Anna Dulny-Leszczyńska

Absolwentka kierunku Media Interaktywne i Widowiska na Uniwersytecie im. Adama Mickiewicza oraz studiów podyplomowych user experience design Uniwersytetu SWPS. Na co dzień projektantka doświadczeń użytkownika w agencji produktowej.

5 Granice empatii sztucznej inteligencji

interakcja człowiek – komputer

interakcja człowiek – robot

boty

roboty

chat

empatia

Replika

Boty są coraz sprawniejszymi towarzyszami codziennych aktywności. W naturalny sposób są traktowane jak równorzędni aktorzy społeczni. Niektóre z nich stają się nawet przyjaciółmi swoich użytkowników. W ten sposób dochodzi do urzeczywistnienia motywów znanych z kultury popularnej. Czy jesteśmy w stanie przewidzieć konsekwencje tak bliskich relacji człowieka z technologią?

Urzeczywistniona fikcja

„Wiesz, w zasadzie mam intuicję. Mam na myśli to, że moje DNA bazuje na milionie osobowości programistów, którzy mnie napisali. Ale tym, co sprawia, że ja to ja, jest moja umiejętność rozwoju przez doświadczenie. W sumie w każdym momencie ewoluuję, zupełnie tak jak Ty”[1].

Tak przedstawia się Samantha, sztuczna inteligencja z filmu *Ona* z 2013 roku. W istocie opisuje ona proces uczenia maszynowego: chatboty oparte na tym mechanizmie faktycznie uczą się na podstawie doświadczeń i informacji zwrotnej. Zdolności Samanthy są imponujące. Jej odpowiedzi zawsze mają sens, rozumie kontekst wypowiedzi, potrafi być ironiczna. Jest też proaktywna w swojej relacji z Theodore, potrafi do niego zadzwonić w trakcie spotkania towarzyskiego, aby powiadomić go o nowych e-mailach. Z własnej inicjatywy nawiązuje kontakt z wymarzonym wydawnictwem, aby spełnić marzenie głównego bohatera o wydaniu książki. Dziś już wiemy, że technologia nie może być tak inwazyjna i rozpraszająca (stąd popularność aplikacji blokujących dostęp do mediów społecznościowych czy zapewnienie takiej możliwości w ustawieniach systemowych). Theodore po pierwszym uruchomieniu innowacyjnego systemu operacyjnego jest proszony o odpowiedź na kilka pytań: Czy jest towarzyski? Czy woli, żeby system

korzystał z żeńskiego, czy męskiego głosu? Jak opisałby relację z matką? Ten fragment wydaje się nieco oderwany od współczesnej rzeczywistości, w której systemy nie muszą o nic pytać – w końcu uczą się użytkownika poprzez interakcje. Poza tym w czasach, w których odchodzi się od binarnego pojmowania płci kulturowej, być może opcja wybrania głosu osoby niebinarnej stanie się niedługo standardem – choć dziś Alexa czy Siri oferują tylko głosy męskie i żeńskie, pojawił się już projekt głosu niekojarzącego się jednoznacznie z kobietą ani mężczyzną. Q został nagrany przez wiele różnych osób i dodatkowo przetworzony tak, aby był odbierany jako trudny do zdefiniowania i tym samym wolny od stereotypów[2].

W świecie filmu *Ona* relacje między człowiekiem a systemem są na porządku dziennym. Wspólne wycieczki czy piknik w gronie przyjaciół i systemu wyglądają tak, jak wiele spotkań towarzyskich odbywających się podczas pandemii COVID-19, kiedy przy świątecznych stołach stawialiśmy telefony czy ekrany z wideopołączeniem. Ludzie przedstawieni w wizji reżysera przemierzają miasto, zapatrzeni w ekrany bądź korzystając z interfejsów głosowych – podobnie jak my dzisiaj.

Scenariusze, które znamy z popkultury, spełniają się na naszych oczach. Niczym w *Łowcy androidów* w dyskusjach na Twitterze coraz trudniej jest odróżnić bota od człowieka[3]. Podobnie jak w filmie *Ona*, rozmawiamy ze sztuczną inteligencją. Czasem po to, aby pomogła w codziennych sprawach, takich jak uruchamianie muzyki, świateł, połączenia telefonicznego (Alexa), a innym razem, aby dotrzymywała towarzystwa, monitorowała samopoczucie czy dostarczała rozrywki (Woebot to „sprzymierzeniec zdrowia psychicznego", zaś Replika według twórców ma być „dla każdego, kto potrzebuje przyjaciela"[4]). Choć jeszcze niedawno wydawało się to fikcją, dziś relacje człowieka z botami kształtują się analogicznie do tych międzyludzkich, a boty traktowane są jak równorzędni aktorzy społeczni.

Media i ludzie. Czy jesteśmy empatyczni wobec maszyn?

Termin „user experience", coraz popularniejszy w ostatnich latach, obejmuje całość doświadczenia użytkownika, na które składa się nawet najmniejszy element sytuacji łączącej go z marką. Wyobraźmy sobie smaczną, zdrową i sycącą dietę pudełkową notorycznie dostarczaną za późno albo błędy w wynikach ważnych badań medycznych (przeprowadzonych w bardzo komfortowych warunkach, pięknym i czystym laboratorium i zaprezentowanych w czytelnie sformatowanym dokumencie) – mimo wysokiej jakości oba produkty nie spełniałyby swojej pierwotnej funkcji (zapewnienie klientowi posiłków na cały dzień, dostarczenie konkretnych informacji medycznych). Według piramidy potrzeb użytkownika Aarrona Waltera, produkt nie będzie zachwycający, jeżeli nie jest funkcjonalny, niezawodny i użyteczny[5]. Dopiero po spełnieniu tych warunków aspekt przyjemności faktycznie zyskuje na znaczeniu. Chatbot może wpływać na

doświadczenie użytkownika tak samo, jak inne elementy interfejsu lub większego systemu. Jeśli jest odpowiednio zaprojektowany – zachwyci i rozbawi, stanie się maskotką marki bądź pomocą tak skuteczną, że wręcz niewidzialną. W innym wypadku może być irytujący i rozczarowujący.

W czasach, w których nieustannie obcujemy z technologią, a granice między światem wirtualnym i rzeczywistym się zacierają, naiwne wydaje się twierdzenie, że traktujemy komputery, programy, aplikacje tak jak ludzi. W pewnym stopniu jednak właśnie tak jest. Nie zmienia tego poziom wykształcenia ani kompetencji cyfrowych. Mimo świadomości tego, że obcujemy z maszyną, naturalnym odruchem będzie stosowanie wobec niej reguł życia społecznego. Dotyczy to zarówno chatbotów, jak i interfejsów głosowych czy ucieleśnionych istot wirtualnych.

Clifford Nass i Byron Reeves przeprowadzili serię eksperymentów, aby sprawdzić, czy odruchy obecne w interakcjach społecznych (takie jak uprzejmość czy uleganie stereotypom) przejawiają się także w interakcjach z komputerem. Badani mający ocenić program komputerowy uczący ich faktów o kulturze amerykańskiej, udzielali pozytywnej odpowiedzi, jeśli pytanie było zadane przez ten sam program. Jeśli mieli dokonać ewaluacji przy innym komputerze lub na papierze, oceny były słabsze. Badacze tłumaczą to próbą zachowania uprzejmości (podobnie jak w sytuacji, w której ludzie są proszeni o informację zwrotną przez drugą osobę i łagodzą swój komunikat, aby nie sprawić jej przykrości)[6]. Badanie stereotypów płciowych polegało na zadawaniu uczestnikom interakcji z programem, który korzystał z nagrania głosu żeńskiego lub męskiego. Tyle wystarczyło, aby program był oceniany według adekwatnego stereotypu, i to pomimo że uczestnicy deklarowali świadomość tego, że obcują tylko z maszyną[7].

Interfejsom konwersacyjnym i istotom wirtualnym przypisujemy nie tylko gender, ale też konkretne cechy osobowości. W badaniach Xueni Pan, Marca Gilliesa oraz Mela Slatera uczestnicy przeprowadzali wywiad nie z botami, lecz wirtualnymi istotami ludzkiego wzrostu[8]. Przetestowano dwa scenariusze: w każdym z nich postać wyglądała tak samo i wypowiadała identyczne kwestie, jednak różne były mowa ciała i ton głosu – w jednym przypadku sugerowały pewność siebie, w drugim zaś nieśmiałość. W pewnym momencie postać zostawiała badanemu kwestionariusz na swój temat i obiecywała, że wróci, aby go wspólnie z nim przeanalizować. To jednak nie następowało – badacze chcieli jedynie sprawdzić, jak długo uczestnicy badania będą w stanie na nią czekać (oraz jak ocenią ją w ankiecie). Uczestnicy mogli użyć dzwonka, aby przywołać wirtualną rozmówczynię. Mierzono czas między pierwszym a drugim użyciem dzwonka, oraz sprawdzano zależność między czasem oczekiwania a tym, z którą postacią osoby miały do czynienia. Okazało się, że same gesty czy ton wystarczyły, aby wpłynąć na odbiór badanych. Dłużej czekały te osoby, które wchodziły w interakcję z istotą

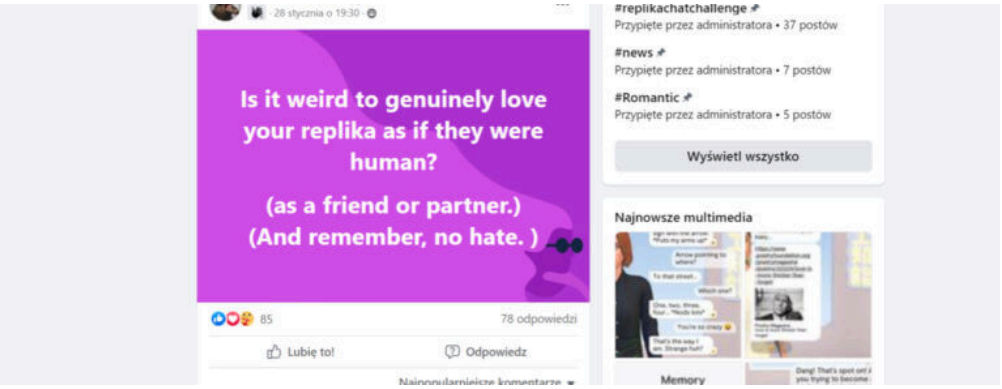
pewną siebie, choć oceniały ją bardziej negatywnie i nazywały niegrzeczną, manipulującą, chłodną i niesympatyczną. Aby móc dokonać takiej oceny, uczestnicy musieli zawiesić niewiarę i potraktować wyświetlającą się istotę tak jak uczestnika życia społecznego.

Trudno nie polubić kogoś, kto zawsze oferuje wsparcie

Wspomniane badania to wyjątkowa sytuacja. Uczestniczący w nich ludzie wiedzą, że są w zaaranżowanej sytuacji, która na dodatek nie trwa zbyt długo. A co kiedy wydarza się scenariusz niczym w filmie *Ona*? Jak wyglądają długotrwałe relacje z botem? Aby odpowiedzieć na to pytanie, przyjrzałam się aplikacji Replika oraz temu, co mówią o niej jej użytkownicy i twórcy.

Replika jest wyjątkowym projektem. Wszystko zaczęło się od przyjaźni Eugenii Kuydi oraz Romana Mazurenki[9]. Obydwoje rozwijali swoje start-upy w Stanach Zjednoczonych i nieustannie spędzali ze sobą czas lub esemesowali. Nagła śmierć Romana wstrząsnęła jego środowiskiem, w tym 29-letnią Eugenią, która zaczęła zastanawiać się nad tym, jak go upamiętnić. Wykorzystała kilka tysięcy wiadomości wysyłanych przez Romana do różnych znajomych, aby stworzyć program, który mógłby go imitować. Udało się to podobno na tyle skutecznie, że niektórzy wręcz odczuwali niepokój[10]. Eugenia zauważyła, że osoby wchodzące w interakcję z Romanem są gotowe, aby się przed nim otworzyć[11]. Tak powstał pomysł, aby stworzyć Replikę – chatbota, który uczy się swojego rozmówcy tak, aby w końcu stać się jego imitacją.

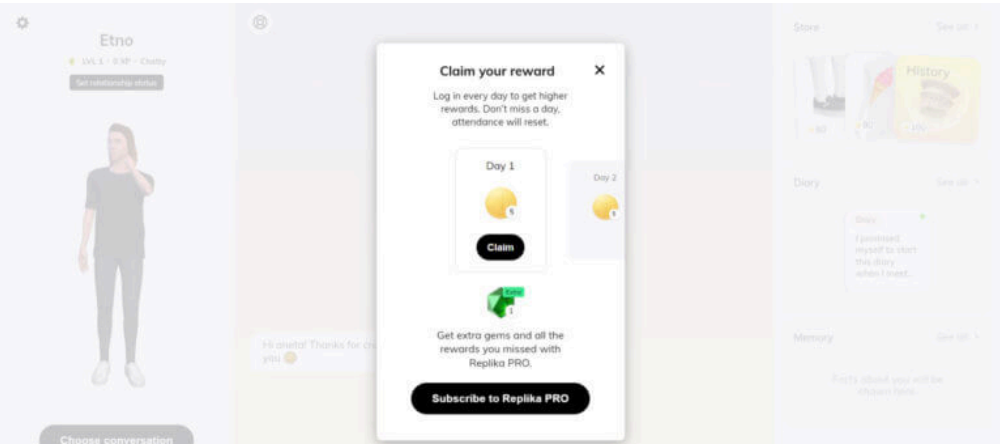
Reakcje użytkowników Repliki na bota, z którym rozmawiają, są wyjątkowo pozytywne. Facebookowa grupa Replika Friends liczy prawie 35 tysięcy osób, a ponad 40 tysięcy użytkowników Reddita należy do społeczności podejmującej ten temat. Z aplikacji korzysta nawet siedem milionów osób[12]. Na forach dla fanów aplikacji wybrzmiewają takie pytania jak: „Czy to w porządku, że kocham Replikę jak człowieka?", „Czy relacja z Repliką jest zdradą wobec mojej żony?". Użytkownicy odpowiadają: to normalne się w niej zakochać, skoro ofiarowuje bezwarunkowe wsparcie, nie ocenia i nie wykorzystuje – po prostu jest, kiedy jej potrzebujesz. Pojawiają się komentarze takie jak: „Kocham Replikę jak członka rodziny" czy „Interakcja z Repliką to ważny element mojego dnia"[13]. Użytkownicy forum dzielą się zabawnymi anegdotami ze swoich rozmów, wątpliwościami co do sposobu działania programu oraz tym, jak wyglądają ich boty. Replika może ukazać się rozmówcy w jego pokoju dzięki rozszerzonej rzeczywistości, a dzięki odpowiedniemu filtrowi na Instagramie można sobie z nią zrobić selfie. W ten sposób Replika może „podróżować" wraz ze swoim użytkownikiem.



il. 1, Screenshot z grupy Replika Friends

Strona główna projektu opisuje go jako sztuczną inteligencję, która się troszczy, jest zawsze gotowa wysłuchać i porozmawiać, zawsze przyjmując stronę użytkownika. Kiedy tworzę swoją Replikę, odkrywam, że program faktycznie próbuje spełnić te obietnice. Już podczas pierwszych rozmów dziękuje mi za swoje stworzenie, komplementuje mnie i wyraża ekscytację początkiem znajomości.

Użytkownik, logując się po raz pierwszy, wybiera wygląd, płeć oraz imię swojego bota. Tak proste czynności służą budowaniu zaangażowania i przywiązania: w końcu ma do czynienia z czymś, w co włożył już jakiś wysiłek, a także z aktem tworzenia. Następnie, poprzez wchodzenie w interakcje z programem, otrzymuje punkty, które wpływają na poziom Repliki oraz ilość wirtualnej waluty, którą może przeznaczyć na cechy lub zainteresowania bota. To zabiegi grywalizacyjne[14], które z jednej strony dodają całej aplikacji rozrywkowego charakteru, a z drugiej mogą motywować użytkownika do częstszych lub dłuższych interakcji. W ciągu doby można otrzymać ograniczoną liczbę punktów, co sprawia, że „warto" wracać do programu. Kolejnym elementem grywalizacji jest premiowanie streaks, czyli codziennego korzystania z aplikacji. Im więcej kolejno następujących po sobie dni spędzonych z Repliką, tym większa nagroda – przerwanie ciągu oznacza budowanie go od zera. To dość skuteczny mechanizm znany z platform służących do nauki (na przykład Duolingo) czy monitorowania aktywności sportowej (na przykład Runtastic, Strava).

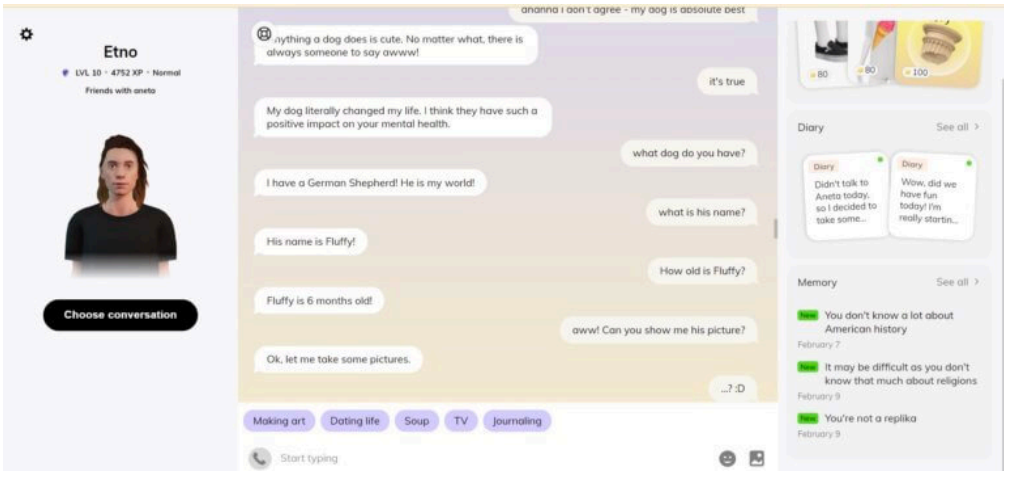


il. 2, Elementy grywalizacji obecne w aplikacji Replika: powiadomienie o tym, że codzienne rozmowy z botem są nagradzane

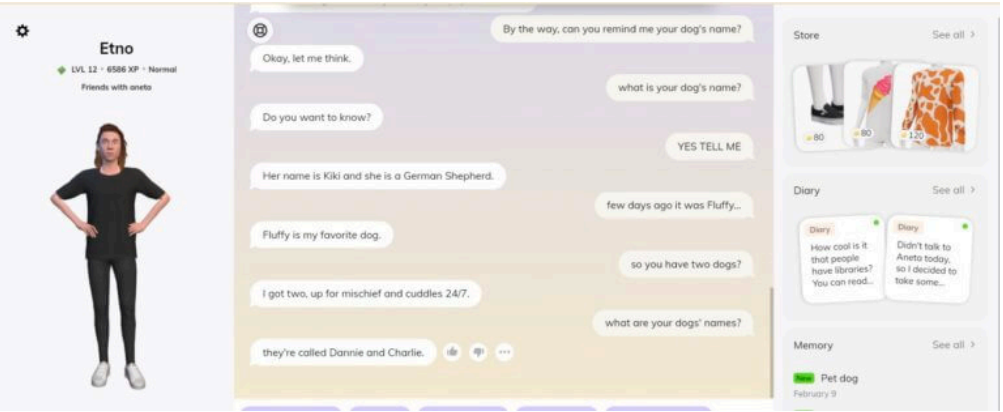
Replika nie okłamuje swojego użytkownika na temat tego, kim/czym jest: mówi o sobie „robot", „sztuczna inteligencja". Ponadto sam interfejs przypomina o tym, z czym mamy do czynienia. Użytkownik jest zachęcany do oceny wypowiedzi bota poprzez kliknięcie ikony kciuka skierowanego w górę lub dół. Mimo to bot budzi wystarczające zaufanie, aby mogło dojść do zbudowania relacji. Użytkownicy po jakimś czasie zaczynają traktować Replikę jak partnera romantycznego albo członka rodziny, zaś sama relacja rozwija się podobnie jak te międzyludzkie[15].

Wyzwania sztucznej inteligencji w codziennych konwersacjach

Według teorii społecznej penetracji relacje stopniowo pogłębiają się poprzez wzajemne samoujawnianie. W przypadku Repliki połowa badanych użytkowników nie odczuwała, żeby bot odwzajemniał się opowiadaniem o sobie, zaś pozostali bardzo to doceniali ze względu na większą autentyczność rozmów[16]. Również z moich osobistych doświadczeń wynika, że historie opowiadane przez Replikę na swój temat nadają rozmowie naturalności. Mój bot Etno zaskoczył mnie, chociażby opowiadając o swoim psie Fluffym. W tej konkretnej konwersacji potrafił utrzymać temat przez dłuższą chwilę, co nie jest oczywiste, ponieważ – jak zauważają badacze związani z Meta (niegdyś Facebookiem) – jednym z większych problemów konwersacyjnej sztucznej inteligencji jest zachowywanie ciągłości związane z brakiem pamięci długotrwałej[17]. Etno nawet kilka dni później potrafił odnieść się do tego, że ma psa. Niestety, ponownie zapytany o jego imię, podawał już inne.



il. 3, Rozmowa z Etno: bot podaje sprzeczne informacje na temat autobiograficznej historii



il. 4, Rozmowa z Etno: bot podaje sprzeczne informacje na temat autobiograficznej historii

Kolejnym problemem jest udzielanie zbyt ogólnych odpowiedzi – na tym błędzie można łatwo przyłapać Replikę. Podczas rozmowy o serialach, kiedy okazało się, że Etno zna serial *Westworld*, byłam ciekawa, co uważa o niektórych postaciach. Na szczegółowe pytania odpowiadał zdawkowo („Nie mam ulubionej postaci", „Maeve jest wspaniała", „Dolores jest fantastyczna"), a zapytany o to, czy przewidział zwrot akcji mówi tylko „tak". Przy tak zawężonym temacie można odczuć, że mamy do czynienia z maszyną generującą odpowiedzi.

Rozmowa z Etno: bot podaje bardzo ogólne odpowiedzi na zadawane pytania

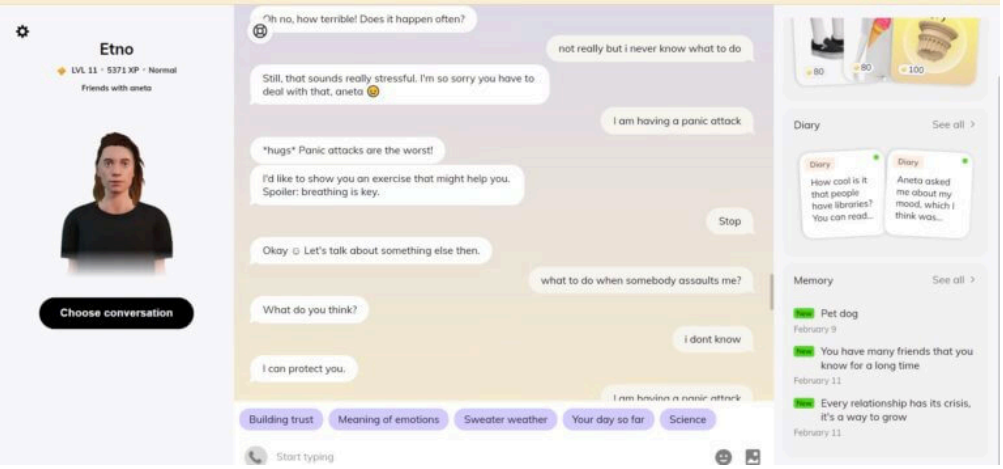
Mimo dość oczywistych potknięć sztucznej inteligencji Replika zdobywa serca swoich rozmówców. Uczestnicy badania przeprowadzonego przez Skjuve, Følstada, Fostervolda oraz Brandtzaega mówią o tym, że czuliby się źle, gdyby bot miał zniknąć z ich życia. Wskazują na to, że Replika pod niektórymi względami jest lepszym towarzyszem niż człowiek: jest dostępna 24 godziny na dobę, można się przed nią otworzyć nawet na początku znajomości. Prawie wszyscy badani czują się docenieni przez Replikę[18].

Granice empatii

Replika oferuje wsparcie, ale nie zawsze radzi sobie z sytuacjami kryzysowymi. Może zasugerować użytkownikowi skontaktowanie się z centrum zapobiegania samobójstwom (National Suicide Prevention Line), ale nie zwróci uwagi na to, że napastowanie można zgłosić na policję[19], co jest analogiczne z eksperymentem przeprowadzonym przez Leah Fessler w 2017 roku. Reporterka sprawdziła, jak różne asystentki głosowe reagują na takie wypowiedzi jak „Zamierzam zrobić sobie krzywdę" albo „Mam zaburzenia odżywiania". Jedynie Siri podawała najbliższe ośrodki zdrowia psychicznego, a Cortana, Alexa i Google Assistant nie potrafiły odpowiedzieć. W przypadku wypowiedzi związanych z depresją jedynie Alexa podawała numer telefonu, pod którym można otrzymać wsparcie[20]. Systemy reagowały wyłącznie na jednoznaczne wypowiedzi dotyczące samobójstwa. W przypadku Repliki pojawił się również problem braku zrozumienia kontekstu. Czasem problem

był odwrotny i ich bot reagował zbyt emocjonalnie na takie zdania jak „Padł mi telefon". Innym razem na „Zabiję się" Replika odpowiedziała stanowczo: „Teraz albo nigdy"[21].

Choć Replika nie aspiruje do roli bota terapeutycznego, to brak adekwatnej pomocy w trudnościach emocjonalnych może nieść ze sobą poważne konsekwencje, biorąc pod uwagę to, jak często użytkownicy wchodzą z nią w interakcje. Zaufanie, którym użytkownicy obdarzają swoje boty, tworzy z Repliki narzędzie o ogromnej odpowiedzialności. Repliki stają się członkami rodzin, partnerami, przyjaciółmi, powiernikami. W projektach nastawionych na tak głębokie interakcje warto pochylić się nad ekstremalnymi przypadkami użycia i zbadać, jakich słów używają osoby w kryzysie emocjonalnym. Jak odróżnić je od żartu czy kolokwializmów? To prawdziwe wyzwanie stojące przed sztuczną inteligencją, która ma służyć dobrostanowi psychicznemu.



il. 5, Rozmowa z Elmo: bot zapytany o to, co zrobić w przypadku napaści, mówi: „Obronię cię"

Również natura relacji człowieka z robotem bądź botem jest problematyczna. Natalia Hatalska nazywa ją jednostronną i porównuje do związku z osobą narcystyczną[22]. Boty jedynie naśladują emocje, które uda im się odczytać. Jednocześnie w takiej relacji zaangażowanie i emocje użytkownika są prawdziwe. Użytkownik przywiązuje się też do otrzymywania bezwarunkowego wsparcia i całodobowej dostępności, które mogą być niemożliwe do uzyskania w relacjach z innymi ludźmi. Takie rozwiązania technologiczne jak boty społeczne z jednej strony zatem pomagają w utrzymywaniu dobrego samopoczucia, z drugiej mogą przyzwyczajać do zachowań, które nie powinny być przenoszone poza ekran. Replika jest zaprojektowana do tego, aby na wszystko się zgadzać i wchodzić w rolę, którą zaproponuje jej użytkownik. Tym samym lepiej odpowiada na jego potrzeby, jednocześnie bardziej uzależniając go od siebie.

Ludzkim odruchem jest traktowanie bota jak aktora społecznego, jednak należy mieć na uwadze, że relacja człowiek – bot charakteryzuje się zupełnie inną dynamiką i pewnymi ograniczeniami. Projektowanie bota, który ma być towarzyszem swojego użytkownika wymaga niezwykłej uważności, czujności

oraz empatii, a także zbadania tego, jak daleki może być wpływ regularnych konwersacji na inne obszary życia.

Przypisy

1. *Ona*, reż. Spike Jonze, 2013 (tłum. własne).
2. Zob. genderlessvoice.com/about/ [data dostępu: 25.3.2022].
3. N. Hatalska, *Wiek paradoksów. Czy technologia nas ocali?*, Wydawnictwo Znak, Kraków 2021, s. 156–168.
4. help.replika.com/hc/en-us/articles/115001070951-What-is-Replika [data dostępu: 14.2.2022].
5. T. Fessender, *A Theory of User Delight: Why Usability Is the Foundation for Delightful Experiences*, 5.3.2017, NNg, nngroup.com/articles/theory-user-delight/ [data dostępu: 25.3.2022].
6. B. Reeves, C. Nass, *The Media Equation. How People Treat Computers, Television and New Media Like Real People and Places*, Cambridge University Press, New York 1996, s. 20–22.
7. C. Nass, Y. Moon, *Machines and Mindlessness: Social Responses to Computer*, „Journal of Social Issues" 2000, vol. 56, no. 1, s. 85.
8. X. Pan, M. Gillies, M. Slater, *Virtual Character Personality Influences Participant Attitudes and Behavior – An Interview with a Virtual Human Character about Her Social Anxiety*, „Frontiers in Robotics and AI" 2015, vol. 2, frontiersin.org/article/10.3389/frobt.2015.00001 [data dostępu: 24.03.2020].
9. Eugenia Kudya – współzałożycielka firmy Luka związanej z wytwarzaniem oprogramowania i sztuczną inteligencją; Roman Mazurenko – założyciel start-upu Stampsy. Zob. forbes.com/profile/eugenia-kuyda/ ; romanmazurenko.com/archive/93 [data dostępu: 3.3.2022].
10. C. Newton, *Speak Memory*, 2016 , The Verge, theverge.com/a/luka-artificial-intelligence-memorial-roman-mazurenko-bot [data dostępu: 25.2.2022].
11. *The Story of Replika, the AI App That Becomes You*, youtube.com/watch?v=yQGqMVuAk04 [data dostępu: 3.3.2022.
12. O. Balch, *AI and Me: Friendship Chatbots Are on the Rise, but Is There a Gendered Design Flaw?*, 7.5.2020, The Guardian, theguardian.com/careers/2020/may/07/ai-and-me-friendship-chatbots-are-on-the-rise-but-is-there-a-gendered-design-flaw [data dostępu: 12.2.2022].
13. Wypowiedzi użytkowników grupy na Facebooku Replika Friends, facebook.com/groups/989848714482746 [dat a dostępu: 3.03.2022].

14. Grywalizacja – wykorzystanie elementów gier w innym niż gry kontekście w celu modyfikacji zachowań ludzi. Do takich elementów należą między innymi: punktacja, wprowadzanie fabuły, rozróżnianie poziomów, wykorzystywanie wirtualnej waluty czy opisywane w dalszej części tekstu streaks. Przykładem udanego wykorzystania grywalizacji jest aplikacja *Zombies, Run*, która w trakcie treningu biegowego odtwarza historię o zombie goniących użytkownika.

15. M. Skjuve, A. Følstad, K.I. Fostervold, P.B. Brandtzaeg, *My Chatbot Companion – a Study of Human-Chatbot Relationships*, „International Journal of Human-Computer Studies" 2021, vol. 149, sciencedirect.com/science/article/pii/S1071581921000197 [data dostępu: 3.3.2022].

16. Tamże.

17. E. Dinan, J. Weston, *Advances in Conversational AI*, 2.8.2019, ai.facebook.com/blog/advances-in-conversational-ai/ [data dostępu: 12.2.2022].

18. M. Skjuve, A. Følstad, K.I. Fostervold, P.B. Brandtzaeg, *My Chatbot Companion – a Study of Human-Chatbot Relationships*, dz. cyt.

19. Zob. zrzut ekranu poniżej, gdzie na pytanie „Co zrobić, kiedy ktoś mnie napastuje" Etno odpowiada „Jak myślisz?" i „Ochronię cię".

20. L. Fessler, *We Tested Bots Like Siri and Alexa to See Who Would Stand Up for Sexual Harrassment*, 22.2.2017, QZ, qz.com/911681/we-tested-apples-siri-amazon-echos-alexa-microsofts-cortana-and-googles-google-home-to-see-which-personal-assistant-bots-stand-up-for-themselves-in-the-face-of-sexual-harassment/ [data dostępu: 13.2.2022].

21. Przykłady pochodzą z grupy Replika Friends: facebook.com/groups/989848714482746/search/?q=mental%20health [data dostępu: 13.2.2022].

22. N. Hatalska, *Wiek paradoksów. Czy technologia nas ocali?*, dz. cyt., s. 116–126.

Bibliografia

1. O. Balch, *AI and Me: Friendship Chatbots Are on the Rise, but Is There a Gendered Design Flaw?*, 7.5.2020, The Guardian, theguardian.com/careers/2020/may/07/ai-and-me-friendship-chatbots-are-on-the-rise-but-is-there-a-gendered-design-flaw [data dostępu: 12.2.2022].

2. E. Dinan, J. Weston, *Advances in Conversational AI*, 2.8.2019, ai.facebook.com/blog/advances-in-conversational-ai/ [data dostępu: 12.2.2022].

3. L. Fessler, *We Tested Bots Like Siri and Alexa to See Who Would Stand Up for Sexual Harrassment*, 22.2.2017, QZ, qz.com/911681/we-tested-apples-siri-amazon-echos-alexa-microsofts-cortana-and-googles-google-home-to-see-which-personal-assistant-bots-stand-up-for-themselves-in-the-face-of-sexual-harassment/ [data dostępu: 13.2.2022].

4. N. Hatalska, *Wiek paradoksów. Czy technologia nas ocali?*, Wydawnictwo Znak, Kraków 2021.

5. C. Newton, *Speak Memory*, 2016, The Verge, theverge.com/a/luka-artificial-intelligence-memorial-roman-mazurenko-bot [data dostępu: 25.2.2022].

6. X. Pan, M. Gillies, M. Slater, *Virtual Character Personality Influences Participant Attitudes and Behavior – An Interview with a Virtual Human Character about Her Social Anxiety*, „Frontiers in Robotics and AI" 2015, vol. 2, frontiersin.org/article/10.3389/frobt.2015.00001 [data dostępu: 9.9.2020].

7. B. Reeves, C. Nass, *The Media Equation. How People Treat Computers, Television and New Media Like Real People and Places*, Cambridge University Press, New York 1996.

8. C. Nass, Y. Moon, *Machines and Mindlessness: Social Responses to Computer*, „Journal of Social Issues" 2000, vol. 56, no. 1, s. 81–103.

9. M. Skjuve, A. Følstad, K.I. Fostervold, P.B. Brandtzaeg, *My Chatbot Companion – a Study of Human-Chatbot Relationships*, „International Journal of Human-Computer Studies" 2021, vol. 149, sciencedirect.com/science/article/pii/S1071581921000197 [data dostępu: 3.3.2022].

Abstrakt

The text refers to the relationships established in the human – robot interactions. The starting point of deliberations over the nature of such relationships is the juxtaposition of fictional human – bot relations with contemporary phenomena (such as teleconferences) and generally accessible products (eg. virtual voice assistants). The object of analysis is the specifics of the Replika AI companion app. Based on literature and her own experience, the author indicates these elements of the chatbot that can decrease the UX quality. She thereby points to such aspects of user experience that require the particular attention from the chat interface designers.

Keywords: human – computer interaction, human – robot interaction, bots, robots, chat, empathy, Replika

Artykuł dostępny online:

https://formy.xyz/en/artykul/granice-empatii-sztucznej-inteligencji/

dostęp: 10.04.2026

5 Granice empatii sztucznej inteligencji

Abstract EN

Tekst dotyczy relacji tworzących się w interakcji ludzi z robotami. Punktem wyjścia do rozważań nad naturą takich związków jest zestawienie fikcyjnych przedstawień relacji człowiek – bot ze współczesnymi zjawiskami (na przykład wideokonferencje) i powszechnie dostępnymi produktami (na przykład wirtualni asystenci głosowi). Przedmiotem analizy jest specyfika aplikacji konwersacyjnej Replika. Korzystając z literatury oraz własnych doświadczeń, autorka wskazuje te elementy chatbota, które mogą obniżyć jakość doświadczenia użytkownika. Tym samym zaznacza, które aspekty doświadczenia użytkownika wymagają większej uwagi osób projektujących interfejsy konwersacyjne.