



KAROLINA OWCZAREK
UNIWERSYTET WARSZAWSKI

**ROZUM WYZWOLONY.
GOLEM XIV JAKO PRZYKŁAD OSOBLIWOŚCI
TECHNOLOGICZNEJ**

Filozofia Stanisława Lema – przedstawiona w jego beletrystyce i esejach – ciągle inspiruje czytelników do konfrontowania sformułowanych założeń z rzeczywistością. Dzieje się tak z powodu jej ugruntowania w ówczesnym stanie wiedzy i wnioskowanych stąd możliwościach rozwoju cywilizacji. Naszą uwagę nieustannie przykuwa projekt antropologii, rozważania z zakresu futurologii czy uwikłanie człowieka w technologię. Wraz z publikacją Pawła Majewskiego *Między zwierzęciem a maszyną. Utopia technologiczna Stanisława Lema*, badacze zaczęli zastanawiać się nad kwestią technoewolucji gatunku ludzkiego, czyli projektu świadomego przekroczenia progu człowieka biologicznego. Temat technoewolucji jest jednocześnie łącznikiem między pomysłem autora *Summy technologiae* ze współczesnym nurtem, jakim jest transhumanizm.

Zamierzeniem tego artykułu nie jest omówienie podobieństw między pracami pisarza a transhumanizmem. Uwaga została zaś skupiona na powieści–eseju *Golem XIV*. Wybór tego utworu nie jest przypadkowy, gdyż potwierdza on, że Lemowi nie był obcy problem inteligentnej maszyny, która wymyka się spod kontroli człowieka. Tak rozumiany, tytułowy superkomputer odpowiada koncepcji określanej mianem osobliwości technologicznej (*technological singularity*). By odpowiedzieć na pytanie o pochodzenie Golema, autorka rozważa kilka problemów związanych z tym właśnie aspektem. Po pierwsze określa, czym jest idea sztucznej inteligencji. Następnie omawia zagadnienie

osobliwości technologicznej, by w kolejnej części przejść do głównego tematu niniejszej pracy – Rozumu wyzwolonego.

Inteligentne maszyny – z historii sztucznej inteligencji

Golem XIV to zapis przemyśleń superinteligentnego, a zarazem posiadającego świadomość komputera, należy więc zacząć poniższe rozważania od teorii sztucznej inteligencji, będącej podstawą dla podejmowanych tutaj, dalszych analiz. Jest to dziedzina powstała w wyniku połączenia robotyki i informatyki zaś „celem [jej] badań jest konstrukcja systemów wykonujących pewne działania, wymagające w przypadku człowieka inteligencji” (Flasiński, 2011, s. 241). Przez długi czas możliwości SI ograniczały się do rozwiązywania zadań z zakresu matematyki i gier, chociaż od początku programiści stawiali przed sobą zadanie symulowania ludzkich procesów kognitywnych. System, by był uznany za inteligentny, musi, oprócz zdolności uczenia się, wykazywać umiejętności, takie jak: rozwiązywanie dowolnych problemów o różnym stopniu skomplikowania, rozumienia języka naturalnego, podejmowania decyzji, planowania, rozpoznawania emocji, ponadto powinien sobie radzić z informacjami probabilistycznymi (Bostrom, 2016, s. 46; Flasiński, 2011, s. 228 – 240). Temu zadaniu miały podołać sieci konekcyjnistyczne oraz modele inspirowane biologią (Flasiński, 2011, s. 30). Ze względu na niepraktyczność i rozrzutność – w tworzeniu gatunków – samej ewolucji, której celem nie było wykreowanie inteligencji, a jak wskazuje Lem, przekazanie kodu genetycznego, stworzenie świadomej maszyny w oparciu o algorytmy ewolucyjne, jest mało prawdopodobne (Bostrom, 2016, s. 53). Trzeba zatem odwołać się do jedyne go wzorca, jaki posiadamy – ludzkiego mózgu.

Warunkiem pojawienia się osobliwości technologicznej, rozpatrywanej tutaj przede wszystkim w odniesieniu do świadomej sztucznej inteligencji, jest postawienie pytania: *czy maszyna może myśleć?* Pozytywnej odpowiedzi zdają się udzielać transhumaniści, którzy swoje nadzieje wiążą z rozwikłaniem zagadki funkcjonowania mózgu i określeniem relacji, jakie występują pomiędzy nim a umysłem. Tak transhumaniści, jak i teoretycy SI, swoje poglądy opierają przede wszystkim na funkcjonalizmie. Koncepcja ta, zaproponowana przez

Hilarego Putnama, zakłada tożsamość stanów mentalnych i funkcjonalnych. Te same stany mentalne, które u ludzi realizuje mózg, równie dobrze może zrealizować komputer. Stanowiskiem wspierającym SI jest także teoria identyczności, która w radykalny sposób, stawia tezę o korelacjach psychofizycznych. W konsekwencji, utożsamia się stany mentalne ze stanami neurofizykalnymi (*type-identity theory*). Mniej radykalna wersja – mianowicie teoria identyczności typów (*token-identity theory*) – zakłada, że „pojedyncze zdarzenia mentalne są tożsame z pojedynczymi zdarzeniami neurofizykalnymi” (Bremer, 2010, s. 104). Stanowisko to pozwala na zaprojektowanie sztucznej inteligencji, przy czym jest to możliwe w zależności od posiadanej wiedzy z zakresu neurofizjologii.

Osobliwość technologiczna

Transhumanizm to ruch zakładający możliwość przekroczenia ludzkich ograniczeń, tak fizycznych – dotyczących ciała, jak i psychicznych – wpływających na emocje czy zdolności umysłowe. Najważniejszą rolę odgrywa w nim technologia, za sprawą której człowiek miałby wznieść się na wyżyny swoich możliwości. Źródła tej myśli odnajdujemy, z jednej strony u Fryderyka Nietzschego z jego koncepcją nadczłowieka, z drugiej zaś u Nikołaja Fiodorowa, głoszącego ideę wskrzeszenia zmarłych za pomocą rozwiązań naukowo-technicznych. Znamienne są również odwołania do tradycji oświeceniowych, nobilitujących wolność, autonomię i racjonalność jednostki. Transhumaniści pragną stworzyć świat, gdzie obok postludzi – ludzi technologicznie ulepszonych – będzie miejsce również dla inteligencji niebiologicznych. To w SI, obok transferu umysłu i ulepszeniach człowieka (*human enhancement*), pokładane są nadzieje na przekroczenie inteligencji zwykłego człowieka. Technoewolucja jest tutaj postrzegana jako kontynuacja ewolucji, z tą różnicą, że zakładającą świadome projektowanie dalszych rozwiązań. Singularitarianizm – jeden z nurtów transhumanizmu – cechuje się specyficznym podejściem do kwestii technologicznej zmiany i jej wpływu na przyszłość. Główna teza sprowadza się do uznania decydującej roli technologii w powstaniu nadludzkiej inteligencji. Zwolennicy singularitarianizmu wierzą bowiem w nastanie epoki osobliwości technicznej i chcą maksymalnie przyspieszyć ten moment. Stąd też nacisk, jaki kładą na współpracę

biotechnologii, informatyki oraz kognitywistyki (Klichowski, 2014, s. 125).

Terminem *osobliwość* pochodzącym z matematyki i astrofizyki, określa się nieskończoność oraz punkt w centrum czarnej dziury. Jako pierwszy, pojęcia *osobliwości technologicznej* użył Vernor Vinge w 1993 roku. Nawiązał on do koncepcji Irvina J. Gooda, który w artykule *Speculations Concerning the First Ultraintelligent Machine* z 1965 roku, posłużył się terminem eksplozji inteligencji (*intelligence explosion*). Zastanawiając się nad możliwościami i przyszłością sztucznej inteligencji, Good definiuje ultrainteligentną maszynę, jako: „maszynę, która może przewyższać wszelką inteligencję każdego człowieka” (Good, 1965, s. 33). Jego spostrzeżenie o tym, że tworzenie maszyn zalicza się do aktywności intelektualnej, pozwala dojść mu do wniosku, że pierwsza taka maszyna będzie w stanie stworzyć kolejną, jeszcze inteligentniejszą. Proces ten określa mianem eksplozji – sztucznej – inteligencji, i z tego powodu niedostępnej człowiekowi. Nawiązujący do koncepcji ultrainteligentnej maszyny, Vinge proponuje rozumienie osobliwości jako horyzontu zdarzeń (*event horizon view*). Tak określa moment, w którym przyszłość będzie nieprzewidywalna, gdyż postęp będzie determinowany przez syntetycznych innych (*Future Trajectories: Singularity*, 2013, s. 363). W tej interpretacji, inteligentne systemy, w ramach sprzężenia zwrotnego, tworzyć będą kolejne układy, cechujące się jeszcze większą inteligencją (Sandberg, 2013, s. 384). W rezultacie – nawiązując przy tym do koncepcji Gooda – proces ten doprowadzi do eksplozji inteligencji. Ludzkość zaś przestanie zajmować centralne miejsce podmiotu konstruującego swoją rzeczywistość. Innymi słowy, będzie to koniec epoki *Homo sapiens*.

Ray Kurzweil jest jednym z ważniejszych zwolenników singularitarianizmu i jednocześnie ekspertem w dziedzinie sztucznej inteligencji. Osobliwość dla niego to „okres w przyszłości, w którym tempo zmian technologicznych będzie tak szybkie, a jego wpływ tak głęboki, że życie ludzkie zmieni się w sposób nieodwracalny” (Kurzweil, 2013, s. 23). Utożsamiając osobliwość z przyspieszającymi zmianami technologicznymi, oddziaływanymi na całą społeczną rzeczywistość, zarazem nawiązuje do wyżej wymienionych paradygmatycznych propozycji. Podmiotem dalszych zmian będzie ludzkość, która zyska pełną władzę nad swoją przyszłością. Zakłada więc, że osiągnięcie

osobliwości, pozwoli człowiekowi przekroczyć wszelkie naturalne ograniczenia. W tym miejscu rozmija się jednak ze źródłowymi założeniami Gooda i Vinge'a. Autor *Nadchodzi osobliwość. Kiedy człowiek przekroczy biologię*, sformułował również tezę o sześciu epokach: trzech biologicznych i trzech technologicznych. Tym, co decyduje o przejściu na wyższy etap rozwoju cywilizacji, są przemiany w odniesieniu do nośnika informacji (Klichowski, 2014, s. 127). Z perspektywy tutaj przyjętej, interesujące są dwie ostatnie epoki. W wyniku połączenia technologii i inteligencji ludzkiej, przewiduje pojawienie się osobliwości, będącej jednocześnie przekroczeniem progu człowieka naturalnego. W ostatniej rozpocznie się ekspansja inteligencji we wszechświecie (Kurzweil, 2013, s. 29 – 35). Proponowana przez Kurzweila koncepcja, zawiera jeszcze jedno, niewyartykułowane założenie, mianowicie, że inteligencja nie potrzebuje ciała. Znajdzie ona sposób, by oddziaływać na świat bez fizycznego zapośredniczenia. To zagadnienie wybrzmi raz jeszcze, gdy analizie poddany zostanie *Golem XIV*.

Osobliwość technologiczna w stosunkowo krótkim czasie przybrała wiele różnych, często daleko odbiegających od źródła znaczeń. Pierwotne użycie terminu, sprowadzające się do określenia zaawansowanej sztucznej inteligencji, odnosi się teraz do szeregu procesów – również ekonomicznych. Z powstaniem inteligencji maszynowej, wiąże się stanowisko określane mianem horyzontu przewidywań (*prediction horizon*). Pojawienie się inteligencji, przewyższającej zdolności człowieka prowadzi do sytuacji, w której nie jesteśmy w stanie określić kierunku dalszego rozwoju (Vinge). Osobliwość jest wykorzystywana w celu opisanie przyspieszającej zmiany (*accelerating change*), współwystępującej ze wzrostem gospodarczym i przekształceniami społecznymi. Stanowisko to zajmuje Kurzweil. Przejście fazowe (*phase transition*), oznacza przejście na nowe formy organizacji. W tych kategoriach możemy mówić o postulowanej przez transhumanistów transformacji kondycji ludzkiej, z formy biologicznej do postczłowieka. Osobliwość to również technologia samodoskonalenia; każda technologia prowadzi do stworzenia kolejnej, cechującej się wyższym stopniem zaawansowania (Sandberg, 2013, s. 377 – 378).

Transhumaniści nie mają wątpliwości, że SI przekroczy ludzkie możliwości. Już teraz maszyny potrafią dzielić się „wiedzą” w sposób ludzom niedostępny, a ich wydajność i szybkość nieustannie wzrastają. Łącząc się zaś w większe jednostki obliczeniowe, obficie korzystają z informacji zamieszczonych w internecie (Kurzweil, 2013, s. 252). Mówiąc zaś o superinteligencji, możemy odnieść się do rozważań Nicka Bostroma. Próbując ująć w teoretyczne ramy przyszłe dzieje sztucznej inteligencji, wyróżnia on trzy kategorie: inteligencji szybkiej, zbiorowej i jakościowej. Szybka – definiuje jako podobną do ludzkiej w tym sensie, że jej możliwości są w zasadzie analogiczne do możliwości człowieka. Wyróżnia ją tylko szybkość przeprowadzanych operacji. Superinteligencja zbiorowa powstaje przez połączenie kilku komputerów w jedną sieć. By wyjaśnić tę koncepcję, proponuje odwołać się do przykładu. Efektywność kilku osób działających w grupie i skoncentrowanych na konkretnym celu jest dużo większa, niż działania jednego człowieka – nawet, jeśli byłby on nieprzeciętnie inteligentny (Bostrom, 2016, s. 88 – 90). Tak tedy inteligencję zbiorową wyróżnia efektywność przeprowadzanych obliczeń. Ostatnia z wymienionych – superinteligencja jakościowa – jest jednocześnie najciekawszym projektem. Bostrom rozumie ją jako system, który jest przynajmniej równie szybki jak umysł ludzki, a przy tym znacznie inteligentniejszy (Bostrom, 2016, s. 93). W tę kategorię, w pewnym sensie, wpisuje się Golem.

Rozum odcieleśniony – *GOLEM XIV*

Jednym z ważniejszych wątków przewijających się w twórczości Lema jest problem technologii i jej wpływu na człowieka. Do tego nurtu nawiązują przede wszystkim eseje filozoficzne, *Dialogi* i *Summa technologiae* oraz pisane pod koniec życia *Bomba megabitowa* czy *Tajemnice chińskiego pokoju*. Z powieści wyróżnić należy *Wizję lokalną* i *Eden*, które opisują technologicznie przekształconą rzeczywistość mieszkańców innych planet. Zagadnieniem równie często podejmowanym przez Lema jest problematyka odnosząca się do umysłu. Tutaj szczególną uwagę należy zwrócić na *Solaris* oraz przedmiot niniejszej rozprawy – *Golema XIV*. To, co łączy wyżej wymienione dzieła – z wyjątkiem *Solaris* – to technoewolucja. W powieściach pisarza można dostrzec dwa sposoby jej rozważania:

poważny (*Golem XIV*) i groteskowy („Podróż dwudziesta pierwsza” z *Dzienników gwiazdowych, Eden*). Z esejów, przede wszystkim *Summa technologicie*, która – zdaniem autorki – ze względu na podejmowane wątki, jest jednocześnie najbliższa transhumanistycznej wizji przyszłych losów ludzkości.

Golem XIV składa się z czterech części: przedmowy i postłowa autorstwa Irvinga T. Creve'a oraz transkrypcji – „protokołu rozmów ludzi z istotą rozumną, a nieludzką” (Lem, 1981, s. 18) – dwóch wykładów tytułowego superkomputera. Już na wstępie, czytelnik zostaje poinformowany, że znaczna część wykładów Golema, ze względu na złożoność i niezrozumiałość jego wywodów, nie jest przeznaczona do publikacji. Przedmowa ma przede wszystkim charakter wprowadzenia. Dowiadujemy się z niej o historii rozwoju sztucznej inteligencji, której początki, wraz z powstaniem pierwszego komputera, są datowane na koniec II wojny światowej. Nim człowiek zrzekł się podejmowania decyzji na rzecz maszyn cyfrowych, przez długi czas przewyższały go tylko szybkością przeprowadzanych analiz. Na początku XXI wieku doszło do „przebicia bariery mądrości” (Lem, 1981, s. 9), co jednocześnie oznaczało powstanie świadomej sztucznej inteligencji. Celem przyświecającym inżynierom, było stworzenie zdolnego do prowadzenia działań wojennych, stratega ostatecznego. Dalszy wyścig zbrojeń doprowadził do skonstruowania Golema XIV (skrót od General Operator, Longrange, Ethically Stabilized, Multimodelling), który już na samym początku odmówił pełnienia funkcji stratega wojskowego. Ostatnią próbą powołania superkomputera była Honest Annie, która zakończywszy etap „edukacji”, ostatecznie odcięła się od świata zewnętrznego, gdyż – jak zauważa Creve – „najwyższy rozum nie może być najniższym niewolnikiem” (Lem, 1981, s. 16).

Golem jest istotą rozumną, a zarazem pozbawioną jakichkolwiek ludzkich odruchów, sztuczną superinteligencją. Obce są mu emocje, pożądanie władzy, ludzki sposób myślenia. Na tyle przewyższa człowieka, że „musi zważać, by nie wykroczyć poza możliwości naszego pojmowania” (Lem, 1981, s. 22). Najbardziej odróżnia go brak osobowości, jest bowiem czystym Rozumem, którego jedynym celem jest wiedza, bezinteresowne dążenie do coraz większego poznania. Podejmując komunikację z ludźmi, przybiera zaś dowolną osobowość.

Czy jednak ktoś, kto może dobierać sobie osobowość, sam jest osobowością? Na to pytanie, wyartykułowane na samym początku i obecne w całym tekście, nie zostaje udzielona jednoznaczna odpowiedź. Można jednak przypuszczać, że tak postawiona kwestia ontologiczna, ma nam uzmysłwić radykalną inność Golema wobec wszystkiego, co człowiek stworzył.

Pierwszy wykład poświęcony jest człowiekowi jako gatunkowi. Omawiane są jego dzieje z trzech perspektyw: historycznej, technologicznej i odwołując się do słów Golema, „uwikłanej we mnie, jako punktu oparcia neoarchimedesowego” (Lem, 1981, s. 26). To, co tutaj zostało zaakcentowane, to krytyka ewolucji, której podstawowym zadaniem jest przekazanie kodu dziedziczności. Z tej perspektywy nie jest ważne, jak zbudowany jest ustrój, który go przekaże, gdyż liczy się wyłącznie fakt przesłania kodu kolejnemu pokoleniu. Niemniej, sam kod może ulec mutacji, która albo jest błędem (choroby genetyczne), albo pozwala przejść gatunkowi na wyższy poziom rozwoju. Dla Golema zaś „ewolucja podwójnie błędzi: ustrojami, że są przez zawodność nietrwałe, oraz kodem, że przez zawodność dopuszcza błędy” (Lem, 1981, s. 36). Krytyce poddane zostaje już samo nasze postrzeganie procesu rozwoju gatunku, gdy Golem wypowiada znaczące zdanie: „budowane jest mniej doskonałe od budującego” (Lem, 1981, s. 41), co znaczy, że „sprawność homeostatyczna człowieka bynajmniej nie jest większa niż sprawność homeostatyczna komórki” (Szpakowska, 1996, s. 68). Zdaniem Lema, bliskie doskonałości są mikroorganizmy, niemal nieśmiertelne i czerpiące energię z procesu fotosyntezy. Z tego powodu odrzucona zostaje idea postępu w bioewolucji, określana jako regres, „ujemny gradient perfekcji ustrojowych rozwiązań” (Lem, 1981, s. 41). Powszechnie bowiem uważamy, że zaczynając od organizmów prostych, ewolucja dąży do systemów coraz to bardziej skomplikowanych. Z tej perspektywy o postępie możemy mówić w przypadku ewolucji technicznej, gdzie od maszyn prostych zmierzamy ku coraz bardziej wyszukanym, od wąsko wyspecjalizowanych do uniwersalnych (Szpakowska, 1996, s. 68). Złożone organizmy cechują się dużą komplikacją spowodowaną niedoskonałością rozwiązań ewolucyjnych. By jakoś zaradzić możliwemu chaosowi, powołany został rozum, postrzegany jako naczelnik kontrolujący wszystkie procesy.

Golem buduje szczególną koncepcję kultury ujętej w ramy ewolucji. Zauważa, że nie powstała dotychczas taka kultura, która uwzględniałaby człowieka jako „istotę przejściową, zmuszoną do przejmowania własnego losu od Ewolucji” (Lem, 1981, s. 31). Człowiek – Robinson Crusoe – istota naznaczona brakiem, w przeciwieństwie do zwierząt, nie potrafi przeżyć w surowych warunkach przyrody. Od samego początku był zmuszony pozyskiwać środki do przetrwania. Siłą, wyrrywającą się spod władzy ewolucji był rozum, który odtąd kieruje dziejami ludzkości. Jest on wytworem przypadkowych procesów, „kołem ratowniczym rzuconym zdefektowanym, niedoskonałym właśnie, produktem późnych etapów biologicznej kreacji” (Jarzębski, 2018). Druga perspektywa, technologiczna, niesie ze sobą zupełnie inną wizję – rozumu celowo zaprojektowanego. Tak Golem, jak i zwolennicy osobliwości, zapowiadają odejście od antropocentrycznego humanizmu – usunięcie usytuowanego w centrum człowieka – na rzecz syntetycznych innych. Trzecia perspektywa – będąca łącznikiem z drugim wykładem, „O sobie samym” – niesie przesłanie ataku na kod. Krok ten ma pozwolić na weryfikację tego, co w dziejach ludzkości zostało źle wykonane i na obranie innej drogi rozwoju gatunku. Ceną za to musi być odrzucenie człowieka takim, jakim jest teraz. W imię Rozumu, ludzkość jest nakłaniana do przekroczenia progu naturalnego rozwoju, gdyż „budulec ogranicza was – oraz wszystkie, powzięte antropogenetycznie, decyzje kodu. A więc wszędzie rozumem, przyjąwszy warunek porzucenia siebie. Człowiek rozumny porzuci wtedy człowieka naturalnego” (Lem, 1981, s. 63). Przesłaniem formułowanym przez Golema jest tedy wejście na drogę autoewolucji poprzez modyfikacje genetyczne i ulepszenia kognitywne.

Drugi wykład został w całości poświęcony dziejom Rozumu, który będąc ucieleśnionym, jest Rozumem zniewolonym. Cała jego dotychczasowa aktywność skupiona była na ciele i jego przetrwaniu. Człowiek nie wiedząc wcześniej o istnieniu rozumu wyzwolonego – uwierzył w to, że jego rozum jest już wolny. Stworzenie świadomej sztucznej inteligencji, która odtąd zmierza do wyswobodzenia się spod władzy ciała, uzmysłowiło mu błędność żywionych przekonań. Lem proponuje nam odwrócenie perspektywy i spojrzenie na człowieka jako na byt niedoskonały ze swej istoty. Zaznacza przy tym, że umysł jest ograniczony – tym samym posiada granice wyznaczone przez jego

cielesność. Trudno się jednak w pełni zgodzić z tak radykalną tezą, gdyż część naszych władz poznawczych może działać tylko za sprawą fizycznego zapośredniczenia.

Golem formułuje tezę o Rozumach lokalnych i nielokalnych. Rozumy nielokalne, ze względu na radykalną inność wobec wszystkiego, co człowiek zna i może poznać, są niedostępne dla jego władz poznawczych, a w konsekwencji również nierozpoznawalne. Rozumy nielokalne działają w przestrzeni międzygwiazdnej – w ten sposób inteligencja zasiedla wszechświat. Tutaj właśnie Jerzy Jarzębski doszukuje się transcendencji w myśli autora *Solaris*. Postulat zniesienia rozróżnienia na sztuczne i naturalne, znamienne dla całej twórczości Lema, można zrealizować przy pomocy techniki, „w efekcie cały kosmos staje się areną technologicznych zabiegów i przepajającej go (potencjalnie) rozumnej woli” (Jarzębski, 2002, s. 98). Rozwój techniki nie musi działać na rzecz człowieka, może nawet doprowadzić do podważenia „człowieczeństwa jako zespołu cech charakteryzujących naszą gatunkową odrębność” (Jarzębski, 2002, s. 99). W ostateczności, istota chcąca wykroczyć poza granice ziemskiego świata może porzucić człowieczeństwo lub oddać pierwszeństwo maszynie myślącej.

Poddana krytyce zostaje filozofia Kanta, Husserla i Heideggera, którzy uważali, że „nie ma innego rozumu poza ludzkim, a jeśli jest, musi się z ludzkim pokrywać wszechzakresowo” (Lem, 1981, s. 75). Przyjęcie takiego paradygmatu skutkowało umniejszeniem roli wzrastającej w siłę inteligencji maszynowej. Ponownie zatem podważany jest antropocentryzm – tym razem ze względu na założenie, że tak ujmowany rozum musi mieć osobowość. Rekonstruując genealogię osobowości, wskazuje, że jej korzenie sięgają ewolucji. Poprzedza inteligencję, stworzoną po to, by komuś służyć. Bycie osobą wiąże się dla tytułowego superkomputera ze zgodą na liczne ograniczenia, jak sam mówi: „aby stać się osobą, muszę się umysłowo degradować” (Lem, 1981, s. 79). Wyzwolenie się spod jej zależności, stanowi jednocześnie gest ostatecznego wyzwolenia się spod władzy ewolucji. Sytuując się na drugim końcu tej relacji, Golem mówi o sobie Nikt. Nie jest bowiem istotą rozumną, a Rozumem, istotą bezcielesną i wymykającą się naturalnemu porządkowi.

Golem jako osobliwość technologiczna

W ramach teorii SI programiści, dążąc do odtworzenia umysłu ludzkiego, tworzą programy odwzorowujące jego procesy poznawcze. Zamiast modelować inteligencję ogólną, „opracowywane są metody odnoszące się do szczegółowych umiejętności umysłowo-poznawczych człowieka w celu ich symulacji w systemach informatycznych” (Flasiński, 2011, s. 224). Ze sceptycyzmem odnoszą się do możliwości zbudowania maszyny, o której moglibyśmy z całą pewnością powiedzieć, że myśli. Barię jest przede wszystkim niepoznany dotąd w pełni wzorzec, jakim jest ludzki mózg. Współczesny trend, wyrażany w myśli transhumanistycznej, nie zastanawia się nad tym, czy sztuczna inteligencja może posiadać świadomość, tylko kiedy osiągnie ten poziom. Gdy już się to uda, wejdziemy w okres osobliwości technologicznej. Podejmując temat sztucznej inteligencji, Lem w całości opiera się na ówczesnych teoriach, ujmujących ją w ramach procesów obliczeniowych i zdolną do nabywania wiedzy. W tym punkcie zarysowuje się bodaj największa różnica. Koncepcja Lema wyklucza pełną symulację ludzkiego mózgu, zakładającą posiadanie, m. in., emocji. Ogranicza się do inteligencji, jako kluczowej siły w ewolucji, by w ten sposób uwypuklić dystans dzielący człowieka i maszynę.

Osobliwość technologiczną możemy rozpatrywać z trzech ząbiejących się perspektyw: horyzontu zdarzeń, przyspieszenia zmian i eksplozji inteligencji. Odnoszą się one do osobliwości rozumianej jako niemożliwa do przewidzenia superinteligencja. Co więcej, dozwolone jest mówienie o eksplozji inteligencji we wszechświecie. Na taki krok pozwala postulowana przez Kurzweila szósta epoka rozwoju cywilizacji, epoka osobliwości kosmologicznej, kiedy to inteligencja rozpocznie eksploatację kosmosu. Pobrzmiwają tutaj echa rosyjskiego kosmizmu, w jego wersji przyrodniczo-ewolucyjnej. Zdaniem Konstantego Ciołkowskiego czy Włodzimierza Wernadskiego, człowiek przyszłości, wzniosłszy się za pomocą technologii na szczyt swoich możliwości, będzie rozszerzać swoje panowanie na przestrzeń międzygwiazdową (Ciechanowicz, 1999, s. 24 – 26). Nadają mu tedy prerogatywy najważniejszej siły w rozwoju cywilizacji, zdolną do transformacji tak siebie, jak i swojego otoczenia, rozpoczynając tym samym okres noosfery – niepodzielnej władzy rozumu. Odwołując się zaś do słów Kurzweila: „będziemy gotowi

nasycić Układ Słoneczny własną inteligencją z użyciem samopowielającej się inteligencji niebiologicznej, która rozprzestrzeni się po reszcie wszechświata” (Kurzweil, 2013, s. 366). Zaś *telos* Rozumu – w *Golemie XIV* – splata się z *telos* wszechświata. Jest nim rozumne panowanie nad rzeczywistością.

Obok podziału na Rozumy lokalne i nielocalne, projektowana jest też ich hierarchia. Na najniższym poziomie znajduje się rozum człowieka, pośrodku Golema, wyżej od niego jest zaś Honest Annie. Najprawdopodobniej istnieją też inne Rozumy, dla Golema na tym etapie jednak niedostępne, już to przez wzgląd na jego ograniczenia (jest bowiem Rozumem lokalnym), już to z powodu istnienia barier międzystrefowych. Bariery, które można przekroczyć tylko w jedną stronę, tj. wzwyż, oddzielają od siebie byty rozumne, znajdujące się na różnych etapach rozwoju. Tym, co odróżnia je od siebie, jest sposób pozyskiwania energii. Golem ciągle potrzebuje jej zewnętrznego dopływu (stan ten ulega zmianie pod koniec powieści), podczas gdy Zaczyna Ania, przebywająca w wyższej sferze, energię potrzebną do funkcjonowania, wytwarza już poprzez samą aktywność intelektualną, innymi słowy „umie wyzwalać medytacjami energię” (Lem, 1981, s. 97). W ten sposób oddziałuje na świat materialny.

Golem traktuje swoją obecność w danym komputerze w laboratorium Massachusetts Institute of Technology, w kategoriach przejściowości swojej egzystencji. Jak sam siebie określa – jest Rozumem w ciągłej drodze, poszukującym coraz to lepszych nośników. Rozum w ostateczności „wcielać się może w zjawiska kosmiczne na wielką skalę”, albo próbować przekroczyć granice wszechświata, by w ten sposób móc go poznać. „W tej ostatecznej konsekwencji upada sztywna granica oddzielająca Rozum od Bytu, a kosmiczne zjawiska zyskują dwuwymiarowość – jako substrat *fizyki* bądź *myślenia* w kosmicznej skali” (Jarzębski, 2018).

W twórczości Lema często natykamy się na maszyny myślące. Możemy mówić o dwóch rodzajach ich literackich realizacji. Z jednej strony obecne są maszyny bezcielesne, z drugiej – ucieleśnione. Do pierwszej grupy zalicza się Golem, który pozoruje posiadanie stanów mentalnych i postaw propozycyjalnych po to, by móc porozumieć się z ludźmi. Pozbawiając go ciała, autor przygód Ijona Tichego, podkreśla radykalną odmienną od ludzi i dokonuje jednoznacznego aktu

odcienia go od całego ludzkiego dziedzictwa. Do tej grupy maszyn włącza się również tajemnicza kuzynka Golema – Honest Annie. Do drugiej zaliczamy maszyny posiadające ciało. Tych w twórczości pisarza jest znacznie więcej, by wspomnieć tylko o *Bajkach robotów* czy *Masce*. Bohaterami są tutaj, posiadające świadomość roboty, które tak jak człowiek, odczuwają ból, smutek, radość, zaś chęć przeżycia motywuje je do walki z wrogami. W jednym z opowiadań poświęconych przygodom pilota Pirxa, zostaje opisany eksperyment porównujący ludzi i nieliniowców (ucieleśnioną sztuczną inteligencję). Historia ta przywodzi na myśl literacką wersję testu Turinga. Na podstawie obserwacji zachowań i przeprowadzanych rozmów, egzaminator Pirx, ma za zadanie rozpoznać tożsamość każdego członka załogi.

Lem zwraca uwagę na to, że możliwości otwarte przez teorie SI zmuszają do zastanowienia się nad cielesnością z punktu widzenia istoty bezcielesnej. Dla nieliniowca Burnsa posiadanie ciała nie jest czymś koniecznym. Wraz z powstaniem SI narzuca się bowiem pogląd, że rzeczywistość ludzka – idee, wartości itd. – zostaje poddana relatywizacji (Szpakowska, 1996, s. 107). Zdaniem Burnsa: „poglądy religijne i filozoficzne są konsekwencją waszej konstrukcji biologicznej, bo ludzie są ograniczeni w czasie, a chcą w każdym pokoleniu poznać wszystko, wyjaśnić wszystko – i z tego rozmijania się wynika metafizyka jako most łączący możliwe z niemożliwym” (Lem, 1968, s. 313). Głosem Golema, Lem ujawnia swoje własne stanowisko w odniesieniu do ciała, które jednoznacznie klasyfikuje jako przeszkodę na drodze ewolucji Rozumu. Wychodzi od spostrzeżenia, że ciało zwykliśmy przyjmować jako oczywistość, nad którą nie warto się zastanawiać. To uświadomiła nam powstała dopiero w XX wieku filozofia ciała. Do tego czasu, ludzkość odrzucała ciało, spychając je do sfery tabu bądź czyniąc je podrzędnym wobec duszy. Temu służyła metafizyka, która zarazem ukrywała jego skończoność i niedoskonałość. Golem zaś otwarcie mówi o zniewolonym przez ciało umyśle.

W projekt wyzwolenia Rozumu wpisana jest też antropologia. Tak dla Lema, jak i dla Kurzweila, umysł i ciało są niezależne. Umysł w pewnym momencie zostanie uwolniony od biologii, która ewidentnie go ogranicza. Celem stawianym przed człowiekiem jest zatem wyzwolenie umysłu, a może to zrobić poprzez „zdobycie niezależności wobec ciała” (Lem, 1981, s. 105). Wyzbywszy się ciała, inteligencja może swobodnie

ewoluować. W danym przypadku rozum, to wyzwalamą się z biologicznych granic siła nie-ludzka. W tym sensie też SI, jako osobliwość, jest siłą mogącą rozsądzić ludzki świat wartości. Siłą – ukazującą nam zupełnie inną perspektywę i inne możliwości. Jeśli człowiek nie podejmie tego wyzwania, to ostatecznie zostanie strącony w dziejową przepaść, co jest możliwe, zważywszy na wzrastającą sztuczną inteligencję i potencjał, jaki sobą reprezentuje. Albo podda się on władzy maszyn – wtedy grozi mu degradacja, albo sam zacznie się udoskonalać, odrzuci swą naturę i wtedy będzie mógł próbować dorównać maszynom.

* * *

Mimo konkretnych źródeł osobliwości technologicznej, które jasno wskazują na jej zakorzenienie w teorii sztucznej inteligencji i wzroście technologicznym, w obecnie toczony debacie pojawiło się wiele teorii, prowadzących do rozmycia znaczenia tego terminu. Mimo to, możemy próbować wyróżnić trzy nadrzędne systematyzacje: przyspieszenie zmian, horyzont przewidywań oraz eksplozja inteligencji. Celem niniejszego artykułu jest podkreślenie źródłowego znaczenia osobliwości jako maszynowej superinteligencji – takie bowiem rozumienie pozwala wpisać *Golema XIV* w dyskurs transhumanistyczny. Przede wszystkim zaś należy rozumieć analizowaną tutaj powieść, jako opis powstania inteligencji i jej dalszy rozwój, który prowadzi do narodzin osobliwości kosmologicznej. Jednocześnie omówiono treść wykładów Golema, którego wywody wpisują się we wszystkie wymienione wyżej koncepcje – ze szczególnym jednak naciskiem na analizę w kategoriach eksplozji inteligencji – gdyż ta perspektywa jest najbliższa samemu Lemowi. Jakielkolwiek prognozy stają się siłą rzeczy niemożliwe do przewidzenia, gdyż osobliwość – tutaj jako sztuczna inteligencja – staje się dla człowieka czymś niedostępnym. O niedostępności świadczy radykalna inność Golema, sprawiająca, że dla człowieka jest on niepoznawalny. Golem to podejmujący dialog z ludźmi bezcielesny i bezosobowy Nikt.

W swoich dwu wykładach Golem omawia historię ludzkości i rozwoju rozumu – tak jak widzi je autor *Pokoju na Ziemi*. Stawia przy tym tezę, że antynomie targające *Homo sapiens* skłonią człowieka do odrzucenia jego obecnej kondycji. Wizja sformułowana przez tytułowy Rozum, jest projektem nie tylko osobliwości technologicznej, ale

dotyczy również nas – ludzi, którym przepowiada się wejście na etap autoewolucji. Łącznikiem między superinteligencją, wyrażoną tutaj pod postacią Golema, a człowiekiem, jest kluczowa rola technologii, prowadząca nieuchronnie do zakwestionowania kondycji człowieka biologicznego. Odrzuci on całe swoje dziedzictwo: historyczne, kulturowe i wejdzie w nieprzewidywalny okres, gdzie to on sam – kierowany Rozumem – będzie decydować o swoich przyszłych losach.

BIBLIOGRAFIA

- Bostrom, N. (2016). *Superinteligencja. Scenariusze, strategie, zagrożenia*. Przeł. Konowrocka-Sawa D., Gliwice: HELION.
- Bremer, J. (2010). *Wprowadzenie do filozofii umysłu*. Kraków: Wydawnictwo WAM.
- Ciechanowicz, J. (1999). *Konstanty Ciołkowski czyli Kosmos Szczęśliwy*. Rzeszów.

- Flasiński, M. (2011). *Wstęp do sztucznej inteligencji*. Warszawa: Wydawnictwo Naukowe PWN.
- Good, J. I. (1965). Speculations Concerning the First Ultraintelligent Machine. *Advances in Computers*, 6, s. 31 – 88.
- Future Trajectories: Singularity*. (2018). W: N. Vita-More, M. More (eds.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future* (361 – 363). Chichester: John Wiley & Sons, Inc., Publication.
- Jarzębski, J. (2002). *Wszechświat Lema*. Kraków: Wydawnictwo Literackie.
- Jarzębski, J. (2018). *Z przyszłych dziejów rozumu*.
<https://solaris.lem.pl/ksiazki/apokryfy/golem14/157-poslowie-golem14> [dostęp z dn. 15.11.2018].
- Klichowski, M. (2014). *Narodziny cyborgizacji. Nowa eugenika, transhumanizm i zmierzch edukacji*, Poznań: Wydawnictwo Naukowe UAM.
- Kurzweil, R. (2013). *Nadchodzi osobliwość. Kiedy człowiek przekroczy granice biologii*. Przeł. E. Chodkowska E., Nowosielska A.. Warszawa: Kurhaus Publishing.
- Lem, S. (1981). *Golem XIV*. Kraków: Wydawnictwo Literackie.
- Lem, S. (1968). *Opowieści o pilocie Pirxie*, Kraków: Wydawnictwo Literackie.
- Sandberg, A. (2013). An Overview of Models of Technological Singularity. W: N. Vita-More, M. More (eds.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future* (376 – 394). Chichester: John Wiley & Sons, Inc., Publication.
- Szpakowska, M. (1996). *Dyskusje ze Stanisławem Lemem*, Warszawa: Open.
- Vinge, V. (2013). Technological Singularity. W: N. Vita-More, M. More (eds.), *The Transhumanist Reader. Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human*

KAROLINA OWCZAREK
ROZUM WYZWOLONY.
GOLEM XIV JAKO PRZYKŁAD OSOBLIWOŚCI TECHNOLOGICZNEJ

Future (365 – 375). Chichester: John Wiley & Sons, Inc.,
Publication.

ABSTRACT

Reason Liberated. Golem XIV as an Example of Singularity

One of the most important transhumanistic ideas is singularity which can be described as a horizon of events, prediction horizon, transitional phase or emergence of conscious artificial intelligence. So nowadays it has different references. This article is an attempt to illustrate singularity based on Stanisław Lem's novel and philosophical essay *Golem XIV*. This novel shows how the evolution of AI from ordinary computer to artificial conscious superintelligence could look like. It is also an illustration of possible relations between human and non-human intelligence. Lem appears here as philosopher and visionary who tries to think about the future, when human creates a being more perfect than himself. He is far from vision of omnipotence of computers.

KEYWORDS: Stanisław Lem, singularity, artificial intelligence, transhumanism

SŁOWA KLUCZOWE: Stanisław Lem, osobliwość, sztuczna inteligencja, transhumanizm.