

# Joanna Łapińska

---

## Świadectwo afektu : spotkania ludzi i sztucznej inteligencji w filmach "Ex machina" i "Uncanny"

---

Kultura Popularna nr 2 (48), 106-115

---

2016

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej [bazhum.muzhp.pl](http://bazhum.muzhp.pl), gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Joanna Łapińska

# Świa- dectwo afektu.

## *Spotkania ludzi i sztucznej intelligen- cji w filmach Ex ma- china i Uncanny*

*„Kiedyś sztuczna inteligencja będzie traktować nas tak, jak my traktujemy skamieniałości znalezione w Afryce. Wyprostowane małpy żyjące w kurzu, porozumiewające się prymitywnymi językami i narzędziami. Musimy wyginąć.”*

Nathan (*Ex Machina*, 2015)

## Wstęp

W niniejszym artykule przeanalizuję przedstawienia sztucznej inteligencji (ang. *Artificial Intelligence; AI*) znane z dwóch współczesnych filmów amerykańskich: *Ex Machina* (2015) Alexa Garlanda i *Uncanny* (2015) Matthew Leutwylera. Interesują mnie wyobrażenia autorów wymienionych filmów na temat sztucznej inteligencji, jej roli oraz (romantycznych) związków z człowiekiem, a także umieszczenie ich w szerszym kontekście badań posthumanistycznych oraz badań nad afektami. Moim celem jest zbadanie wyżej wymienionych przedstawień w kontekście dawania przez bohaterów opowiadanych historii świadectwa na temat istnienia afektu u robotów humanoidalnych obdarzonych sztuczną inteligencją.

Filmowe wyobrażenia sztucznej inteligencji umiejscawiam w dwóch głównych kontekstach. Przede wszystkim dostrzegam w nich przyczynek do badań posthumanistycznych, stawiających w centrum swojego zainteresowania już nie tylko człowieka, lecz także pozostałe elementy uniwersum. Posługuję się tutaj kategoriami wprowadzonym przez Gilles'a Deleuze'a i Felixa Guattariego, a także ich rozszerzeniem i interpretacjami zaproponowanymi, między innymi, przez Rosi Braidotti i Agnieszkę Jelewską. Drugi niezmiernie ważny kontekst, zasygnalizowany już w tytule artykułu, to aspekt afektywny. Tzw. zwrot afektywny w badaniach humanistycznych wydaje się dostarczać interesujących narzędzi do badania tekstów filmowych traktujących o związkach człowieka z maszyną. Za kluczową w moich badaniach i szczególnie mi bliską uznaję teorię afektów Silvana Tomkinsa, w duchu której postaram się interpretować przedstawienia sztucznej inteligencji na filmowym ekranie.

W mojej analizie wyobrażeń sztucznej inteligencji na ekranie posługuję się także – *last but not least* – kategorią „bycia świadkiem” (ang. *witnessing*), której używam w kontekście testu Turinga, któremu poddawane są roboty w filmach. Kategoria ta wydaje mi się szczególnie przydatna w analizie wybranych przeze mnie dzieł, gdyż bohaterowie obrazów są bardzo wyraźnie stawiani w szczególnej sytuacji „bycia świadkiem” i „dawania świadectwa” na temat sztucznej inteligencji. Uważam, że bohaterowie filmów, wchodząc w interakcje z humanoidalnymi robotami, są świadkami afektu, który rodzi się w obserwowanych robotach na oczach ludzi i który wydaje się konieczną i integralną częścią sztucznej inteligencji.

Zauważamy także, że inteligencja człowieka może być traktowana jako coś w rodzaju maszyny, lecz z elementem dodatkowym, wynikowym, nadanym – czyli właśnie, jak sądzę, z afektem. Autorzy filmów pytają o to, czy sztuczna inteligencja może także być stworzona na podobieństwo ludzkiej oraz jakie niesie to za sobą konsekwencje dla człowieka. Wszystko to w duchu posthumanistycznym – człowiek nie jest już w centrum wszechświata, lecz znajduje się pośród innych elementów uniwersum, z którymi musi współpracować i z którymi łączy go przeróżne związki.

## Od automatu do komputera

Zacznijmy od bardzo krótkiego zarysu historii przedstawień sztucznego człowieka w kulturze.

Sztuczny człowiek nie zawsze musiał posiadać, według jego twórców, coś na kształt inteligencji czy świadomości, a przynajmniej nie były te przymioty nazywane w ten sposób. Wieki XVIII i XIX to czas triumfu tzw. „automatów” – urządzeń niemających świadomości, lecz służących ludziom, np.

Joanna Łapińska

– magister Filologii  
Polskiej w zakresie  
Filmoznawstwa, Telewizji  
i Kultury Medialnej.  
Doktorantka na Uniwer-  
sytecie Humanistyczno-  
społecznym SWPS  
w Warszawie. Przygotowuje rozprawę doktorską na temat związków miłosnych między człowiekiem a robotem we współczesnym kinie.  
E-mail: jlapinska1  
@st.swps.edu.pl

w gospodarstwie domowym czy w celu rozrywkowym. To właśnie wtedy powstały setki „samoporuszających się androidów” (Radkowska-Walkowicz, 2008) – figurek tworzonych przez naukowców i rzemieślników ku uciesze władców i gawiedzi. Warto wspomnieć tu o bodaj najślynniejszym, żyjącym w XVIII wieku twórcy automatów – Jacques’u de Vaucansonie, którego dziełami: bębniarzem, flecistą, rysownikiem oraz chyba najbardziej znaną kaczką wyposażoną w mechaniczny układ pokarmowy pozwalający na jedzenie i wydalanie, zachwycała się swego czasu cała Europa.

Koncepcja sztucznego człowieka jest oczywiście równie stara jak cywilizacja. Już w *Iliadzie* i *Odysei* Homer pisał o automatycznych niby-robotach służących Hefajstosowi w pracowni, a także o humanoidalnych urządzeniach-kobietach, które wyposażone były nawet w inteligencję i umiejętność mówienia, czyli w swego rodzaju świadomość (Mendelsohn, 2015). Przez wieki uformowały się dwie ścieżki, którymi podążały późniejsze narracje dotyczące związków ludzi i robotów: pierwsza, której bohaterami są przedmioty ożywione służące ludziom, oraz druga – gdzie maszyny zdobywają inteligencję oraz świadomość i stają się równe ludziom, a w końcu zmieniają się w ich wrogów.

Era komputeryzacji dała przyczynę do tworzenia się nowej literatury *science fiction* – z nowymi narracjami, podyktowanymi pojawieniem się komputera, któremu najbliżej chyba do czystej inteligencji, a więc też do człowieka. Komputery stały się z pewnością towarzyszami człowieka w codziennym życiu, co pociągnęło za sobą wzrost liczby filmów mierzących się z tematyką związków człowieka z maszyną w tych nowych okolicznościach. Do takich filmów zaliczymy obrazy, którym w dalszym ciągu przyjrzymy się nieco uważniej, tj. *Ex Machina* (2015) Alexa Garlanda oraz *Uncanny* (2015) Matthew Leutwylera.

Filmy te cierpią pełnymi garściami z wcześniejszych przedstawień istot sztucznych i ich związków z ludźmi na szklanym ekranie, np. z filmów 2001: *Odyseja kosmiczna* (1968) Stanleya Kubricka, *A.I. Sztuczna inteligencja* (2001) Stevena Spielberga i innych. Jednak filmy, które mnie zainteresowały, znacząco różnią się od wymienionych obrazów, a także wyrażają to, czego tamte nie starały się ukazać – czyli związku miłosnego i uwikłania romansowego między robotem/sztucznym wytworem a człowiekiem. W *Lowcy androidów* (1982) Ridleya Scotta problem ten jest także istotny, lecz wydaje się, że proza Philipa K. Dicka oraz film na jej podstawie kładą większy nacisk na obawę dotyczącą niemożności rozróżnienia między człowiekiem a androidem – i to jest główny temat utworów. To w *Lowcy androidów* właśnie problem zostaje też po raz pierwszy przedstawiony w duchu posthumanistycznym – człowiek nie jest jedyny w swoim rodzaju, nie jest panem i centrum wszechświata. Jest tylko jednym z elementów i być może wcale nie najważniejszym.

## W duchu posthumanizmu

Gilles Deleuze i Felix Guattari to nazwiska, które badaczom identyfikującym się z badaniami posthumanistycznymi z pewnością nie są obce. Koncepcje obu filozofów odbiły się szerokim echem na całym świecie, a ich szerokie zastosowanie w badaniach nie tylko humanistycznych jest już faktem. Co jest w nich tak rewolucyjnego?

Deleuze i Guattari znoszą tradycyjny podział na świat ludzki i przyrodniczy. Człowiek nie jest już w centrum wszechświata – przestaje być miarą wszechrzeczy, człowiekiem witruijańskim z ryciny Leonarda da Vinci. Nowy podmiot stoi teraz na równi z elementami wszechświata-universum (czy, jak

chce Agnieszka Jelewska, sensorium), a jego obecność w nim jest definiowana przez zdolność tworzenia powiązań w złożonej, nieskończonej sieci zmysłów i połączeń (Jelewska, 2012). To, co nie-ludzkie (roślinne, zwierzęce, mineralne, ale też sztuczne, cyborgiczne, mechaniczne...) zajmuje pozycję tuż obok człowieka, a życie – rozumiane już nie jako *stricte* ludzki *bios*, lecz jako nie-ludzkie, nieosobowe *zoe* – jest przedmiotem refleksji posthumanistycznej. Takie podejście do rzeczywistości daje nam możliwość myślenia o podmiotach jako o „posiadających przepuszczalne granice, stanowiących złożenie nieosobowych i heterogenicznych sił, czasem współdziałających, a czasem ścierających się w burzliwy sposób” (Bednarek, 2014: 29).

Dlaczego kontekst posthumanistyczny jest dla nas tak istotny? Głównymi bohaterami filmów *Ex Machina* i *Uncanny* są postacie o statusie nie-ludzkim. W *Ex Machina* jest to Ava – żeńska wersja androida obdarzonego przez swojego twórcę niebywale błyskotliwą sztuczną inteligencją i pięknym, ponętym ciałem. Status ontologiczny Avy jest niejasny – jest ona jak gdyby zawieszona pomiędzy światem ludzi a przedmiotów. Z jednej strony nie można odmówić jej inteligencji (również emocjonalnej), która predestynowałaby ją do statusu odrębnej, wolnej, niezawisłej jednostki mogącej o sobie decydować. Z drugiej strony jednak wciąż jest maszyną całkowicie zależną od swego stwórcy – robotem służącym do zabawy, do sprawiania mu przyjemności. Jak dowiemy się z historii opowiedzianej w filmie, Ava zawalczy o swoją godność i prawo do samostanowienia. Zniszczy swojego twórcę, wykorzysta człowieka, który się w niej zakochał i umożliwił jej ucieczkę – wszystko po to, aby udać się na „ruchliwe skrzyżowanie ulic” (tam właśnie chciałaby pójść Ava, gdyby mogła opuścić swój pokój – do czego przyznaje się Calebowi podczas jednej z ich rozmów. Twierdzi, że skrzyżowanie stanowi ciągle zmieniający się obraz ludzkiego życia).

Podobnie jest w filmie *Uncanny* – tu głównym bohaterem jest android już zupełnie nieodróżnialny od człowieka (historia w filmie jest opowiedziana w ten sposób, że ten, którego bierzemy za człowieka, okazuje się być maszyną i odwrotnie). Pytanie, o to, czy przedmioty nieożywione powinny mieć status podobny do ludzi, czy powinny być bohaterami historii, czy mają prawo (!) być traktowane jak odrębne jednostki, jest stawiane w obu filmach. Co więcej, wydaje mi się, że maszyny jako odrębne podmioty są w opowiedzianych historiach o wiele bardziej interesujące niż ludzie, którzy są nudni, przewidywalni i naiwni.

Oba filmy stawiają w centrum zainteresowania podmioty posiadające sztuczną inteligencję i ich relację z człowiekiem, przy czym wyjątkowy status tego ostatniego jest tu dyskusyjny. Przeczytajmy choćby hasła promocyjne z plakatów filmowych obu obrazów:

It's only human. (*Uncanny*)

There is nothing more human than the will to survive.

(*Ex Machina*)

Związłość i dosadność pierwszego z haseł podkreślają zwątpienie w wyjątkowość gatunku człowieka. Szczególnie interesująco zestawione zostaje ono z postaciami dwóch mężczyzn na plakacie, którzy na pierwszy rzut oka mogą wydawać się jedną osobą o dwóch twarzach – albo zrośniętymi głowami braćmi syjamskimi. Czy sztuczna inteligencja to druga (mroczna?) strona człowieka? A może to ten ostatni jest „tylko człowiekiem”, tj. przedmiotem zależnym od innej jednostki, która może nim obracać niczym głową?

Drugie z haseł odnosić się może do woli przetrwania, którą wykazywała

Ava. Pragnęła ona bowiem za wszelką cenę uwolnić się od swojego bezdusznego kreatora (albo „ojca”), zyskać wolność, decydować o swoim losie, przerwać niewolę, w której tkwiła. To pragnienie – obce tradycyjnym maszynom, lecz dobrze znane ludziom – ma podkreślać bardziej „ludzki” niż „maszynowy” ontologiczny status Avey. Oczywiście jest sprawą dyskusyjną, czy Avey zależało na tym, aby nazwać jej pragnienia „ludzkimi”. Twórcy filmowi nie odpowiadają na to pytanie wprost, lecz opowieść sugeruje, że Avey było to obojętne – gdyż człowiek nie był dla niej ani miarą wszechrzeczy ani wzorem do naśladowania. Wydaje się, że maszyna pragnie sama wypracować sobie swój status. Ludzie jednak potrafią ją ocenić tylko z własnej perspektywy. Dlatego Caleb ufa Avey i wydaje mu się, że ona także zapalała do niego uczuciem. Nasz bohater nie widzi, że jest tylko pionkiem w jej grze, gdyż ocenia robotę z własnej, ludzkiej perspektywy; projektuje swoje miłosne uczucia na Avey, wyobrażając sobie, że ona także darzy go uczuciem.

Humanistyka nie-antropocentryczna, jak rozumie ją Ewa Domańska, dopuszcza myśl o nietypowym, nieantropocentrycznym traktowaniu innych figuracji podmiotowości, w tym przedmiotów nieożywionych, zwierząt, cyborgów, maszyn (Domańska, 2008). Tradycyjna humanistyka już z definicji miała stawiać w centrum swojego zainteresowania człowieka, dlatego tak trudno obecnie jej przedstawicielom zrewidować wielowiekowy dorobek. Deantropocentryzacja humanistyki jest zatem sporych rozmiarów wyzwaniem.

Wraz ze zmieniającym się przedmiotem zainteresowań narasta potrzeba stworzenia nowych metod, teorii, sieci pojęciowych, które wyjdą poza semiotykę czy konstruktywizm. Domańska widzi możliwość skorzystania z innych dziedzin nauki, np. powstałej w latach 80. technonauki (ang. *technoscience*) czy z doświadczeń sztuki. Może tu być przydatne myślenie w kategoriach intencjonalności przedmiotów nieożywionych – np. nadanie statusu aktywnych podmiotów sprawczych rzeczom, siłom natury, i innym. Nie chodzi tu oczywiście o wiarę w animalizm, lecz po prostu o podkreślanie złożonych związków między przedmiotami nieożywionymi a podmiotami ludzkimi, o ich wzajemne oddziaływanie na siebie, ich współistnienie – bez wskazywania na nadrzędność jednego z nich – w skomplikowanej sieci połączeń.

Drugim sposobem, według Domańskiej, jest „wyrwanie” przedmiotów z ich otoczenia, odarcie ich z funkcjonalności i przyjrzenie się ich nagiemu, „biednemu stanowi” (zgodnie ze słowami Tadeusza Kantora). Fascynacja przedmiotem niedostępnym, odległym, nieobciążonym znaczeniem symbolicznym to też jedna z dróg rozwoju współczesnej humanistyki nie-antropocentrycznej, zajmującej się materią nieożywioną. Należy zaakceptować fakt, że przedmioty często nie są zdyscyplinowane i niekoniecznie wyrażają to, co widzi w nich człowiek. Personifikacja rzeczy nie powinna być jedyną drogą badania ich statusu i znaczenia.

Podkreślę raz jeszcze: wydaje mi się, że główni bohaterowie analizowanych filmów świetnie pasują do definicji „innych figuracji podmiotowości”, o których mówi Domańska. Zarówno Ava, jak i android z *Uncanny* są niczym przedmioty wyrwane z ich otoczenia, są zarazem ludźmi, lecz także maszynami. Ich funkcjonalność jest niejasna – czy mają służyć ludziom? A może powinny same stanowić o sobie? I jaki to ma związek z afektem? Zobaczmy.

## Dać świadectwo afektowi

Słowo, które robi obecnie zawrotną karierę w humanistyce, czyli afekt, to drugą ciekawą kategorią, które pomoże nam zinterpretować wybrane filmy

o związkach maszyn i ludzi. Antonio R. Damasio stwierdził w słynnej książce *Błąd Kartezjusza*, że ciało i emocje współtworzą świadomość (Damasio, 2011). Położył tym samym kres naiwnym przekonaniom o rozdzielności ciała i umysłu oraz prymatu tego ostatniego nad pierwszym. Dzisiaj już wiemy, że – parafrazując tytuł książki badacza – Kartezjusz się pomylił, oddzielając grubą kreską światły i nadrzędny umysł od grzesznego, marnego, skazanego na zagładę ciała. Nasza świadomość jest tworzona również przez emocje (a także przez ciało, odczuwające je w sposób fizyczny, namacalny), które przez wiele lat spychane były na niższą pozycję w hierarchii ważności, daleko za umiejętnościami analizowania, kalkulowania i racjonalnego podejmowania decyzji „z zimną krwią”. Umiejętność rozdzielenia uczuć od rozumu stała się najwyższym celem człowieka mającego się przeciw za istotę myślącą.

Termin „afekt” i to, jak należy go rozumieć, nie jest wciąż jednoznaczne. Jedną z najważniejszych teorii w tym zakresie jest ta Silvana Tomkinsa – amerykańskiego psychologa i filozofa. Według niego człowiek rodzi się z zestawem dziewięciu afektów, do których należą: 1) zainteresowanie-ekscytacja (ang. *interest-excitement*), 2) radość (ang. *enjoyment-joy*), 3) zaskoczenie-zdziwienie (ang. *surprise-startle*), 4) wzburzenie (ang. *distress-anguish*), 5) gniew (ang. *anger*), 6) strach (ang. *fear*), 7) wstyd (ang. *shame-humiliation*), 8) obrzydzenie (ang. *disgust*) oraz 9) reakcja na przykre zapachy (ang. *dissmell*). Afekty, według Tomkinsa, mają charakter zarazem biologiczny, jak i kulturowy. Są one głównym motywem działania organizmów, a dzięki tzw. sprzężeniu zwrotnemu (ang. *feedback loop*) poznajemy świat, który nas otacza, reagujemy na różne sytuacje, które nas spotykają i nadajemy im osobiste znaczenia. Sprzężenie zwrotne działa na zasadzie swoistej wstęgi: bodziec (z zewnątrz) – afekt (w naszym ciele/umyśle) – reakcja (na bodziec). Nasze osobowość i świadomość są właśnie wynikiem tego procesu zachodzącego w ciągu naszego życia nieskończoną ilość razy (Kosofsky-Sedgwick i Frank, 2014).

Od uznania wagi emocji w codziennym życiu już tylko krok do zwrotu afektywnego w humanistyce. Według Zofii Budrewicz i innych zainteresowanych tym tematem badaczy, kategoria afektów może być niezwykle przydatnym narzędziem do badania tekstów kultury. W nowej humanistyce liczy się bowiem to, jak dzieło sztuki działa na odbiorcę; najistotniejsze jest to, co „niewyjaśnialne, nieujmowalne językowo, niewypowiadalne” (Sendyka, 2014: 15). Afekty, uczucia, emocje stają się kategoriami, za pomocą których można badać dzieła literackie i filmowe. Jak zauważa Roma Sendyka, jest to obszar niepoznany, dziewiczy, dziki i „nieomapowany”; prawdziwa terra incognita, po której poruszać się możemy prawie że po omacku, korzystając tylko z nielicznych drogowskazów (Sendyka, 2014: 17).

Nowe podejście do badań humanistycznych wynikające między innymi z porzucenia narracyjnych form i tradycyjnych, prostych modeli dychotomicznych, pozwala, a nawet nakazuje sięgnięcie do nietradycyjnych metod badawczych. Prostota i linearność powoli zastąpione zostają „siecią” (znaną z koncepcji Bruno Latoura), w której sprawczością charakteryzują się również aktorzy „nie-ludzcy”, np. komputery i maszyny, czy „polem” (według Pierre’a Bourdieu). Historia nie jest już wielką narracją, a najważniejsze staje się dążenie do tego, aby poczuć się jak w danym czasie historycznym. Nieprzypadkowo stawia się tutaj na odczucia – to właśnie one grają pierwsze skrzypce w afektywnym zwrocie humanistycznym.

Afekty regulują stosunki między człowiekiem a światem zewnętrznym. Mogą być przydatne do zbadania czegoś, co zalega głęboko, czego nie sposób

wyrazić słowami jasno i wyraźnie. Przekaz słowny i komunikacja za pomocą znaków nie spełniają w takich wypadkach swojego zadania. Kategoria afektu może być przydatna w badaniu problematyki, gdzie tradycyjne ujęcia się nie sprawdzają – np. w studiach nad wojenną traumą, lecz nie tylko.

Uważam, że afekt w interesujących nas filmach ściśle łączy się z kategorią „bycia świadkiem”. Afekt w *Ex Machina* i *Uncanny* przestaje być cechą przynależną tylko podmiotowi ludzkiemu. To figuracje nie-ludzkie (maszyny obdarzone sztuczną inteligencją) okazują się podmiotami wyposażonymi w emocje i uczucia (a przynajmniej świetnie je symulują – czy to jednak sprawia, że są one mniej prawdziwe, mniej odczuwalne, mniej namacalne? Jest to jedno z trudniejszych pytań w badaniach nad sztuczną inteligencją w ogóle. David L. Levy w książce *Love & Sex with Robots* przekonuje nas, że symulacja stanów emocjonalnych przez maszynę może mieć taki sam wpływ na człowieka, jak obcowanie z żywym organizmem, którego emocje nie są „udawane”, i że w ogólnym rozrachunku fakt, że emocje są symulowane, nie ma większego znaczenia dla człowieka obcującego z robotem). Być może są to afekty swoiście „robotyczne” – przynależne tylko wytworom sztucznym i nie wydaje się słuszne przykładanie do nich ludzkiej miary (zwłaszcza w kontekście humanistyki postludzkiej). Jednak faktem jest, że ludzie bohaterowie filmów stający oko w oko ze sztuczną inteligencją (Caleb w *Ex Machina* oraz dziennikarka Joy Andrews w *Uncanny*) są świadkami istnienia świadomości czy też afektu w obserwowanych robotach. Bohaterowie mają zadanie do wykonania – Caleb musi przetestować Avę przy pomocy testu Turinga i zdać relację z postępów swojemu przełożonemu, Nathanowi, zaś dziennikarka Joy dostaje zlecenie napisania artykułu na temat sztucznej inteligencji do specjalistycznego czasopisma, co również jest formą dawania świadectwa.

Afekt ma związek z aspektem romantycznym/miłosnym w analizowanych przez nas filmach. Miłość jest tu niczym ostateczna odpowiedź w teście Turinga – odpowiedź twierdząca (wyobraźmy sobie ją następująco: „Tak, inteligentna maszyna wzbudza we mnie uczucia podobne do tych, które wzbudzać mogą inni ludzie. Tak, zakochałem się w maszynie i maszyna mnie kocha. Uznaję więc maszynę za jednostkę potrafiącą rozbudzać uczucia i uczucia posiadającą”). Do testu Turinga jeszcze wrócimy.

## AI? Nie pytaj „czy”, lecz „kiedy”!

Sama historia opowiedziana w filmie *Ex Machina* jest dość prosta. Otóż Nathan (grany przez Oscara Isaaca), wizjoner i genialny twórca informatycznego imperium, zaprasza do tajnej posiadłości jednego ze swoich pracowników – informatyka Caleba (granego przez Domhnalla Gleesona), gdzie czeka na niego żeński android (postać brawurowo odtwarzana przez Alicję Vikander), którego ten ma poddać testowi Turinga.

Podobną sytuację mamy w *Uncanny*. Tu dziennikarka Joy Andrews (grana przez młodą aktorkę, Lucy Griffiths) dostaje zlecenie napisania artykułu na temat badań nad sztuczną inteligencją prowadzonych przez wybitnego naukowca, Davida Kressena (w tej roli Mark Webber). Joy każdego dnia przez tydzień odwiedza Davida w jego laboratorium i prowadzi rozmowy z nim oraz z androidem, Adamem (David Clayton Rogers). Artykuł Joy ma mieć charakter popularyzujący osiągnięcia badacza w dziedzinie sztucznej inteligencji i ma być świadectwem tego, czego Joy doświadczyła na własnej skórze. Jej zawód – dziennikarza-reportera – już od razu narzuca nam pewną optykę spojrzenia.



Reporter bowiem z definicji ma za zadanie przekazywać obiektywną prawdę o opisywanych zjawiskach, dawać jej świadectwo. Obiektywizm reporterski przegra tu z uczuciem, gdy Joy zakocha się w genialnym robotyku i da tym samym swoje subiektywne świadectwo afektywnej relacji łączącej ją z robotem.

W *Ex Machina* Caleb zaczyna więc testować Avę. Bohaterowie oddają się długim rozmowom, które są rejestrowane przez kamery w celu późniejszej analizy przez Nathana. Caleb jest wyraźnie zafascynowany inteligencją Avy, jej zadziwiającymi umiejętnościami prowadzenia ujmującego dialogu, a także swoistą wrażliwością i ciekawością świata. W miarę upływu czasu Ava zmienia się pod wpływem rozmów – wydaje się, że jej inteligencja, również emocjonalna, budzi się do życia, dojrzewa. Ava wzbudza w naszym bohaterze całą gamę uczuć: począwszy od kompletnego zaskoczenia i sceptycyzmu, poprzez zainteresowanie, sympatię, zachwyt, aż po totalną ekscytację i oczarowanie. Później dojdą do tego totalnie odwrotne uczucia – strach, niedowierzanie, wstręt, wściekłość. Można by tu wręcz właściwie wymienić całą gamę afektów ze słownika Tomkinsa! Obserwujemy, jak bohater – który jest przecież człowiekiem inteligentnym, wykształconym i zdającym sobie sprawę, że ma do czynienia ze sztucznym wytworem – w miarę rozwoju akcji po prostu zakochuje się w androidzie, którego miał poddać testowi.

Oczywiście test Turinga w tym przypadku – o czym wspomina również Nathan w początkowych sekwencjach filmu – to nie jest już ten klasyczny test. Nie ma on polegać li tylko na prostej zgadywance, czy mamy do czynienia z robotem, czy z człowiekiem. Rozumiemy przecież, że w tej sytuacji nie chodzi o klasyczną Turingowską *imitation game* (Turing, 1995). Caleb wie i zdaje sobie sprawę z tego, że ma rozmawiać z androidem, więc pierwotny test Turinga nie ma tutaj racji bytu. Teraz chodzi już o to, aby zbadać, jakie możliwości może mieć taka sztuczna inteligencja, która wprawdzie została zbudowana przez człowieka, ale nie może on jej kontrolować w stu procentach. Douglas Hofstadter w książce *Gödel, Escher, Bach: An Eternal Golden Braid* pytał o to, w jaki sposób tworzy się nasza świadomość. Był to bowiem dla niego proces zupełnie niewiarygodny i fenomenalny. Doszedł do wniosku, że proste elementy, które budują naszą świadomość i które oczywiście same żadnej świadomości nie mają (mowa tu o atomach czy innych cząstkach elementarnych), tworzą razem swego rodzaju kompletną całość z wartością dodatkową, naddaną. Suma elementów początkowych nie równa się więc sumie elementów wyjściowych! Ta wartość dodatkowa właśnie ma tworzyć inteligencję i świadomość. W *Ex Machina* podobne pytanie jest stawiane z całą mocą – czy możliwe jest powstanie sztucznej inteligencji, która będzie posiadać tę właśnie wartość dodatkową (i czy jest to afekt)? Jeśli tak, to jak ludzie będą się wobec niej zachowywać? A jak sztuczna inteligencja zachowa się wobec nas? Zarówno Hofstadter, jak i Turing, są zdania, że jest możliwe stworzenie tego typu inteligencji – tzn. złożenie wielu prostych elementów, które zaowocują w kreacji pewnej całości, która będzie czymś więcej niż składanką pojedynczych puzzli. Tym nieuchwytnym elementem jest być może afekt (rozumiany w duchu Tomkinsowskim) u Avy, którego świadkiem jest w filmie Caleb. On bowiem, zakochując się w Avie i wierząc, że Ava to uczucie odwzajemnia, zaświadcza, że – jego zdaniem – Ava jest istotą posiadającą emocje do złudzenia przypominające ludzkie.

Bycie świadkiem interesuje nas tutaj najbardziej. Dlaczego kategoria bycia świadkiem i dawania świadectwa jest szczególnie ważna? To, jak opowiadamy o danej historii, którą zapamiętaliśmy, oraz to, w jaki sposób dajemy jej świadectwo – świadczy nie tylko o tym, co jest dla nas ważne, ale również

o okolicznościach, w których się znajdujemy. Sytuacja, w której dajemy świadectwo, może również wpływać na nie samo (pomyślmy np. o składaniu zeznań pod przysięgą w sądzie czy o wymuszaniu zeznań torturami. Są to sytuacje, które nie pozostawiają świadka obojętnymi i mogą mieć wpływ na kształt zeznań). Caleb nie opowiada przecież wszystkiego, czego się dowiedział od Avy, Nathanowi – wręcz przeciwnie. Zachowuje jej historię dla siebie w nadziei, że może przechrzcić swojego chlebobdawcę i uwolnić ukochaną z więzienia.

Hayden White w eseju *The Modernist Event* pisał o stylu filmowym, który mógłby – według krytyków filmu Olivera Stone’a *JFK* – „wypaczyć” zdolność widza do formułowania własnych wniosków, własnych prawd na temat historycznego wydarzenia (White, 2000). Wydaje mi się, że nie tylko o stylu filmowym możemy tak powiedzieć. Jest oczywiście wiele elementów, które wprowadzają swoiste alteracje w naszej percepcji zdarzeń. Odnajdujemy je na poziomie opowiadanych historii w *Ex Machina* oraz *Uncanny*. Tam bohaterowie-świadkowie są jednocześnie bohaterami oszukanyymi. Ich świadectwa są wypaczone przez: 1) emocjonalny stosunek do obiektu badanego, 2) nieświadomość istnienia pewnych zatajonych przez osoby trzecie elementów rzeczywistości (przez Nathana w *Ex Machina* i Davida w *Uncanny*), 3) nieznamość reguł gry prowadzonych przez same maszyny (czy to na polecenie przełożonego czy przeciwko niemu). Czy możemy powiedzieć, że bohaterom udało się w końcu dać świadectwo spotkaniu ze sztuczną inteligencją? Śmierć Caleba w *Ex Machina* jest w pewnym, być może okrutnym, sensie takim świadectwem. Nie wiemy, jak potoczyły się dalsze losy Joy, gdy dowiedziała się, że jej wybranek, w którym zdążyła się zakochać i z którym uprawiała seks, nie jest człowiekiem, lecz robotem. Artykuł do czasopisma w każdym razie nie powstał.

## Podsumowanie

Filmy *Ex Machina* i *Uncanny* przedstawiają wizję naszej przyszłości, w której istnienie maszyn obdarzonych inteligencją jest faktem. Są one tutaj wynikiem skomplikowanych doświadczeń laboratoryjnych i badań naukowych. W filmach jesteśmy świadkami momentów przełomowych, swoistych punktów zwrotnych w naszym doświadczeniu ze sztuczną inteligencją – w tym momencie jest ona już faktem, a nie tylko teorią. Jest oczywiste, że człowiek, próbując się odnaleźć w tej nowej sytuacji, błądzi i szuka rozmaitych „punktów zaczepienia”, chcąc poradzić sobie z czymś wyjątkowym i nieznanym. Emocje towarzyszą nam od zawsze, a afekty są wrodzonym i stałym elementem naszego istnienia – chyba oczywistym jest więc, że to do nich człowiek będzie się uciekał, aby stworzyć, poznać i zrozumieć sztuczną inteligencję, zwłaszcza po porażce związanej z tzw. podejściem tradycyjnym do AI, traktowanej jako *stricte* logiczna i „obliczeniowa” figura<sup>1</sup>.

Filmowe wizje, przeanalizowane w niniejszym artykule, świadczą o wzrastającym zainteresowaniu sztuczną inteligencją przez artystów filmowych – zwłaszcza jej aspektem afektywnym, a także możliwościami i potencjalnymi zagrożeniami z nią związanymi. Afekt, jako nieodłączna część sztucznej

1 O kluczowej roli emocji w budowaniu sztucznej inteligencji pisze Elizabeth A. Wilson w książce *Affect & Artificial Intelligence*. Przekonująco dowodzi ona, że element afektywny/emocjonalny był obecny w badaniach nad sztuczną inteligencją od zawsze, lecz był spychany na margines na korzyść tradycyjnej, „obliczeniowej” (ang. *computational*) sztucznej inteligencji. Według badaczki jest to błąd – afekt z całą pewnością powinien być wpisany w badania nad AI.

inteligencji, to element kluczowy w opowiadanych historiach. Świadcami istnienia afektu w maszynach są Caleb i Joy – główni „ludscy” bohaterowie analizowanych filmów. Ich reakcje są różne – od zdziwienia poprzez entuzjazm, aż po strach i wstręt. Próbuje oni radzić sobie w nowych sytuacjach, lecz wydaje się, że oboje ponoszą porażkę – ich czujność zostaje uśpiona, a świadectwa stają się w pewnym sensie wypaczone. Lecz może właśnie to jest ich wartością – fakt, że bohaterowie byli w stanie zakochać się w maszynach, staje się swego rodzaju świadectwem, które daje na temat sztucznej inteligencji.

Oba filmy są znakomitymi przykładami kina w duchu posthumanistycznym, gdzie to podmiot „nie-ludzki” gra pierwsze skrzypce, a człowiek – mimo pozorów sprawczości – jest tylko jednym z elementów uniwersum, wcale nie najważniejszym. Takie postrzeganie świata zyskuje dzisiaj na popularności. Czy w rzeczywistym świecie androidy obdarzone sztuczną inteligencją, podobne tym z filmów, staną się wkrótce częścią uniwersum? Czas pokaże.

#### BIBLIOGRAFIA

- Arbiszewski K. (2008). *Poznanie, zbiorowość, polityka. Analiza teorii aktora sieci Bruno Latoura*. Kraków.
- Bednarek J. (2014). Przedmowa do polskiego wydania, [w:] Braidotti R., *Po człowieku*. Warszawa.
- Damasio A. (2011). *Błąd Kartezjusza. Emocje, rozum i ludzki mózg*. Poznań.
- Domańska E. (2008). *Humanistyka nie-antropocentryczna a studia nad rzeczami*. „Kultura Współczesna”, 3, (9–21).
- Freud S. (1997). *Niesamowite*, [w:] tegoż, *Pisma psychologiczne*, t. 3, cz. II. Warszawa.
- Hofstadter D.R. (1984). *Gödel, Escher, Bach: An Eternal Golden Braid*. New York.
- Jelewska A. (2012). *Sensorium. Eseje o sztuce i technologii*. Poznań.
- Kosofsky-Sedgwick E. i Frank A. (2014). *Wstyd w czasie cybernetycznej analogii*. Czytając Silvana Tomkinsa, [w:] Budrewicz Z., Sendyka R., Nycz R. (red.), *Pamięć i afekty*. Warszawa.
- Levy D. (2009). *Love and Sex with Robots: The Evolution of Human-Robot Relationships*. London.
- Mendelsohn D. (2015). *The Robots Are Winning!* „The New York Review of Books”, 4 June.
- Radkowska-Walkowicz M. (2008). *Od Golema do Terminatora. Wizerunki sztucznego człowieka w kulturze*. Warszawa.
- Sendyka R. (2014). *Nowe przestrzenie humanistyki: pamięć, afekty i inne terytoria*, [w:] Budrewicz Z., Sendyka R., Nycz R. (red.), *Pamięć i afekty*. Warszawa.
- Thomson P. (2008). *Field*, [w:] Grenfell M. (red.), *Pierre Bourdieu. Key Concepts*. Acumen.
- Turing A.M. (1995). *Maszyna licząca a inteligencja*, [w:] Chwedeńczuk B. (red.), *Filozofia umysłu*. Warszawa.
- White H. (2000). *The Modernist Event*, [w:] tegoż, *Figural Realism: Studies in the Mimesis Effect*. Baltimore & London.
- Wilson E.A. (2010). *Affect & Artificial Intelligence*. Seattle and London.

#### FILMOGRAFIA

- Ex Machina*. (2015). USA.
- Łowca androidów (Blade Runner)*. (1982). Hong Kong, USA, Wielka Brytania.
- Uncanny*. (2015). USA.