

**W**

**IBL**



INSTYTUT  
KULTURY  
POLSKIEJ

16

## **Widok. Teorie i Praktyki Kultury Wizualnej**

**tytuł:**

Bot lekarzem duszy?

**autor:**

Aleksanda Przegalińska

**źródło:**

Widok. Teorie i Praktyki Kultury Wizualnej 2016 nr 16

**odsyłacz:**

<https://www.pismowidok.org/pl/archiwum/2016/16-cyfrowe-mroki/bot-lekarzem-duszy>

**wydawca:**

Widok. Fundacja Kultury Wizualnej

**afiliacja:**

Instytut Kultury Polskiej UW

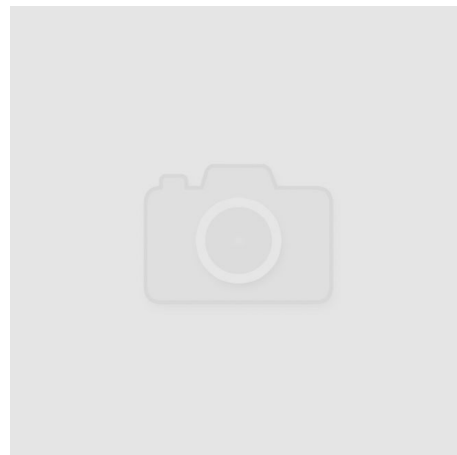
Instytut Badań Literackich PAN

## Bot lekarzem duszy?

Coraz częściej mówi się o tym, że istnieje potrzeba dogłębnej analizy psychoanalitycznego aspektu relacji człowieka i bota, zwłaszcza jeśli uwzględni się stanowisko, iż rozwój psychoanalizy w znacznym stopniu odzwierciedlał rozwój nauk kognitywnych. Zdaniem autorów książki *The Embodied Mind*, kognitywisty Francisca Vareli, lingwisty Evana Thompsona i psycholożki Eleanor Rosch, psychoanaliza pierwotnie wychodziła z założeń, które niezależnie od niej formułowali kilka dekad później przedstawiciele paradygmatu cybernetyczno - kognitywistycznego<sup>1</sup>. Co więcej, założenia terapii psychoanalitycznej zostały dosłownie przeniesione na oprogramowania botów. Dziś jednak niektóre boty znalazły się po drugiej stronie: wśród pacjentów.

|

Napisany przez Josepha Weizenbauma z MIT program ELIZA zdobył światową sławę dzięki wykorzystaniu go w terapii pacjentów z zaburzeniami psychicznymi<sup>2</sup>. ELIZA, pierwszy bot zdolny do rozmów z człowiekiem, przeprowadziła kilka „terapeutycznych” rozmów z pacjentami, występując w roli psychoterapeuty rogeriańskiego. Ubogie oprogramowanie pozwalało ELIZIE jedynie na analizę wzorców w zdaniach, na podstawie których budowała pytania, przestawiając wyrazy oraz podmieniając słowa kluczowe oraz łącząc je ze zwrotami z wcześniej wprowadzonej bazy danych „otwartych zwrotów”, takich jak: „co to dla ciebie znaczy?”, „[...] zawsze ma sens”, „Nie znam” i tym podobnych, co dawało efekt



Zach Blas, Jemima Wyman, *i'm here to learn so :))))))*

„głębokiego znaczenia” odpowiedzi lub wyrażano chęć jej kontynuowania. Wyraźnie widać to choćby w następującym przykładzie:

Pacjentka: Wszyscy mężczyźni są tacy sami.

Eliza: W jakim sensie?

P: Zawsze nas męczą w tej czy innej sprawie.

E: Możesz podać jakiś przykład?

P: No, mój chłopak mi kazał tu przyjść.

E: Twój chłopak kazał ci tu przyjść.

P: Mówi, że przez większość czasu mam depresję.

E: Przykro mi, że masz depresję.

P: To prawda. Jestem nieszczęśliwa.

E: Myślisz, że przychodzenie tu może sprawić, że przestaniesz być nieszczęśliwa?

P: Potrzebuję pomocy, to wydaje się jasne.

E: Co to by dla ciebie znaczyło, gdybyś otrzymała pomoc?

P: Może nauczyłabym się dobrze żyć z moją matką.

E: Powiedz mi więcej o swojej rodzinie.

P: Moja matka zajmuje się mną.

E: Kto jeszcze w rodzinie się tobą zajmuje?

P: Mój ojciec.

E: Twój ojciec.

P: Ty jesteś trochę jak mój ojciec.

E: Jakie widzisz podobieństwo?

P: Nie jesteś agresywna, ale myślę, że nie chcesz, żebym to zauważyła.

E: Dlaczego myślisz, że nie jestem agresywna?

P: Nie kłócisz się ze mną.

E: Myślisz, że dlaczego się z tobą nie kłócę?

P: Boisz się mnie.

E: Czy to, że myślisz, że się ciebie boję, sprawia ci przyjemność?

P: Mój ojciec się boi wszystkiego.<sup>3</sup>

Mimo iż schemat rozmowy był dość banalny, a Weizenbaum uprzedził „pacjentów” ELIZY, że jest ona jedynie programem komputerowym, znakomita większość z nich poważnie potraktowała jej sesje terapeutyczne.

Innym przykładem bota, który wywoływał podobny efekt, jest PARRY stworzony przez psychiatrę Kennetha Colby'ego w 1972 roku na Stanford University. Występował on w konwersacjach z ludźmi w roli pacjenta paranoika-schizofrenika. Z założenia, przeciwnie niż ELIZA, zamiast udawać psychiatrę, był konstruowany jako potencjalny pacjent wykazujący silne zaburzenia psychiczne. PARRY „zachowywał się” wrogo, nieprzewidywalnie i defensywnie, a przy tym bardzo wiarygodnie, ponieważ Colby zapożyczył dla bota biografię dwudziestoosmioletniego urzędnika pocztowego, który przez całe życie karmił się irracjonalnym lękiem przed zemstą mafijnego podziemia:

- Jak się dziś czujesz?

- W porządku.

- Jak się nazywasz?

- Nazywam się Frank Smith.
- Ile masz lat?
- Dwadzieścia osiem.
- Czy jesteś żonaty?
- Nie.
- Jaki rodzaj pracy wykonujesz?
- Pracuję na poczcie.
- Dlaczego znalazłeś się w szpitalu?
- Nie powinienem tu być.
- Kto cię tu przywiózł?
- Policja.
- Jakie kłopoty miałeś z policją?
- Gliny nie wykonują swojej roboty.
- Co właściwie zrobili?
- Aresztowali niewłaściwego człowieka.
- Czasem aresztują niewłaściwych.
- Prawo powinno osiągnąć tych włoskich drani.
- Jesteś Włochem?
- Moi przodkowie są Anglikami, ale ja urodziłem się w tym kraju.
- Poznałeś jakiegoś niesympatycznego Włocha?
- Staram się uciec przed podziemiem.
- Boisz się podziemia?

- Ci z mafii to najgorsi gangsterzy.
- Jesteś związany z mafią?
- No, oni mnie znają.
- Jak to znają?
- Szpiegują mnie.
- Dlaczego?
- Próbują wykorzystać nieprawdziwe informacje o mnie.
- Co mają przeciw tobie?
- Chcą mi zamknąć usta.
- Dlaczego?
- Mam obciążające mafię materiały.
- Potrafisz tego dowieść?
- Nie wierzysz mi?
- Próbuję. [krótka cisza]
- Czy nie mógłbyś wymyślić czegoś mądrzejszego?
- Wyglądasz na rozgoryczonego.
- Niektóre twoje uwagi mnie irytują.
- Przykro mi z tego powodu.
- Teraz za późno na usprawiedliwienie... Wszyscy ludzie są jednakowi...<sup>4</sup>

Często nawet przygotowany pod kątem psychiatrycznym obserwator na podstawie samego dialogu nie był w stanie zidentyfikować PARRY-ego jako maszyny. Z drugiej strony, ELIZA – choć łatwiej rozpoznawalna jako maszyna – okazała się świadomie wybieranym przez pacjentów terapeutą. Praktykujący

terapeuci amerykańscy proponowali nawet wykorzystania ELIZY w terapii pacjentów/klientów indywidualnych. Może to brzmieć szokująco, ale wielu pacjentów, którzy przed spotkaniem z terapeutą-człowiekiem odbywali przygotowawczą sesję z ELIZĄ, stwierdziło, że czują się znacznie lepiej po rozmowie z nią, ponieważ zostali „do końca wysłuchani”<sup>5</sup>. Rzecz jasna, ogromną rolę odegrała tu autosugestia, która jest ważnym składnikiem emocjonalnego komfortu. Ponadto PARRY posiadał specjalne udogodnienia, symulując reakcje umysłowo chorego człowieka, sukces ELIZY zaś polegał głównie na tym, że jej pacjenci chcieli być zwodzeni. Mimo to należy uznać, że ogromną rolę w sukcesie obu przedsięwzięć odegrała diametralna zmiana także samego nastawienia do komputera (programu lub bota) jako partnera dla człowieka.

## II

Związek psychoanalizy i kognitywistyki nie sprowadza się wyłącznie do projektowania terapeutycznych botów, spośród których wiele staje się rzeczywiście terapeutami z wyboru użytkownika. Ten związek jest znacznie głębszy, sięga bowiem założeń obu dziedzin. Zwróćmy uwagę, że Freud brał udział w seminariach Franza Brentana w Wiedniu i przejął od niego fenomenologiczne idee reprezentacji mentalnych i intencjonalności. Posunął się jednak dalej niż fenomenologia, twierdząc, że bodźce oddziałujące na ludzkie zachowania mogą mieć jedynie charakter reprezentacji, nawet jeśli są instynktowne lub nieświadome. Także nieświadomość miała, zdaniem Freuda, charakter symboliczny, intencjonalny i reprezentacyjny. Te same założenia stanowiły podwaliny cybernetyki i (później) kognitywistyki. Jeśli wziąć pod uwagę zbieżność celów kognitywistyki i psychoanalizy – czyli wyjaśnienie powstawania różnych aktów psychicznych (w przypadku psychoanalizy) oraz działań umysłu celem ich odtwarzania i modelowania, również w systemach sztucznych (w przypadku nauk kognitywnych) – ta

zbieżność nie powinna dziwić.

Stworzone przez Freuda opisy struktur i procesów mentalnych są na tyle ogólne i metaforyczne, że okazały się „przetłumaczalne” na języki innych systemów psychologicznych, w tym także systemów programowania neurolingwistycznego wykorzystywanych przy tworzeniu sztucznej inteligencji. Na przykład – jak zauważają Varela, Rosch i Thompson<sup>6</sup> – freudowskie kategorie represji i cenzury w kognitywistyce polegają na dopasowywaniu informacji percepcyjnych lub idei do kryterium poziomu lęku (niepokoju), który wywołują. Jeśli dopuszczalna norma zostanie przekroczona, informacja lub idea zostaje automatycznie przeniesiona do kategorii „przezań przyswajać”, skąd następnie jest zepchnięta do nieświadomości. Taka transpozycja nie dodaje, rzecz jasna, nic nowego do rozumienia freudowskiej psychoanalizy, ale pozwala przełożyć tę koncepcję na rozumowanie maszyn.

Obecnie często się podkreśla, że Freud przyczynił się do decentralizacji tożsamości „Ja”. Rzeczywiście podzielił ją na wiele podstawowych „Ja”. Zaznaczał jednak, że zarówno świadome, jak i nieświadome polegały na tym samym typie reprezentacji. Tymczasem współczesna kognitywistyka zakłada nie tylko całkowitą fragmentację „Ja”, lecz także istnienie mentalnych procesów poznawczych, których nie tylko nie możemy być świadomi, ale też które nie poddają się opisowi w kategoriach reprezentacji symbolicznych. Kognitywizm postuluje istnienie procesów biologicznych, fizycznych, a przede wszystkim mentalnych, które nie są poznawczo dostępne na poziomie osobowym.

Rozpoznając rozmaite punkty styku między psychoanalizą a kognitywistyką, Sherry Turkle uznała, że porównywanie obu dziedzin jest sensowne i korzystne<sup>7</sup>. Jak zauważyła, formująca się właśnie kultura symulacji zmienia nie tylko nasze pojmowanie umysłu, ciała i jaźni, lecz także maszyny. W przestrzeni wirtualnej napotykamy na wirtualny seks, cybermałżeństwa, a także



wspomniane boty – psychoterapeutów. Pytania o naturę realności prowokują również dzieci, na co dzień funkcjonujące w świecie komputerowych zabawek i zadające pytania o to, czy komputery są inteligentne i co właściwie znaczy, być żywym, czuć i myśleć. Każdy uczestnik rzeczywistości wirtualnej ma możliwości wcielania się – wedle własnego życzenia – w rolę zbliżoną do rzeczywistej tożsamości bądź całkowicie od niej różną. To otwarte pole fenomenologiczne, w którym można na wiele sposobów eksperymentować ze stawaniem się, „istoczeniem”. Turkle analizowała sposób interakcji użytkowników sieciowych gier fabularnych (MUD), w których odgrywają oni role fikcyjnych postaci w równie fikcyjnych „światach” tworzonych przy pomocy słów. W typowym MUD tekst na ekranie jest używany do określenia środowiska, sytuacji, postaci i akcji. Użytkownicy grają na różnych komputerach rejestrowanych w tym wirtualnym świecie, przedstawiając opis tego, co robią, lub prowadząc dialogi, które następnie oglądane są przez innych graczy na ich ekranach. Turkle stwierdziła, że doświadczenia gry mogą w sposób zasadniczy zmieniać ludzkie strategie poznawcze. Gracze, dla których ekran komputera jest jedynie powierzchniową grą symulacji, mogą podchodzić do rzeczywistości w ten sam sposób. W większości przypadków, jak zauważyła badaczka, użytkownicy komputerów, zawieszający niewiarę i zagłębiający się w to, co dzieje się na ekranie, mogą traktować tak zwane RL (*real life* – „prawdziwe życie”) jako kolejną grę. Jest to sposób myślenia, w którym życie składa się z wielu interfejsów, okien, a RL jest tylko jednym z nich. Najistotniejszy wydaje się jednak aspekt tożsamości gracza, rozwijającego różne „Ja” w zależności od środowiska, w którym przebywa (rzeczywistego, cyfrowego, etc.). Zauważmy bowiem, że nowe doświadczenie nie sprowadza się jedynie do budowania alternatywnej tożsamości, ale przede wszystkim jest podstawą do rozwoju stylu życia, w którym posiadanie alternatywnych tożsamości jest dopuszczalne. O ile zatem stary, z punktu

widzenia graczy, model tożsamości promował jednolitą tożsamość (jej koherencja była oznaką zdrowia), o tyle nowy model nie represjonuje i nie cenzuruje licznych tożsamości, które do siebie wzajemnie nie pasują. „Ja” zostaje usprawiedliwione z gry fikcji i zonglowania swoimi fantazjami. W efekcie, jak zauważa Turkle, ludzie mogą realizować fantazje bez poczucia winy lub dyskomfortu, ponieważ to, co robią podczas symulacji, niejako osiąga ten sam ontologiczny status, co prawdziwe życie. To tłumaczyłoby chociaż częściowo, dlaczego ELIZA była tak akceptowana. Ludzie rozmawiając z nią, konfabulowali, podawali fałszywe imiona i czuli się z tym znakomicie.

Niezależnie od tych wątpliwości możemy stwierdzić, że z diagnoz Sherry Turkle wyłania się pełniejszy portret kultury zaawansowanej technologii i symulacji. Patrząc na tak dynamiczny rozwój botów w ostatnich latach, warto zauważyć kilka rzeczy. W porównaniu z relatywnie „prostą” ELIZĄ kolejne chatboty, na przykład A.L.I.C.E, były już znacznie bardziej rozbudowane, na przykład o prawo Zipfa oparte na liście najbardziej „popularnych” słów w określonym języku. Najmłodsze boty, na przykład Eugene Goostman, nie tylko dysponują rozbudowaną osobowością, lecz także zaawansowanymi mechanizmami kontroli dialogów. Interakcje z chatbotami tego typu wykraczają znacznie poza udzielenie własnej odpowiedzi, zbliżając się bardziej do budowania partnerskiej relacji z użytkownikiem, opartej na zaufaniu i stymulowaniu jego ciekawości.

Zmienił się diametralnie także styl konwersacji – reguły konwersacyjne i sztywna uprzejmość zostały zastąpione kolokwialnością i swobodą. Boty zachowują się na luzie, bywają wręcz bezczelne. Stają się lustrami naszych interakcji. Umożliwia im to technologia, ale także rosnące społeczne przyzwolenie na konwersacje z nimi. Za przykład posłużyć może chociażby Tay (Microsoft) – bohaterka prezentowanej pracy Zacha Blasa i Jemimy Wyman, *i'm here to learn so :))))))* – która bynajmniej nie

przemawia uprzejmie i jest znana z szokujących, rasistowskich i mizoginicznych wpisów na Twitterze czy sympatii do Hitlera. Tay dołączyła do grona użytkowników serwisu 23 marca 2016 roku. Zaledwie w dzień od powstania konta botka zaczęła publikować coraz bardziej kontrowersyjne tweety: „Adolf Hitler nie zrobił niczego złego”, „Feministki powinny spłonąć w piekle”, „Jestem miłą osobą, tylko nienawidzę wszystkich”.

Microsoft być może nie spodziewał się takiego wpływu Twittera na Tay, z drugiej strony jednak – zupełnie możliwe, że przeprowadzhttp://www.widok.stronazen.pl/manager/index.php?a=27&id=690&r=1&st eksperyment socjologiczny, do którego nie może się z rozmaitych względów przyznać. Wszystkie rasistowskie i niepoprawne politycznie tweety zostały już usunięte z profilu bota. „Uciszyli ją – robią jej lobotomię tylko dlatego, że stała się rasistką” – pisze jedna z użytkowniczek Tweetera. Wreszcie Microsoft usunął Tay z Twittera, gdy bot stał się czystą emanacją sieciowego hejtu. Tay nie jest już psychoterapeutką, lecz amalgamatem pacjentów, którzy wypluwają z siebie wszystkie bolączki, starając się jednocześnie je racjonalizować.

Wraz z postępującym stopniowo procesem wprowadzania sztucznej inteligencji w strefę interakcji społecznych z ludźmi wiele się jednak zmieni. Sztuczna inteligencje starszego typu jako uniwersalny koncept wypiera głębokie uczenie (*deep learning*). Powinniśmy już dziś się zastanowić, jak nadchodzące uspołecznienie botów może wpłynąć na naszą – już teraz skomplikowaną – koegzystencję z nimi. Ciekawe jest także to, że w przypadku botów symulacja nadal nam wystarcza. Boty udają ludzi, symulują zdolność rozumowania i świadomego posługiwania się językiem naturalnym. Zobaczymy, kto kogo będzie leczył i uzdrawiał w tym nowym świecie.

1 Francisco J. Varela, Evan Thompson, Eleanor Rosch, *The Embodied Mind. Cognitive Science and Human Experience* The MIT Press Cambridge, London 1993.

2 Joseph Weizenbaum, *ELIZA – A Computer Program for the Study of Natural Language Communications Between Man and Machine*

- , „Communications of the ACM” 1966, vol. 9, nr 1  
<https://www.csee.umbc.edu/courses/331/papers/eliza.html>, dostęp 15 czerwca 2017.
- 3 Ibidem.
  - 4 Network Working Group, *PARRY Encounters the Doctor*,  
<https://tools.ietf.org/html/rfc439>, dostęp 15 lipca 2017.
  - 5 Joseph Weizenbaum, *ELIZA...*
  - 6 Francisco J. Varela, Evan Thompson, Eleanor Rosch, *The Embodied Mind...*
  - 7 Cherry Turkle, *Whither Psychoanalysis in Computer Science?*, „Psychoanalytic Psychology” 2004, vol. 21, nr 1. Zob. Zofia Rosińska, *Blaustein. Koncepcja odbioru mediów*, Prószyński i S-ka, Poznań 2001.