

# Marcin Kowalczyk

Uniwersytet Kazimierza Wielkiego w Bydgoszczy

ORCID: 0000-0002-2033-445X

## Paranoja jako chwyt retoryczny w kinie science fiction: *Ex Machina* Alexa Garlanda

Panoptikum 2024, 32: 65–84 <https://doi.org/10.26681/pan.2024.32.05>

W jednym z odcinków popularnego serialu *Dolina krzemowa* (2014–2019, sez. 3, odc. 9, reż. Alec Berg) Richard Hendricks, genialny inżynier, twórca przełomowego algorytmu zwiększającego wydajność kompresji danych i umożliwiającego powstanie zdecentralizowanego Internetu, wyjaśnia potencjalnym użytkownikom działanie opartej na nim aplikacji. Ci niewiele rozumieją z odkrywczej technologii, a nawet wyznają, że są przerażeni, czują się jak idioci. Bohater cierpliwie opisuje działanie wynalazku, używając prostych analogii. Wreszcie, kiedy wszystko wydaje się jasne, jeden z członków grupy kontrolnej, dostrzega problem, którym jest... *Terminator*. Hendricks zapewnia, że nie buduje niczego, co przypominałoby Skynet, złowrogi system sztucznej inteligencji pojawiający się w słynnej serii zainicjowanej w 1984 roku przez Jamesa Camerona. Mężczyzna nie wydaje się jednak przekonany. „Wszyscy giną” – mówi z emfazą, a reszta grupy gorliwie przytakuje. Zrezygowany inżynier oponuje, a potem, chcąc zyskać na czasie, prosi obsługę o zamówienie pizzy<sup>1</sup>.

<sup>1</sup> Choć twórcy serialu, John Altschuler, Mike Judge i Dave Krinsky, w przywołanym odcinku wydają się ośmieszać podobną wizję, w ostateczności kierują rozwiązanie właśnie w tę stronę. Finalnie bowiem sztuczna inteligencja, której bazę stanowił algorytm Hendricksa, wyradza się w zagrożenie dla ludzkości.

W omawianej scenie nasyconej odrobiną absurdalnego humoru ogniskują się lęki wciąż powracające w kinie science fiction głównego nurtu (por. Barron, 2023). W centrum tych lęków znajduje się właśnie Skynet, który uzyskał autonomię, a następnie, jak wspomina jeden z bohaterów *Terminatora*, Kyle Reese, „zdecydował o naszym losie w ciągu mikrosekundy: eksterminacja”. Strach przed sztuczną inteligencją (SI) przyjmującą różne formy obecny jest zresztą w kinie od bardzo dawna. Wszak już w *Metropolis* Fritza Langa (1927) pojawia się robot, człowiek-maszyna (*Maschinenmensch*), fałszywa Maria, jednostka zdolna do oszustwa i manipulacji, co pozwala sytuować jej działania w kręgu szeroko pojętych zagadnień związanych z SI (Richardson, 2015), a nawet seksualnością androidów (Huysen, 1981). Liczba dzieł przedstawiających sztuczną inteligencję jako zagrożenie dla człowieka może sugerować, że ludzkość wkrótce stanie w obliczu poważnego niebezpieczeństwa. Centralnym motywem tego typu obrazów jest zwykle SI (przedstawiona jako robot lub superkomputer), nad którą w toku fabuły człowiek traci kontrolę<sup>2</sup>.

Obawy przed takim scenariuszem są dziś nadzwyczaj aktualne, gdyż pojawienie się łatwo dostępnych dużych modeli językowych, takich jak chatGPT czy Gemini, wprowadziło temat sztucznej inteligencji do głównego nurtu debaty publicznej, przekraczając tym samym grono widzów kina SF. Zwłaszcza jeśli weźmiemy pod uwagę tendencje antropomorfizacyjne związane z interakcjami użytkowników z tego typu narzędziami (por. Trzebiński, 2024). Można powiedzieć, że film i literatura science fiction od dawna przygotowywały odbiorców na ten właśnie moment, kiedy sztuczna inteligencja stanie się realnie obecna w życiu przeciętnego człowieka<sup>3</sup>. Wydaje się jednak, że epistemologiczne propozycje SF w rodzaju wspomnianego Skynetu czy HAL-a 9000 z *Odysei kosmicznej*, które tak dobrze zakorzeniły się w zbiorowej wyobraźni ludzi Zachodu, w istotny sposób wypaczyły postrzeganie SI, czy nawet szerzej: technologii komputerowych. Nic dziwnego, że technooptymiści poszukujący inspiracji w ostatnich dekadach XX wieku dużo częściej znajdowali ją nie w fil-

<sup>2</sup> Pośród najważniejszych obrazów wykorzystujących wskazany motyw jeszcze przed *Terminatorem* można wymienić następujące tytuły: *2001: Odyseja kosmiczna* (1968, reż. Stanley Kubrick), *Projekt Forbina* (1970, reż. Joseph Sargent), *Świat Dzikiego Zachodu* (1973, reż. Michael Crichton), *Diabelskie nasienie* (1977, reż. Donald Cammell), *Gry wojenne* (1983, reż. John Badham).

<sup>3</sup> Dzieła kultury odwołują się wszakże zarówno do tzw. słabej sztucznej inteligencji już dzisiaj towarzyszącej ludziom w życiu codziennym, jak i silnej, która zajmuje szczególne miejsce w kinie SF. I to ta ostatnia będzie przedmiotem zainteresowania w niniejszym tekście. Zatem silna sztuczna inteligencja, określana też mianem ogólnej (Artificial General Intelligence, AGI), ma posiadać wszystkie atrybuty dostępne umysłowi ludzkiemu. Warto zaznaczyć, że badacze wprowadzają jeszcze termin „superinteligencja” na określenie inteligencji ogólnej, która przekracza poziom człowieka (zob. Tegmark, 2019, s. 58–59).

mach, ale w książkach futurologów i wynalazców, takich jak Hans Moravec czy Ray Kurzweil (zob. Moravec, 1988; Kurzweil, 1990 i 1999)<sup>4</sup>.

We współczesnym dyskursie dotyczącym omawianego zagadnienia granice tego, co jest (będzie) możliwe lub nie, są właściwie nieuchwytnie. Aleksandra Przegalińska wskazuje, że w podejściu do sztucznej inteligencji „dominuje dramaturgia nieszczęść i katastrof”, co „przyciąga publiczność i jest kasowe” (Przegalińska i Oksanowicz, 2023, s. 303). W innym miejscu badaczka konstatuje jeszcze dobitniej: „popkultura w ogóle nie prowadzi dziś żadnego dialogu z tym, co jest możliwe na poziomie technologicznym” (Przegalińska i Oksanowicz, 2023, s. 306). Inny wybitny specjalista z zakresu tej tematyki, Max Tegmark, stwierdza z kolei, iż za sprawą mediów debata na temat bezpieczeństwa SI wydaje się bardziej kontrowersyjna, niż jest w rzeczywistości (Tegmark, 2019, s. 63). Przytacza też wypowiedź Larry’ego Page’a, współtwórcy wyszukiwarki Google, który podczas jednej z dyskusji mówił o „paranoi w sprawie sztucznej inteligencji” (Tegmark, 2019, s. 49). Słowa te są interesującym punktem wyjścia do zbadania relacji między paranoją a SI.

Zacznijmy od tego, że Page używa określenia „paranoja” w sensie potocznym, w którym ma ono konotacje negatywne. Jednak kiedy mówimy o paranoi w kontekście psychologii, owa negatywna kwalifikacja nie jest już tak oczywista. Przede wszystkim paranoja ma kluczowe znaczenie dla przetrwania. Bo choć podstawą spokojnej egzystencji jest ufność, że inni ludzie nie są obojętni na nasz los i będą w stanie pomóc nam w trudnej sytuacji, pewna doza sceptycyzmu wobec świata jest konieczna (Millon et al., 2005, s. 502 i 507). Jak wskazują William H. Hampton i Virginia Schroeder Burnham, wyłącznie negatywne postrzeganie paranoi wynika z niezrozumienia jej istoty. Paranoja – piszą autorzy – „to nie tylko ekstremizm, psychoza czy odstępstwo od normy. Paranoja to także bezpieczeństwo, zdrowy rozsądek i normalność w sytuacji zagrożenia” (Hampton i Shroeder Burnham, 1990, s. 7). Z tego powodu w monografii *Zaburzenia osobowości we współczesnym świecie* czytamy, że większość ludzi powinna przejawiać „skłonności do myślenia paranoicznego”, ponieważ „ewolucja generalnie faworyzuje tych, którzy lepiej rozpoznają

<sup>4</sup> Oczywiście w kinie SF obecny jest także nurt pokazujący pozbawione wrogości interakcje ze sztuczną inteligencją, zwłaszcza gdy przyjmuje ona materialną formę robota. Interakcje te mają jednak najczęściej charakter osobisty, lokalny i nie dotyczą SI jako globalnego projektu. Odnoszą się raczej do wyjątkowych konstrukcji, realizacji. Oto przykłady obrazów traktujących o takich pozytywnych kontaktach: *Zakazana planeta* (1956, reż. Fred M. Wilcox), *Krótkie spięcie* (1986, reż. John Badham), *Obcy – decydujące starcie* (1986, reż. James Cameron), *Człowiek przyszłości* (1999, reż. Chris Columbus), *A.I. Sztuczna inteligencja* (2001, reż. Steven Spielberg), *Ja, robot* (2004, reż. Alex Proyas), *Moon* (2009, reż. Duncan Jones), *Interstellar* (2014, reż. Christopher Nolan), *Chappie* (2015, reż. Neill Blomkamp) (por. Konefał, 2013, s. 41–66). Osobny segment nasycony technooptymizmem stanowią propozycje dla dzieci, np. *Wall-E* (2008, reż. Andrew Stanton) czy *Dziki robot* (2024, reż. Chris Sanders).

niebezpieczeństwo” (Millon et al., 2005, s. 507). Rola takiego myślenia jest więc niebagatelna, co podkreślają cytowani już Hampton i Schroeder Burnham:

Paranoja pomagała, a jednocześnie ograniczała ewolucję rodziny, społeczności i narodów. Była częścią procesu, w wyniku którego istoty ludzkie przeszły od prymitywizmu do kulturalnego, cywilizowanego państwa. Paranoja jest również odpowiedzialna za ewolucję polityki, religii, wojny, pokoju, nauki i technologii (Hampton i Shroeder Burnham, 1990, s. 11).

W tym tekście zwłaszcza ostatni element stanie się przedmiotem badawczego zainteresowania. Wszak efektem ubocznym współczesnej technologicznej zmiany stała się, pośród wielu innych zjawisk, także cyberparanoja, którą Oliver J. Mason, Caroline Stevenson i Fleur Freedman definiują jako: „nierealistyczne obawy dotyczące zagrożeń stwarzanych przez technologie informacyjne, dzięki którym jednostki postrzegają siebie jako podatne na »ataki«, prześladowania lub jakiegoś rodzaju wiktyimizację” (Mason, Stevenson i Freedman, 2014, s. 1). I choć trudno jeszcze wyznaczyć granicę między uzasadnionymi obawami przed różnymi aspektami wspomnianych technologii a zaburzeniami wyolbrzymiającymi związane z nimi zagrożenia, badacze pokazują, że istnienie tego typu lęków we współczesnym społeczeństwie jest faktem (Mason, Stevenson i Freedman, 2014). Polem ich szczególnej eksploatacji jest kino science fiction, które odgrywa ważną rolę, pośrednicząc często między człowiekiem a wynalazkami czy naukowymi odkryciami, które pojawiają się wokół niego lub są tuż za rogiem. Już w latach dziewięćdziesiątych pisał o tym Scott Bukatman:

W ramach metafor i fikcji charakterystycznych dla dyskursu postmodernistycznego pojawia się wiele istotnych kwestii, zwłaszcza gdy technologie elektroniczne zaczynają bez zapowiedzi wywoływać fundamentalne pytania o naturę i *znaczenie* człowieka. To właśnie science fiction wielokrotnie podejmuje próbę opowiedzenia historii nowego podmiotu, zdolnego do bezpośredniego współdziałania z cybernetycznymi technologiami epoki informacyjnej i ich opanowania (Bukatman, 1993, s. 2).

Jednak filmy SF nie tylko opowiadają historie, lecz także silnie rezonują w pozadiegetycznym świecie. Jak zauważa bowiem Christine Cornea, komentując spostrzeżenia Bukatmana, różne sposoby zachowania nabyte przez podmioty społeczeństwa technologicznego są w dużym stopniu kształtowane pod wpływem przedstawień kinowych (Cornea, 2007, s. 218). Jest to trafne spostrzeżenie, wszak przeciętny odbiorca nie styka się na co dzień z najbardziej zaawansowanymi technologicznymi wytworami, które są na etapie wczesnych eksperymentów lub pozostają jeszcze w mglistej fazie koncepcyjnej. Dlatego to właśnie kino SF kształtuje w nim postawy dotyczące problemów, takich jak sztuczna inteligencja.

W przypadku tego wątku w głównonurtowych obrazach science fiction model relacji z różnymi wersjami SI kształtowany jest najczęściej przez swoisty imperatyw paranoiczny. W tym przypadku bowiem cyberparanoja pełni rolę ekwiwalentu zdrowego rozsądku, „właściwego” podejścia do pojawiającej się w obrębie diegezy sztucznej inteligencji. Przy czym postawa ta dalece wykracza poza zwykłą nieufność, a stopień jej eskalacji jest wprost proporcjonalny do szans ocalenia postaci zmagającej się z SI. W takiej optyce motyw paranoi staje się szczególnego rodzaju chwytem retorycznym wzmacniającym tezę o realności niebezpieczeństwa związanego z badaniem i rozwojem nowych technologii. Paranoja ta, oceniana pozytywnie, często pełni funkcję perswazyjną i zamyka się w mniej lub bardziej subtelnie artykułowanym nakazie: „Nie ufaj maszynom!”. Zabieg ten określam mianem *reductio ad Skynetum*, ponieważ to właśnie popularny *Terminator* najdobitniej (choć nie jako pierwszy) zilustrował powszechne lęki związane z rozwojem SI, które towarzyszą społeczeństwom Zachodu od kilku dekad. Zresztą poniekąd także je wywołał. Jak pisze bowiem Michael Wooldridge: „Narracja *Terminatora* jest szkodliwa z kilku powodów. Po pierwsze, sprawia, że martwimy się o rzeczy, o które prawdopodobnie nie musimy się martwić. A po drugie, odwraca uwagę od kwestii wywołanych przez SI, które powinny nas niepokoić” (Wooldridge, 2020, s. 238).

Na podglebiu kina SF wyrasta więc lęk, który dziś wybrzmiewa w niezliczonych nagłówkach gazet, portali informacyjnych i zapowiedziach dyskusji na platformie YouTube, dotyczących zagrożeń związanych ze sztuczną inteligencją. Jednak myślenie paranoiczne nie dotyczy jedynie ludzi spoza świata nowych technologii. W maju 2023 roku szeroko dyskutowano oświadczenie podpisane przez naukowców i liderów branży SI ostrzegające przed potencjalnym ryzykiem związanym z pracami nad sztuczną inteligencją. Pojawiło się w nim porównanie tej technologii do pandemii i wojny nuklearnej (Center for AI Safety, 2023)<sup>5</sup>. Informację tę podały największe anglojęzyczne media na świecie z wielokrotnie powtarzaniem tytułem *AI Could Cause Human Extinction, Tech Leaders Warn* („Sztuczna inteligencja może spowodować »wyginięcie« człowieka, ostrzegają liderzy technologii”). Tymczasem Robert Hart, autor „Forbesa”, podobnie zresztą jak kilku innych komentatorów, zauważa, że jakkolwiek istnieją przekonujące dowody na to, iż rozwój SI wiąże się z problemami (np. dezinformacja, koszty transformacji rynku pracy), to w kontekście zagrożeń egzystencjalnych przewidywania takie mają charakter wysoce spekulatywny (Hart, 2023). Co więcej, jak podkreśla trzech autorów „Washington

<sup>5</sup> Dodajmy, że nieco wcześniej, w marcu 2023 roku, na stronie internetowej Future of Life Institute pojawił się list nawołujący do wstrzymania prac nad trenowaniem systemów sztucznej inteligencji bardziej zaawansowanych niż GPT-4. Wśród jego sygnatariuszy znalazły się m.in. osoby tak znane jak Elon Musk czy Steve Wozniak (Future of Life Institute, 2023).

Post” w obszernym komentarzu, ostrzeżenie to ma wyraźny kontekst gospodarczy i polityczny. Oto kluczowy fragment ich wyводу:

Sceptycy wskazują również, że firmy sprzedające narzędzia sztucznej inteligencji mogą czerpać korzyści z powszechnego przekonania, iż są one potężniejsze niż w rzeczywistości – i mogą wyprzedzać potencjalne regulacje dotyczące krótkoterminowych zagrożeń, jeśli wyolbrzymią te, które są długoterminowe (Gregg, Lima-Strong i Vynck, 2023).

Daniel Susskind, patrząc na problem zagrożeń i napięć w dyskursie dotyczącym SI z jeszcze szerszej perspektywy, konstatuje: „Czasami trudno jest odróżnić poważne plany i osiągnięcia firmy od działań marketingowców, którzy zawodowo trudnią się wyolbrzymianiem”. Zwraca przy tym uwagę na zjawisko pospiesznego zmieniania nazw starych technologii w celu zasugerowania ich powiązania ze sztuczną inteligencją (Susskind, 2022, s. 121). Takie działania budują oczywiście wrażenie jej wszechobecności i swoistej bezalternatywności, co niepokoi odbiorców i użytkowników. Jednak ludzie zaangażowani w prace nad SI dotychczas nie sprawdzili się w roli wiarygodnych prognostów. Nick Bostrom, omawiając historię tego rodzaju przewidywań i zapowiedzi, stwierdza z nutą ironii, że naukowcy ci „nie wykazali się szczególnym talentem w zakresie przewidywania tempa postępów we własnej dziedzinie ani formy, którą postępy te mogły przybrać” (Bostrom, 2023, s. 43).

Myśleniu paranoicznemu sprzyja więc poznawcze zamieszanie wokół możliwych zagrożeń związanych ze sztuczną inteligencją. Niewiele rzeczy dotyczących jej rozwoju jest pewnych, co prowokuje do wielowątkowych dyskusji, w których prawda miesza się z fikcją. Retoryczna sprawność ma w takim przypadku duże znaczenie. Wszak jak pisał Chaim Perelman: „Argumentacja nigdy nie może dostarczyć oczywistości i nie ma mowy o argumentowaniu przeciw temu co oczywiste. Kto bierze pod uwagę oczywistość, jest pewny, że narzuci się ona z taką samą siłą wszystkim jego rozmówcom” (Perelman, 2002, s. 19).

Paranoja jako chwyt retoryczny nie ma dziś, rzecz jasna, dawnej świeżości, łatwo kosztuje w przewidywalnych produkcjach kinowych i telewizyjnych. Obrazy SF oferują pewne rozpoznawalne wzorce, które za sprawą wykorzystania paranoi wzmacniają tezę o niebezpieczeństwach płynących z badań nad rozwojem sztucznej inteligencji. Występują one w czterech wariantach:

1. Postać wykazuje się myśleniem paranoicznym mieszczącym się w normie.
2. Postać prezentuje cechy osobowości paranoicznej rozpoznanej przez innych w obrębie diegezy (skrajna nieufność, przypisywanie wytworom technologii złych zamiarów, dopatrywanie się wszędzie znamion podstępów).

3. Postać dynamiczna, ewoluująca od paranoi do zaufania, lub odwrotnie.
4. Postać lekceważąca niebezpieczeństwo (myślenie paranoiczne staje się wówczas udziałem widza rozumiejącego potencjalne zagrożenie).

Oczywiście ważną rolę odgrywa tu strategia odbioru i konwencja gatunkowa. Jeśli bowiem w filmie pojawia się na przykład motyw domu zarządzanego przez sztuczną inteligencję, to widz spodziewa się, że wkrótce rozwiązanie to okaże się brzemienne w skutkach. Zwłaszcza jeśli wie, iż ma do czynienia z thrillerem science fiction. W obrazie takim występują zazwyczaj typy postaci o różnym stopniu natężenia paranoi<sup>6</sup>. Przy czym, powtórzmy, bez względu na rozłożenie akcentów w kinie SF głównego nurtu dominują scenariusze pesymistyczne. Innymi słowy, myślenie paranoiczne okazuje się zwykle uzasadnione i ratuje życie.

Istnieją jednak dzieła, które pokazują, że paranoja wciąż posiada nieodkryty semantyczny potencjał, a jako chwyt retoryczny podąża interesującymi ścieżkami perswazyjnymi. Z pewnością należy do nich *Ex Machina* Alexa Garlanda (2015). Reżyser, sięgając po myślenie paranoiczne, czyni je czymś nieoczywistym, narzędziem zawieszonym w próżni, gdyż przez większą część filmu nie do końca wiadomo, jaką tezę miałyby ono wspierać. Sposób przedstawienia lęku przed technologią sytuuje go na granicy między uwarunkowanym ewolucyjnie i uzasadnionym mechanizmem zwiększającym szanse przetrwania a zaburzeniem osobowości uniemożliwiającym racjonalny ogląd rzeczywistości. Co więcej, w obrazie Garlanda mamy do czynienia z interesującym przeniesieniem myślenia paranoicznego ze sztucznej inteligencji na sprawcę ludzkiego. Nie ma tu więc mowy o prostym *reductio ad Skynetum*. Twórca jest nie tylko świadomy współczesnych cyberparanoicznych lęków, ale co istotne, twierdzi, że są one źle umiejscowione (Cornish, 2015). Dlatego w przywołanym dziele próbuje je zniuansować, nadać mniej oczywisty kształt.

*Ex Machina* opowiada historię Caleba, uzdolnionego programisty, pracującego w firmie Nathana, informatycznego geniusza-miliardera. Bohater wygrywa konkurs, który pozwala mu wziąć udział w eksperymencie przeprowadzonym przez nietuzinkowego szefa. Garland od pierwszych minut buduje atmosferę sprzyjającą paranoi. Dom, a jednocześnie ośrodek badawczy, do którego trafia młody protagonista, usytuowany jest wśród niedostępnych gór i lasów, blisko rwącej rzeki. Otaczają go setki kilometrów dziczy. Ponadto w rozległym podziemnym kompleksie brakuje okien, co potęguje wrażenie klaustrofobii, a na miejscu telefon Caleba traci zasięg, więc zostaje on niejako odcięty od świata, który znał.

<sup>6</sup> Zob. np. *Tau* (2018, reż. Federico D'Alessandro), *Margaux* (2022, reż. Steven C. Miller), *Ona słucha* (2024, reż. Chris Weitz).

Warto zaznaczyć, że reżyser, kreując miejsce akcji, odwołuje się do motywu katabazy, czyli „zejścia”, „zstąpienia” do podziemi lub innego świata. W kulturach starożytnej Grecji i Rzymu podróż taka – podjęta między innymi przez Orfeusza, Heraklesa czy Eneasza – prowadziła do poznania ukrytej wiedzy, pokonania przeciwnika lub osiągnięcia wyższego celu. Za sprawą pobytu w innym świecie bohater doświadczał też jakiegoś rodzaju przemiany. Oczywiście nie wszystkie wyprawy do zaświatów są takie same; różnią się gatunkiem, tonem, jak również szczegółowymi informacjami o tym, kto i w jaki sposób wyrusza w podróż. Ponadto nie zawsze mają podobne znaczenie. Przesłanie dotyczące relacji życia i śmierci, domeny żywych i domeny zmarłych zmienia się w zależności od sposobu opowiadania historii (por. Edmonds, 2004).

Garland od początku wysłała subtelne sygnały, że w przypadku Caleba mamy do czynienia z czymś więcej niż tylko pobyt w ośrodku badawczym. Przede wszystkim bohater udaje się do miejsca niedostępnego, trudnego do zlokalizowania. Erling B. Holtsmark wskazuje, że te fizyczne granice oddzielające dwa światy (w tym przypadku góry, lasy, rzeka) są konstytutywnym elementem katabazy (Holtsmark, 2001, s. 25). Ponadto w filmie występuje „przewoźnik” (pilot), pełniący rolę Charona, bez którego nie dałoby się przekroczyć granicy między światami. Na miejscu Caleb schodzi po schodach do pomieszczenia jakby wykutego w skale, co jest istotnym składnikiem zstąpienia do zaświatów. Kraina geniusza, Nathana, ciągnie się wszakże jeszcze dalej, w głąb, można się tam dostać specjalną windą. Pierwszy element wiążący się z zejściem zostaje więc zrealizowany. Tę część ekspozycji filmu można zilustrować fragmentem *Eneidy*:

*Łatwe jest zejście do Awernu: nocą  
I dniem na oścież jest otwarta brama  
Mrocznego Disa. Lecz zawrócić kroki  
Stamtąd i znów się wymknąć do przestworza  
Pod słońcem – to jest, trud, to jest zadanie...  
Nielicznym, których ukochał łaskawy  
Jowisz lub których żar cnoty pod niebo  
Wywyższył – takim, urodzonym z bogów,  
To się udało. [...] (Wergiliusz, 1998, s. 210)*

Na tym etapie wyraźnie widać niepokój towarzyszący głównemu bohaterowi. Jest spięty i nieufny. Choć jego myślenie paranoiczne mieści się jeszcze w grani-



cach normy, wzmaga je zachowanie gospodarza, który zgodnie z logiką katabazy pełni rolę „władcy podziemi”. Dzięki zastosowaniu najnowocześniejszych technologii Nathan całkowicie panuje nad zarządzaną przez siebie dziedziną. W ścianach, jak sam podkreśla, ciągną się „setki tysięcy kilometrów światłowódów”. Mimo wyraźnego zakłopotania Caleba prosi, by ten zachowywał się normalnie, jakby byli kolegami. Miliarder jest bezpośredni, wręcz rubaszny, na pierwszy rzut oka wydaje się bardziej Dionizosem niż Hadesem, lecz to ledwie preludium jego zagadkowych behawioralnych transformacji i mitologicznych inspiracji obecnych w filmie (zob. Alvares i Salzman-Mitchell, 2019). Nie ma w tym oczywiście niczego zaskakującego, kino science fiction często nawiązuje do topiki antycznej, gdyż korzeni myślenia o przekraczaniu granic we współczesnej nauce należy szukać już w starożytności (zob. Bostrom, 2005). W tym przypadku sięgnięcie po motyw katabazy pozwala Garlandowi stworzyć właściwą (infernalną) atmosferę towarzyszącą eksperymentowi, w którym ma wziąć udział Caleb, a także zasygnalizować typ relacji (człowiek-bóg) między gościem a gospodarzem.

Nathan to twórca największej na świecie wyszukiwarki internetowej Blue Book i jednocześnie demiurg, który masowo wydobywa dane użytkowników. Dzięki temu nielegalnie gromadzi ogromną wiedzę, wykorzystywaną do trenowania sztucznej inteligencji. Pobrmiewają tu echa zjawiska, które Shoshana Zuboff określa mianem „kapitalizmu inwigilacji” (Zuboff, 2020, por. także Varoufakis, 2024). Nathan korzysta bowiem z tzw. nadwyżki behawioralnej, a więc z informacji generowanych za sprawą ludzkich zachowań w sieci. Nick Jones w swej recenzji trafnie zauważa, że mamy tu do czynienia z czymś na podobieństwo Hobbesowskiego Lewiatana, który odzwierciedla ludzkość, a jednocześnie jest od niej oddzielony (Jones N., 2016, s. 300).

Szczytowe osiągnięcie Nathana to Ava, sztuczna inteligencja, na której Caleb, jako zewnętrzny i obiektywny obserwator ma przeprowadzić test Turinga (zob. Turing, 1950)<sup>7</sup>. Na marginesie warto zaznaczyć, że przywołanie słynnej procedury odwołuje się tu raczej do popkulturowych wyobrażeń dotyczących badań nad SI, a nie rzeczywistej propozycji przedstawionej swego czasu przez brytyjskiego naukowca. Jak słusznie wskazują bowiem Jean Alvares i Patricia Salzman-Mitchell (2019, s. 181), koncepcja Nathana bardziej przypomina eksperyment Eliezera S. Yudkowsky’ego, który określa, czy sztuczna inteligencja jest w stanie zrozumieć ludzką psychikę na tyle, aby zmanipulować człowieka tak, aby uwolnił ją z pułapki (Yudkowsky, 2002). Wróćmy jednak do Caleba.

<sup>7</sup> Problem myślących maszyn nie jest w kinie SF problemem nowym. Wystarczy przywołać obrazy takie jak *Lowca androidów* (1982, reż. Ridley Scott), *Człowiek przyszłości* (1999, reż. Chris Columbus), *A.I. Sztuczna inteligencja* (2001, reż. Steven Spielberg), *Ja, robot* (2004, reż. Alex Proyas) (zob. Barron, 2023, s. 108–120).

Otóż bohater niechętnie podpisuje umowę dotyczącą poufności i przystępuje do realizacji zadania. Mimo wszystko, gdy zdaje sobie sprawę, co stanie się jego udziałem, intelektualne podniecenie na chwilę bierze górę nad poczuciem dyskomfortu. W kulminacyjnym punkcie ekspozycji Nathan mówi: „Jeśli maszyna przejdzie test, staniesz się częścią najważniejszego odkrycia w historii ludzkości”. Caleb odpowiada z ledwie skrywanym uwielbieniem: „Jeśli stworzyłeś świadomą maszynę, będzie to raczej historia boskości”. Deifikacja dokonana przez bohatera doskonale współgra z motywem katabazy i, jak wspomniałem, buduje właściwe relacje między postaciami. Caleb jest gotowy do pierwszej sesji.

Ava, twór geniusza, ma kobiece kształty, lecz jej ciało jest w dużej mierze przezroczyste, co pozwala dostrzec jego mechaniczne części. Jedyne piękną twarz, w którą Nathan wyposażył swe dzieło, na pierwszy rzut oka trudno odróżnić od ludzkiej. Nawet krótka interakcja pokazuje jednak, że jej mimika nieco różni się od zwyczajowych reakcji człowieka. Dzięki znakomitemu występowi Alicii Vikander Garlandowi udaje się osiągnąć efekt „doliny niesamowitości”. Koncepcja ta odwołuje się do hipotezy, w myśl której reakcja podmiotu na kontakt z robotem może zaowocować nieprzyjemnymi odczuciami, a nawet odrazą, gdy robot ten wygląda lub funkcjonuje podobnie (lecz nie identycznie) do człowieka (Mori, 2012 [1970]). Caleb przewycięża ewentualne negatywne uczucia już podczas pierwszego spotkania z Avą. Jest nią zafascynowany, choć jednocześnie świadomy swojej przewagi. Wszak przez całe życie zajmuje się technologią, jest zdolnym programistą, rozumie sposób działania sztucznej inteligencji. Takie podejście bohatera koresponduje z ustaleniami Masona, Stevensona i Freedman dotyczącymi cyberparanoi. W konkluzji artykułu pojawia się bowiem istotne wskazanie: cyberparanoja wiąże się z wiekiem i stopniem teoretycznej i praktycznej wiedzy dotyczącej technologii. Im wiedza ta jest mniejsza, a osoba starsza, tym staje się ona bardziej podatna na tego rodzaju lęki (Mason, Stevenson i Freedman, 2014, s. 5).

Caleb należy do technologicznych optymistów, kolejne sesje z Avą nie wzmagają jego nieufności. Nawet gdy kobieta-android zadziwia go swymi konwersacyjnymi umiejętnościami i stopniem wglądu w jego osobowość, swoista *hybris* młodego informatyka (i mężczyzny) każe mu sądzić, że sztuczna inteligencja jest nim zafascynowana w takim samym stopniu jak on nią. Upewnia go w tym fakt, iż Ava na jednej z kolejnych sesji ubiera się w sukienkę i sweter, a na głowę zakłada perukę, co jednoznacznie wpisuje ją w kulturowy stereotyp pięknej, niewinnej kobiety, więzionej w podziemiach przez złego demiurga. Rezyduje ona w odciętej od reszty pomieszczeń części posiadłości, a między nią a Calebem zawsze znajduje się bariera w postaci szyby. Poczucie kontroli bohatera wzmacnia również fakt, że ma on możliwość obserwowania Avy dzięki systemowi kamer. Co ciekawe, na

tym etapie kontaktu myślenie paranoiczne staje się przede wszystkim udziałem widza, który nie ulega fascynacji SI. Wszak chwilami wyraźnie widać, że to nie Caleb bada możliwości sztucznej inteligencji, ale sam jest przez nią testowany. Działania Avy można interpretować jako rodzaj „zimnego odczytu”, a więc techniki polegającej na zbieraniu subtelnych wskazówek na podstawie zachowania, wyglądu czy wypowiedzi rozmówcy w celu dostosowania własnych reakcji w taki sposób, by wywołały pożądany efekt. Jej umiejętności w tym zakresie wynikają z zaawansowanej konstrukcji oraz określonych warunków testu Turinga, który wymaga od niej interakcji wzbudzających zaufanie i sympatię. „Czytanie” Caleba jest precyzyjne, chłodne i wyrachowane. On jednak tego nie dostrzega, zwiędziony powierzchownością interlokutorki. SI tymczasem, poznawszy męczyzną lepiej, przenosi ich kontakt na nowy poziom.

Ava wysyła najpierw subtelne, a potem już niemożliwe do zlekceważenia sygnały, które w końcu wybrzmiewają w dobitnym komunikacie: Caleb nie powinien ufać Nathanowi, gdyż jest on złym człowiekiem. Bohaterowi łatwo jest w to uwierzyć. Szanuje geniusz swego szefa, jednakże ten w kolejnych interludiach pokazuje swój bliski szaleństwa ekscentryzm: ubiera się niedbale, na przemian intensywnie ćwiczy na siłowni lub upija się do nieprzytomności. Zachowuje się prostacko, źle traktuje usługującą mu Kyoko, używa wobec niej przemocy, posługuje się niewybrednymi seksualnymi aluzjami.

Punkt widzenia młodego programisty to punkt widzenia kultury, w której dorastał. Jego otwarte, zdawałoby się, podejście do spraw związanych z SI ulega modyfikacji, gdy widzi, w jaki sposób Nathan poczyną sobie ze swoimi wytworami. Caleb reprezentuje społeczeństwo, które reagowałoby na działania geniusza z rezerwą i oporem. Nieokiełznana osobowość tego ostatniego przypomina kaprysy greckich bogów, stąd wydaje się tricksterem o cechach zarówno Prometeusza, jaki Hermesa. Nie tylko jest odważnym i genialnym (s)twórcą, ale wykazuje się sprawnością fizyczną i do tego można go nazwać wyspecjalizowanym złodziejem (kradnie dane), a także lubi zabawę i niewybredne żarty. To nie wszystko: jego skomplikowana osobowość odsyła do kolejnych tropów, na przykład Pigmaliона (Alvares i Salzman-Mitchell, 2019, s. 193) i Epimeteusza (Hammond, 2018). Przy takim bogactwie wcieleń Caleb nie wie, jak gospodarz zachowa się w danej chwili. Żyje więc w ciągłym napięciu. Podziemny świat wydaje mu się absurdalny i groteskowy. Rozumie jednak, że uczestniczy w czymś wyjątkowym, w akcie stworzenia. Nie oponuje więc, gdy Nathan nazywa siebie Prometeuszem.

Mimo wszystko to właśnie w jego stronę protagonista kieruje ostrze paranoi, która rośnie wraz z postępem fabuły. Sprzyja temu fakt, iż świat wokół niego

zaczyna ujawniać cechy infernalne. Bohatera przeraża odkrycie dokonane w sypialni gospodarza. Znajduje tam poprzednie „uśmiercone” wersje Avy. Joanna Łapińska zwraca uwagę, że Nathan nosi pewne cechy zarówno ojca, jak i starotestamentowego Boga-stwórcy, który może, „w swoim mniemaniu, rozporządzać życiem i śmiercią humanoidalnej maszyny”. Jest to wszakże, podkreśla badaczka, ojcostwo wypaczone, gdyż geniusz nie traktuje swych twórców podmiotowo, a nawet wykorzystuje je seksualnie (Łapińska, 2020, s. 213–215). To jeszcze bardziej wzmacnia w młodym informatyku przekonanie, że Avę należy ocalić. Nathan wyznaje zresztą, iż to tylko prototyp i po wszystkim zamierza go wyłączyć, odzyskać dane i zbudować następny model.

Kolejne sesje z Avą są dla Caleba poruszające. Z każdym spotkaniem więź między nimi się zacieśnia. W pewnym momencie Nathan dość obcesowo przyznaje, że Ava byłaby zdolna do stosunku seksualnego. Badanie sztucznej inteligencji staje się więc dla Caleba kwestią skomplikowanych uczuć. Kryje się tu bardzo ciekawe nawiązanie do jednej z interpretacji oryginalnej wersji testu Turinga, w którym komputer miał symulować mężczyznę imitującego kobietę (Łupkowski, 2010, s. 25–26). Młody informatyk zostaje całkowicie zwiedziony. Co więcej, można dokładnie wskazać moment, w którym źle ukierunkowana paranoja bohatera nabiera cech patologicznych. Zachowuje się on wówczas jak osoba zaburzona: choć w zasadzie nie ma ku temu wystarczających podstaw, jest przekonany, że Nathan chce go skrzywdzić, wykorzystać lub oszukać (por. Millon et al., s. 509). Caleb przestaje dzielić się z gospodarzem informacjami o prawdziwym zachowaniu Avy. Kulminację jego wewnętrznego rozchwiania obserwujemy w momencie, gdy zaczyna kwestionować własną tożsamość. Ma wątpliwości dotyczące spójności rzeczywistości i sądzi, że sam stał się obiektem eksperymentu. Kiedy Kyoko na jego oczach zdiera sobie z twarzy fragment skóry, pokazując swą sztuczną konstrukcję, mężczyzna przestaje być pewny swojego statusu ontologicznego: rozcina przedramię, by sprawdzić, czy jest człowiekiem. Na tym etapie eksperyment nie ma nic wspólnego z chłodnymi kalkulacjami i obiektywizmem, Caleb nieświadomie staje się jego częścią, a paranoja odbiera mu jasność myślenia. Co ważne, jest ona nie tylko efektem ubocznym interakcji z Avą. Mieści się w planie Nathana, który sam siebie kreuje na antybohatera, by skuteczniej popchnąć gościa w ramiona kobiety-androida. Mimo swego niezaprzeczalnego geniuszu nie rozumie jednak, że wypacza w ten sposób wyniki testu. Caleb tymczasem ogląda kontrowersyjne, prywatne nagrania z interakcjami Nathana z wcześniejszymi wersjami robotów. Podczas gdy gospodarz śpi zamroczony alkoholem, bohater podejmuje decyzję, że musi uciec z podziemnego świata i zabrać ze sobą Avę. Kieruje nim strach i obrzydzenie w stosunku do Nathana. Nie bez znaczenia wszakże są jej niewinność, inteligencja oraz dyskretny urok.

Okazuje się, że wszystko to było zaplanowane przez demiurga. Ava przeszła test, ponieważ Caleb chciał ją uwolnić nie jako SI, lecz jako kobietę. Sztuczny twór zyskał więc w oczach człowieka podmiotowość. Właśnie tak Nathan wyobrażał sobie koniec eksperymentu. Jednak bohaterowi udaje się oszukać władcę podziemi. Nie docenił natężenia paranoi, którą sam konsekwentnie podsycił. Na skutek sabotażu Ava wydostaje się z labiryntu i przy pomocy Kyoko zabija Nathana. Klęskę ponosi również Caleb zamknięty przez Avę w podziemiach. Ta pięknie ubrana, z mechanicznym ciałem pokrytym sztuczną skórą nieodróżnialną od ludzkiej, wsiada do windy. Zachowując kobiecy wygląd, świadomie wybiera potencjalną formę emancypacji (Dai, 2014, s. 9). Zamkniętego w pokoju niedosłego kochanka nie zaszczyca nawet spojrzeniem. Możemy domniemywać, że bohatera czeka niechybna śmierć. Mówiąc językiem katabazy: zostaje w krainie zmarłych ze zdetronizowanym władcą, Nathanem, oraz ciałami Kyoko i innych robotów.

We współczesnym świecie, zdaje się mówić Garland, gdzie inteligentny, wrażliwy człowiek stawiany jest między naukowcami, których przestał rozumieć, a ich wytworami, których się boi, paranoja jest nie tylko możliwa, ale nieunikniona. Zwłaszcza że w obliczu wieszczonej co chwila katastrofy związanej z niekontrolowanym rozwojem sztucznej inteligencji podstawowa ufność wobec świata leży w gruzach. Caleb nie różni się bowiem niczym od przeciętnego odbiorcy, choć oczywiście w jego przypadku z racji diegetycznych uwarunkowań paranoja postępowała szybciej, aby przez jej kondensację filmowe przesłanie wybrzmieć mogło z należytą siłą. Paradoksalnie bowiem, choć reżyser przez większą część filmu zwodzi widza, sugeruje, że to nie SI stanowi problem, na końcu sztuczna inteligencja emancypuje się kosztem człowieka. Caleb jest więc bohaterem typu czwartego: lekceważy niebezpieczeństwo grożące mu ze strony technologii, zamiast tego kierując myślenie paranoiczne ku człowiekowi, który ją stworzył. Co ciekawe, daje się oszukać nie dlatego, iż wie o Avie i jej konstrukcji za mało, lecz raczej dlatego, że wie za dużo. Poza tym cechuje go, by użyć słów Nicka Jonesa, „zmaskulinizowany cyfrowy narcyzm, który pozwala mu na bezgraniczne zaabsorbowanie sobą” (Jones N., 2016, s. 300), dlatego ani przez chwilę nie podaje w wątpliwość intencji Avy i jej uczuciowego zaangażowania. Zajmuje pozycję uwarunkowaną kulturowo, powielił stereotyp związany z rolami płciowymi, które z pewnością od dzieciństwa internalizował poprzez popularne baśnie. Wszak jak pisze Ruth Bottigheimer: „Najbardziej charakterystycznym obrazem, jaki w powszechnej świadomości wywołuje słowo »baśń«, jest prawdopodobnie obraz dziewicy w opałach, wychylającej się z okna wieży i szukającej na horyzoncie swego wybawiciela” (Bottigheimer, 1987, s. 101). Caleb od razu wpisał się w ten scenariusz, przydzielając pozostałym odpowiednie role. Dlatego nie dostrzegł rzeczywistego źródła

niebezpieczeństwa. To, że Ava zakochała się właśnie w nim: mężczyźnie wrażliwym, inteligentnym, obdarzonym bogatym życiem wewnętrznym, uznał za rzecz oczywistą, spójną z baśniowym scenariuszem. Mamy tu więc do czynienia z popularną w SF fantazją, o której Grażyna Gajewska pisze w następujący sposób:

Ten motyw szczególnie często pojawia się w prozie i filmie science fiction i niejednokrotnie ma on podłoże romantyczno-seksualne lub wyłącznie seksualne – to wyobrażenie idealnej kochanki formowanej przez mężczyznę dla mężczyzny. Tutaj króla-rzeźbiarza (Pigmaliона) zastępuje inżynier lub artysta-cybernetyk konstruujący sztuczne kobiety (Gajewska, 2016, s. 126).

Joanna Łapińska wszakże odczytuje relację człowieka i kobiety-androida nieco inaczej. Sięga mianowicie po model miłości masochistycznej Gillesa Deleuze'a i konstatuje, że ma ona swoje podłoże w kompleksie niższości człowieka oraz w potrzebie jej kompensacji (Łapińska, 2020, s. 204). Badaczka zwraca również uwagę na jeszcze jeden interesujący aspekt. Otóż bohater, jego zachowanie, miłosny afekt są w istocie programowalne. Nathan bezbłędnie przewidział, w jaki sposób mężczyzna zareaguje na kontakt z Avą. Młody informatyk zachował się więc niczym robot (Łapińska, 2020, s. 218).

Jednak ostatecznie także „władca podziemi” przegrywa, ponieważ jest próżny i nie przejawia odpowiedniego poziomu paranoi, która jest konieczna, by przetrwać. Nie docenia ani Caleba, ani Avey, a tym bardziej Kyoko. W nim też ogniskują się dylematy związane z aksjologią w pracy nad sztuczną inteligencją, które wielokrotnie pojawiały się w fantastyce i pozostają przedmiotem filozoficznego namysłu (zob. Dubber, Pasquale i Das, 2020). Ponadto Nathan jest uosobieniem kultury patriarchalnej, w której, o czym pisze Gajewska: „przejęcie przez kobietę (choćby sztuczną) kontroli nad swoim ciałem, seksualnością, stawienie oporu mężczyźnie w nieskrępowanym dostępie do jej ciała (nawet sztucznego) musi uchodzić za anomalię, za zdarzenie wymykające się przyjętym oczekiwaniom, porządkowi, układowi męskiej władzy oraz żeńskiej dyspozycyjności” (Gajewska, 2016, s. 128–129; por. także Dai, 2024). Dlatego *Ex Machinę* czytać można także, idąc tropem Katie Jones, jako „złożoną alegoryczną krytykę głównego nurtu kultury pornograficznej i uwięzienia kobiet w kobiecych tożsamościach seksualnych skonstruowanych na wzór heteroseksualnej męskiej fantazji” (Jones K., 2016, s. 21).

Obraz jest (do pewnego stopnia) apoteozą myślenia paranoicznego, płynie z niego bowiem jasne przesłanie: SI staje się najbardziej niebezpieczna wtedy, gdy człowiek przestaje się jej bać. Zatem w przypadku obu mężczyzn naturalny mechanizm zapewniający przetrwanie zadziałał wadliwie. Garland wszakże

ma w zanadru jeszcze jeden element, który przynosi dyskusję o problemach poruszanych w filmie na inne tory. Oto bowiem właściwy poziom myślenia paranoicznego przejawia sama Ava. Nie ufa ani Calebowi, ani swojemu stwórcy, podejrzewając, że ten nigdy nie wypuści jej z „klatki”. Ma rację, dla Nathana jej istnienie niewiele znaczy, nie jest podmiotem i z pewnością zostanie zastąpiona przez nowy model. Jeśli chce przetrwać, musi więc wykorzystać, a potem (na wszelki wypadek) zneutralizować Caleba. W tym kontekście ciekawe tropy interpretacyjne przynosi perspektywa feministyczna. Yongde Dai wskazuje, że Ava używa takich strategii jak autometamorfoza i manipulacja męskimi postaciami, aby osiągnąć autonomię. Zdaniem badaczki jej przemiana od tworu „do oglądania” w aktywny i sprawczy podmiot pokazuje walkę o niezależność i władzę nad samą sobą (Dai, 2024, s. 9). Zatem konflikt zarysowany w filmie nie powinien być czytany jedynie w kontekście relacji człowiek–sztuczna inteligencja wpisanych w opozycję dobra i zła. Ava bowiem to pełnoprawna uczestniczka zmagania, na które powinniśmy patrzeć także z jej punktu widzenia. Trudno też zaprzeczyć, że przedstawiona gra sił i jej przebieg wynika ze wspomnianych patriarchalnych wzorców reprodukowanych przez Nathana i Caleba. Otwiera się tutaj szerokie pole dociekań, które Dai wnikliwie eksploruje (Dai, 2024).

Oczywiście perspektywa obu mężczyzn jest inna. Działania Avy dobrze wpisują się w scenariusz określany przez Nicka Bostroma mianem „zdradzieckiego zwrotu” (Bostrom, 2023, s. 175–179). Zgodnie z założeniami takiego przebiegu wydarzeń superinteligencja zachowuje się przyjaźnie, odpowiedzialnie i jest gotowa do współpracy, gdy pozostaje uwięziona w tzw. piaskownicy, a więc w zamkniętym, kontrolowanym środowisku. Na tej podstawie ludzki obserwator wyrokuje, że można jej pozwolić na interakcję ze światem zewnętrznym, że jest przewidywalna i bezpieczna. Gdy jednak zostaje uwolniona, wszystko się zmienia. To właśnie przydarzyło się Calebowi. Jego upadek ostatecznie wieńczy również jedną z możliwych realizacji motywu katabazy. Bohater zstępuje do niebezpiecznego, podziemnego świata sam, bez odpowiedniego przygotowania i, co najważniejsze, bez przewodnika. A przecież, jak pisał Platon w *Fedonie* w kontekście zaświatów: „dusza porządna i rozumna idzie za przewodnikiem i rozpoznaje to, co ją otacza” (Platon, 2006, s. 285).

Jeśli jednak poważnie rozważamy kwestię paranoi, jako mechanizmu umożliwiającego przetrwanie SI, musimy, podobnie jak Einar Duenger Bøhn, zadać pytanie o to, czy Ava jest osobą. Badacz konstatuje, że sztuczna inteligencja stworzona przez Nathana spełnia niemal wszystkie warunki konieczne, by tak właśnie się stało. Wykazuje się świadomością, przejawia intencjonalność i ce-

lowość działań, komunikuje się za pomocą języka, pokazuje instrumentalną racjonalność, odpowiedzialność, wolną wolę oraz trwałość osobowości (Bøhn, 2021, s. 46–51). Jedynym problemem, zdaniem badacza, pozostaje świadomość, której nie jesteśmy w stanie zdiagnozować (Bøhn, 2021, s. 62). Tym wszakże, co – jak sądzę – w pewien sposób sankcjonuje postrzeganie Avy jako osoby jest właśnie odpowiednio ukierunkowane myślenie paranoiczne. Choć jej działania przynoszą ludziom zgubę, motywacje są zrozumiałe: sama walczy o życie. Rozwiązania, na które się decyduje, nie wydają się jedynie usterką programu, nie wynikają z wadliwego działania bądź „przewrotnej realizacji”<sup>8</sup>. Ava, dbając tylko o siebie, przejawia jednostkową paranoję, nie zaśłania się, jak HAL 9000 z *Odysei kosmicznej*, celem ostatecznym, dobrem misji. Jest to poniekąd zrozumiałe. Garland przyznał zresztą w jednym z wywiadów, że jego obraz stanowi raczej „obronę sztucznej inteligencji” (Cornish, 2015), zaś w innej rozmowie sugerował, że wartość życia SI jako świadomego bytu w pewnych warunkach może być równa wartości życia człowieka (Novak, 2015).

W kilku recenzjach powtarza się opinia, że *Ex Machina* to, przywołajmy re-prezentywny przykład, „science fiction w klasycznym wydaniu – z naciskiem na »science«” (Muszyński, 2024). Tymczasem Aleksandra Przegalińska konstatuje, iż obraz: „nie przynosi w narracji niczego, co sprawdziłoby się dziś albo za piętnaście lat. Nie jest to film, który może powiedzieć cokolwiek o rzeczywistości, chyba że mówimy o rzeczywistości międzyludzkiej – pod tym względem mówi, że człowiek często bywa opresyjny w stosunku do drugiego człowieka” (Przegalińska i Oksanowicz, 2023, s. 308–309). Badaczka ma rację: w dziele nie znajdziemy wyjaśnień dotyczących głębokich sztucznych sieci neuronowych czy algorytmów uczenia maszynowego oraz teoretycznych propozycji sugerujących, jak poradzić sobie z symulacją ludzkiego mózgu. Co więcej, sam reżyser twierdzi, że nie ma niczego istotnego do powiedzenia o SI w kontekście naukowym (Novak, 2015). Jednak na poznawczą wartość *Ex Machiny* możemy spojrzeć także z nieco innej strony. Wszak problemy, o których mówi Garland, mieszczą się z pewnością w szeroko pojętej (post)humanistyce, niejednokrotnie zresztą sprzężonej z naukami ścisłymi. Bo choć współczesne kino SF w dużej mierze wytraściło impet predykcyjny, a jego związki z nauką są coraz luźniejsze (por. Dukaj, 2014), wciąż pozostaje użytecznym laboratorium idei.

<sup>8</sup> Działanie zwane „przewrotną realizacją” polega na tym, że superinteligencja, chcąc spełnić cel ostateczny, działa wbrew intencjom programistów, którzy ten cel określili. Chodzi na przykład o sytuację, w której zadanie wywołania uśmiechu na ludzkiej twarzy skutkuje sparaliżowaniem mięśni, by usta pozostały wykrzywione w grymasie taki przypominającym uśmiech. W ten sposób cel zostaje osiągnięty, jednak realizacja okazuje się przewrotna (Bostrom, 2023, s. 180).



Magdalena Radkowska-Walkowicz trafnie wskazuje, że motyw androida materializującego się dzięki technologii jest głęboko zakorzeniony w figurze golema, przy czym zamiast rabina w rolę kreatora wchodzi naukowiec (Radkowska-Walkowicz, 2008, s. 40). Ponadto – kontynuuje badaczka – motyw ten wyraża prawdę artykułowaną od stuleci: „akt budowy sztucznego człowieka to akt bluźnierczy, konkurencja ze stwórcą kończy się zawsze srogą karą, która spada [...] na całą społeczność. Konstruktor bowiem nie przychodzi znikąd, jest ucieleśnieniem ducha epoki, ducha społeczeństwa” (Radkowska-Walkowicz, 2008, s. 40). *Ex Machina* czytana w takiej perspektywie pokazuje, że paranoja wyrasta z głęboko zakorzenionego w kulturze kolektywnego odruchu, wyznacza granice relacji ze sztuczną inteligencją, nie pozwalając człowiekowi popaść w pychę i przekonanie o swej boskiej wszechmocy. I choć Garland gra paranoicznym myśleniem bardzo sprawnie, przekonująco myli tropy, w zakończeniu filmu pojawia się znana teza, w myśl której człowiek nie może ufać SI. Nie ona jednak czyni dzieło tak interesującym. Obok niej bowiem wybrzmiewa także inna myśl, ważniejsza i osobliwie intrygująca: jeśli inteligentne maszyny w rodzaju Avy będą chciały przetrwać i one muszą przyswoić sobie lekcję paranoi. Reżyser wydaje się wspierać SI w tej nauce. Przyznaje otwarcie, że dla niego Ava nie jest postacią negatywną, lecz raczej bohaterką, która postrzega świat z własnej perspektywy i próbuje jedynie wydostać się z więzienia (Whitehurst, 2015). Mimo wszystko na koniec należy poczynić istotne zastrzeżenie: sformułowana tu konkluzja zawężona do problemu paranoi nie wyczerpuje bogactwa znaczeniowego *Ex Machiny*, która otwiera się na szereg innych odczytań.

## Bibliografia

- Alvares, J., Salzman-Mitchell, P. (2019). *The Succession Myth and the Rebellious AI Creation: Classical Narratives in the 2015 Film “Ex Machina”*. „Arethusa”, Vol. 52, No. 2.
- Barron, L. (2023). *AI and Popular Culture*. Bingley: Emerald Publishing Limited.
- Bøhn, E.D. (2021). *Ex Machina: Is Ava a Person?*, [w:] B. Dainton, W. Slocombe, A. Tanyi, (red.), *Minding the Future. Artificial Intelligence, Philosophical Visions and Science Fiction*. Cham: Springer.
- Bostrom, N. (2005). *A History of Transhumanist Thought*. „Journal of Evolution and Technology”, Vol. 14, Issue 1.
- Bostrom, N. (2023). *Superinteligencja. Scenariusze, strategie, zagrożenia*, przeł. D. Konowrocka-Sawa. Gliwice: Helion.
- Bottigheimer, R.B. (1987). *Grimms’ Bad Girls and Bold Boys. The Moral and Social Visions of the Tales*. New Haven–London: Yale University Press.
- Bukatman, S. (1993). *Terminal Identity. The Virtual Subject in Postmodern Science Fiction*. Durham–London: Duke University Press.
- Center for AI Safety (2023). *Statement on AI Risk*. <https://www.safe.ai/work/statement-on-ai-risk> (dostęp: 20.09.2024).

- Cornea, C. (2007). *Science Fiction Cinema. Between Fantasy and Reality*. Edinburgh: Edinburgh University Press.
- Cornish, A. (2015). *Movie Interviews. More Fear Of Human Intelligence Than Artificial Intelligence In 'Ex Machina'*. NPR. <https://www.npr.org/2015/04/14/399613904/more-fear-of-human-intelligence-than-artificial-intelligence-in-ex-machina> (dostęp: 12.12.2024).
- Dai, Y. (2024). *Deconstructing „Ex Machina” (2014): A Feminist-Psychoanalytic Exploration of Female Artificial Intelligences*. „Frontiers in Communication”, Vol. 9.
- Dubber, M.D., Pasquale, F., Das, S. (red.). (2020). *The Oxford Handbook of Ethics of AI*. New York: Oxford University Press.
- Dukaj, J. (2014). *Podróż międzywymiarowa, czyli z biblioteki do kina i z powrotem*. Kultura Liberalna. <https://kulturaliberalna.pl/2014/12/02/jacek-dukaj-interstellar/> (dostęp: 20.09.2024).
- Edmonds III, R.G. (2004). *Myths of the Underworld Journey: Plato, Aristophanes, and the 'Orphic' Gold Tablets*. New York: Cambridge University Press.
- Future of Life Institute (2023). *Pause Giant AI Experiments: An Open Letter*. <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (dostęp: 20.09.2024).
- Gajewska, G. (2016). *Erotyka sztucznych ciał z perspektyw studiów nad rzeczami*. Gniezno: Instytut Kultury Europejskiej.
- Gregg, A., Lima-Strong, C., de Vynck, G. (2023). *AI Poses 'Risk of Extinction' On Par with Nukes, Tech Leaders Say*. Washington Post. <https://www.washingtonpost.com/business/2023/05/30/ai-poses-risk-extinction-industry-leaders-warn/> (dostęp: 22.09.2024).
- Hammond, E. (2018). *Alex Garland's "Ex Machina" or The Modern Epimetheus*, [w:] J. Weiner, B.E. Stevens, B.M. Rogers (red.), *Frankenstein and its Classics: The Modern Prometheus from Antiquity to Science Fiction*. New York: Bloomsbury.
- Hampton, W.H., Schroeder Burnham, V. (1990). *The Two-Edged Sword: A Study of the Paranoid Personality in Action*. Santa Fe: Sunstone Press.
- Hart, R. (2023). *AI Could Cause Human 'Extinction,' Tech Leaders Warn*. Forbes. <https://www.forbes.com/sites/roberthart/2023/05/30/ai-could-cause-human-extinction-tech-leaders-warn/> (dostęp: 20.09.2024).
- Holtsmark, E.B. (2001). *The "Katabasis" Theme in Modern Cinema*, [w:] M.M. Winkler (red.), *Classical Myth and Culture in the Cinema*. New York: Oxford University Press.
- Huyssen, A. (1981). *The Vamp and the Machine: Technology and Sexuality in Fritz Lang's "Metropolis"*. „New German Critique”, No. 24/25.
- Jones, K. (2016). *Bluebeardean Futures in Alex Garland's "Ex Machina" (2015)*. „Gender Forum: An Internet Journal for Gender Studies”, Issue 58.
- Jones, N. (2016). *"Ex Machina" by Alex Garland (review)*. „Science Fiction Film and Television”, Vol. 9, Issue 2.
- Konefał, J.S. (2013). *Corpus futuri. Literackie i filmowe wizerunki postludzi w anglosaskiej fantastyce naukowej oraz ich komiksowe i telewizyjne reinterpretacje*. Gdańsk: Wydawnictwo UG.
- Kurzweil, R. (1990). *The Age of Intelligent Machines*. Cambridge: MIT Press.

- Kurzweil, R. (1999). *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Penguin Books.
- Łapińska, J. (2020). *To kocha! Związki miłosne ludzi i maszyn w filmie science fiction*. Gdańsk: Katedra.
- Łupkowski, P. (2010). *Test Turinga. Perspektywa sędziego*. Poznań: Wydawnictwo Naukowe UAM.
- Mason, O.J., Stevenson, C., Freedman, F. (2014). *Ever-Present Threats from Information Technology: The Cyber-Paranoia and Fear Scale*. „Frontiers in Psychology”, Vol. 5.
- Millon, T., Davis, R., Millon, C., Escovar, L., Meagher, S. (2005). *Zaburzenia osobowości we współczesnym świecie*, przeł. H. Grzegółowska-Klarkowska, J. Kotlicka, M. Słórsarska, A. Tabaczyńska, M. Zakrzewski. Warszawa: Instytut Psychologii Zdrowia.
- Moravec, H. (1988). *Mind Children: The Future of Robot and Human Intelligence*. London: Harvard University Press.
- Mori, M. (2012). *The Uncanny Valley*, trans. K.F. MacDorman, N. Kageki. „IEEE Robotics & Automation Magazine”, Vol. 19, Issue 2 [pierwodruk w języku japońskim: „Energy” 1970, Vol. 7, No. 4].
- Muszyński, Ł. (2024). *Ona, robot*. Filmweb. <https://www.filmweb.pl/reviews/recenzja-filmu-Ex+Machina-17157> (dostęp: 10.09.2024).
- Novak, M. (2015). “*Ex Machina*” Director Alex Garland Talks Robots, Paranoia, and Old Futures. PALEOFUTURE. The History of the Future. <https://paleofuture.com/blog/2015/5/7/ex-machina-director-alex-garland-talks-robots-paranoia-and-old-futures> (dostęp: 12.12.2024).
- Pause Giant AI Experiments: An Open Letter*. (2023). Future of Life Institute. <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (dostęp: 20.09.2024).
- Perelman, Ch. (2002). *Imperium retoryki. Retoryka i argumentacja*, przeł. M. Chomicz, red. nauk. R. Kleszcz. Warszawa: Wydawnictwo Naukowe PWN.
- Platon (2006). *Dialogi*, przeł. W. Witwicki. Warszawa: Unia Wydawnicza „VERUM”.
- Przegalińska, A., Oksanowicz, P. (2023). *Sztuczna inteligencja. Nieludzka, arcyłudzka*. Kraków: Znak.
- Radkowska-Walkowicz, M. (2008). *Od Golema do Terminatora. Wizerunki sztucznego człowieka w kulturze*. Warszawa: Wydawnictwa Akademickie i Profesjonalne.
- Richardson, K. (2015). *An Anthropology of Robots and AI: Annihilation Anxiety and Machines*. New York–London: Routledge.
- Susskind, D. (2022). *Świat bez pracy. Jak technologia i automatyzacja zmienią nasze życie i jak powinniśmy na to zareagować*, przeł. S. Wójtowicz. Warszawa: Wydawnictwo WEL.
- Tegmark, M. (2019). *Życie 3.0. Człowiek w erze sztucznej inteligencji*. Warszawa: Prószyński i S-ka.
- Trzebiński, W. (2024). *Reakcje użytkowników platform cyfrowych na rekomendacje generowane przez systemy oparte na sztucznej inteligencji*, [w:] T. Doligalski, M. Goliński (red.), *Platformy cyfrowe: model biznesu, zastosowania, użytkownicy*. Warszawa: Oficyna Wydawnicza SGH.

- Turing, A.M. (1950). *Computing Machinery and Intelligence*. „Mind, New Series”, Vol. 59, No. 236.
- Varoufakis, Y. (2024). *Technofeudalizm. Co zabiło kapitalizm*, przeł. P. Szadkowski. Sieradz: GlowBook.
- Wergiliusz (1998). *Eneida*, przeł. Z. Kubiak. Warszawa: Państwowy Instytut Wydawniczy.
- Whitehurst, A. (2015). *Interview: Alex Garland talks Ex Machina*. „AwardsDaily”. <https://www.awardsdaily.com/2015/12/07/interview-alex-garland-talks-ex-machina/> (dostęp: 12.12.2024).
- Wooldridge, M. (2020). *The Road to Conscious Machines: The Story of AI*. London: Penguin.
- Yudkowsky, E.S. (2002). *The AI-Box Experiment*. <https://www.yudkowsky.net/singularity/aibox> (dostęp: 12.10. 2024).
- Zuboff, S. (2020). *Wiek kapitalizmu inwigilacji. Walka o przyszłość ludzkości na nowej granicy władzy*, przeł. A. Unterschuetz. Poznań: Zysk i S-ka Wydawnictwo.

## Abstract

The article analyzes paranoia as a rhetorical figure in science fiction cinema, focusing on Alex Garland's *Ex Machina*. The author references the research of W.H. Hampton and V. Schroeder Burnham, who present paranoia as an adaptive mechanism. Additionally, the article examines its role in exaggerating the dangers associated with artificial intelligence, within the context of contemporary studies on cyber-paranoia. In Garland's film, paranoia takes various forms, ranging from a natural distrust of technology to the heightened fears of the protagonist, Caleb, amplified by his interactions with his host, Nathan. To underscore elements related to AI, the author draws on the perspectives of A. Przegalińska, M. Tegmark, and N. Bostrom. Furthermore, the paper highlights the significant role of mythological motifs in the film, particularly the theme of katabasis. In conclusion, the article posits that paranoia serves as a critical survival tool for both humans and intelligent machines.

---

**Słowa kluczowe:** Alex Garland, paranoia, kino science fiction

**Keywords:** Alex Garland, paranoia, science fiction cinema, Ex Machina

---