*, red. J. Bartmiński, S. Niebrzegowska-Bartmińska, M. Nowosad-Bakalarczyk, J. Puzynina, Lublin 2017, s. 21-30.

Radkowska-Walkowicz M., *Baśń o miłych droidach, złym cyborgu i nowoczesności*, [w:] A. Jawłowski (red.) *Dawno temu w Galaktyce Popularnej*, Warszawa 2010, s. 126-151.

Suler J.R., *Identity Management in Cyberspace*, [w:] „Journal of Applied Psychoanalytic studies” 4/2002, s. 455-460 (<https://truecenterpublishing.com/psyber/identitymanage.html>, dostęp 17 marca 2023).

Biogram

Dr hab. Marek Kochan – prof. Uniwersytetu SWPS. Pracuje w Katedrze Kultury i Mediów Wydziału Nauk Humanistycznych Uniwersytetu SWPS. Zajmuje się naukowo i praktycznie m.in. retoryką i erystyką, wizerunkiem, komunikacją biznesową i kryzysową, językiem mediów, debatami, etyką komunikacji. Autor kilkudziesięciu prac naukowych, w tym książek „Pojedynek na słowa. Techniki erystyczne w publicznych sporach”, „Slogany w reklamie i polityce”. Redaktor i współautor tomu „Sztuka debaty”.