



Sztuczna inteligencja a tożsamość narracyjna: perspektywa transhumanistyczna

Michał KUMOREK*

ABSTRACT

Artificial intelligence and narrative identity: a transhumanist perspective: The main aim of the article is to show the changes in the structure of narrative identity as a result of combining the human brain with artificial intelligence (AI) under cyborgization. The hermeneutic structure of narrative identity proposed by Paul Ricoeur, considering the internal dialectics of the subject, its capacity to act and narrate itself, its relations with Others and orders of recognition, will be taken as the starting point of these analyses. The effects of changing the experience of unpredictability, uncertainty and memory in the lives of humans assisted by AI will be analyzed. The implications of AI assistance in the context of one's own and other people's cognition and actions, the I — Thou relationship, the mutuality of orders of recognition, and the abilities of subjects will be considered. An attempt will be made to answer the questions of whether AI assistance will increase the scope of responsibility of the subject for the actions taken or rather contribute to its reduction. The danger of losing the subject's capacity to take certain actions, such as acts of promise or forgiveness, which will be reduced to a contract in the presence of AI taking on the role of an institution, will be indicated. Finally, an attempt will be made to describe the changes in the narrative identity of a person gaining new abilities directly from AI rather than in the process of forming mutual recognitions and acquiring self-respect and social esteem, of which the subject has been partially deprived with the possibility to perform certain activities. In this context, the perspective of the formation of a Husserlian, self-transparent Self, deprived of the otherness that exposes it to criticism as a result of the connection of the human brain with AI, will be outlined.

KEYWORDS

AI; cyborgization; memory; otherness; Paul Ricoeur; phenomenology; recognition; self; subject; superintelligence; uncertainty

* Mgr filozofii, doktorant w Szkole Doktorskiej na Uniwersytecie Papieskim Jana Pawła II w Krakowie. E-mail: michal.kumorek@doktorant.upjp2.edu.pl.

WSTĘP

Współcześnie najbardziej zaawansowane badania nad sztuczną inteligencją (*artificial intelligence*, AI), które zarysowują perspektywę ulepszenia kondycji człowieka, koncentrują się na możliwości połączenia ludzkiego mózgu z maszyną. Mają na celu nie tylko leczenie, ale również rozbudowanie możliwości poznawczych, pomoc podmiotowi w podejmowaniu decyzji, a w odleglejszej perspektywie wyposażenie człowieka w nowe zdolności. Podjęty problem dotyczy potencjalnego wsparcia ludzkiego umysłu przez sztuczną inteligencję, która umożliwiłaby mu przekroczenie jego naturalnych zdolności oraz ograniczeń i w konsekwencji prowadziłaby do wykształcenia tzw. superinteligencji (Bostrom, 2016: 87–88). Analizy podjęte w artykule zostaną ograniczone do hipotetycznego przypadku połączenia ludzkiego mózgu ze sztuczną inteligencją, która zyskując wpływ na świadomość człowieka, będzie dostarczała mu nowe dane, wspomagała przeprowadzane analizy, oceniała prawdopodobieństwo zdarzeń oraz proponowała w odpowiedzi na nie zestaw etycznych działań zmierzających do osiągnięcia przez jednostkę przyjętego celu. Artykuł nie porusza kwestii technicznych, struktury mózgu podłączonego oraz utopijnej perspektywy kolektywnej świadomości, która nie tylko zniósłaby płaszczyznę uznania, ale również zagroziła indywidualnej tożsamości osoby (Žižek, 2021: 83–100, 227–228). Podjęte rozważania dotyczą etapu przedcyfrowej apokalipsy, w którym możemy jeszcze wyróżnić ludzki podmiot współpracujący z AI. W związku z tematem sztucznej inteligencji jako asysty umieszczonej w ciele człowieka pominięte zostają kwestie związane z problemem psychofizycznym (*mind-body problem*) oraz pytania o ucieleśnienia sztucznej inteligencji, która będzie otrzymywała informacje o świecie również dzięki ludzkiej percepcji, jednocześnie sama tę percepcję udoskonalając (Przegalińska, 2016: 134–160). Zaproponowane podejście wysuwa na pierwszy plan pytanie o zmiany tożsamości narracyjnej osoby mogące nastąpić w efekcie takiego połączenia, co będzie przedmiotem analiz wkraczających na płaszczyznę transhumanizmu¹.

Za punkt wyjścia przedstawionych analiz przyjęty został opis tożsamości narracyjnej osoby zaproponowany przez Paula Ricoeura. Opisany przez francuskiego filozofa podmiot odnosi się do świata i samego siebie poprzez interpretację, zrozumienie i wyjaśnianie. Koncepcja ta, uwzględniająca współczesną antropologię i wielość możliwych podejść, wskazuje, że rozumienie siebie dokonywane przez podmiot jest niejednokrotnie zapośredniczone przez ujawniającą się w różnych kategoriach inność. Jednocześnie Ricoeurowska teoria opiera się psychofizycznemu redukcjonizmowi, dzięki czemu pozwala przeanalizować

¹ Osobnego opracowania wymagałaby analiza powiązanego z asystą AI poszerzenia rzeczywistości wirtualnej dla podmiotu oraz wpływu na jego doświadczenia, działania i relacje w wirtualnym świecie.

zmiany tożsamości osoby w wyniku asysty AI. Celem artykułu będzie próba nakreślenia zmian postrzegania siebie samego i inności przez podmiot, który działałby w asyście sztucznej inteligencji. Zostanie podjęta próba odpowiedzi na pytania: Czy powiązanie AI z ludzkim mózgiem zmieni zakres odpowiedzialności podmiotu za podejmowane działania? Czy modyfikacja doświadczenia niepewności i przypadku wpłynie na symetryczność relacji z innymi ludźmi, szczególnie tymi pozbawionymi asysty AI? Ostatecznie, czy taki rodzaj cyborgizacji człowieka nie będzie oznaczał pokusy powrotu do Husserlowskiego, samoprzejrzystego Ja, pozbawionego wystawiającej go na krytykę inności?

CYBORGIZACJA JAKO WARIACJA WYOBRAŻENIOWA

Opowieść o świecie, innych ludziach i o samym sobie jest sposobem, który pozwala ująć człowiekowi rzeczywistość oraz samego siebie w narracyjne ramy. Stanowi ona istotę tożsamości narracyjnej uwzględniającej opis podejmowanego przez podmiot działania wraz z jego etycznymi implikacjami. Ricoeur, budując koncepcję tożsamości narracyjnej, odwołuje się do analiz Johna Locke'a, przywołując zakorzenione w filozoficznej tradycji wzory bycia tym samym (*idem*) i bycia sobą (*ipse*). Wzorec *idem* jako „takotożsamość” umożliwia rozpoznanie jakiegoś bytu mimo upływającego czasu jako tego samego, natomiast *ipse* jako „sobość” nie poszukuje rdzenia trwałości ujmowanego z zewnątrz, ale odnosi się do relacji podmiotu z samym sobą, która wyraża się w pamiętaniu, rozumieniu i myśleniu (Ricoeur, 2003: 192; por. Locke, 2011: 387–389). Odwołanie do analiz Locke'a Ricoeur wzbogaca o wzorce charakteru i dochowanego słowa oparte na zapośredniczeniu zrozumienia i myślenia w tym, co nosi wymiar inności. Ta zapoczątkowana przez Georga Wilhelma Friedricha Hegla dialektyka zapośredniczeń świadomości zostaje przez Ricoeura poszerzona o współczesne wątki filozoficzne wskazujące na inność własnego ciała i sumienia. Ricoeurowska teoria przekracza dystans między opowiadaniem życiem a jego doświadczaniem, ukazując wiążące je relacje. Pozwala ująć osobę jako postać, o której się opowiada, z jej własną, indywidualną historią (Ricoeur, 2003: 233–234). Sztuczna inteligencja zaimplementowana w ludzki organizm i mająca wpływ na świadomość człowieka nie pozostałaby bez wpływu na jego samorozumienie. Odpowiedź na kluczowe dla tożsamości narracyjnej pytania: „kto mówi?”, „kto dokonał tego czynu?”, „kogo historię przedstawia opowiadanie?”, „kto jest odpowiedzialny za takie bądź inne działania?” wymagałyby uwzględnienia znaczącej roli AI na wszystkich etapach życia.

Poszukiwania odpowiedzi na te pytania wymagają doprecyzowania funkcjonalności sztucznej inteligencji asystującej w życiu człowiekowi. Implementacja w ludzkim ciele i bezpośredni, aktywny wpływ na ludzką świadomość stanowią podstawę jej rozróżnienia od zewnętrznej formy asysty AI, z którą dzisiaj w różnych formach się spotykamy. Dostarczanie danych i wyników analiz z inicjatywy

sztucznej inteligencji otwiera perspektywę transhumanistyczną. Dostęp do zewnętrznych baz danych a zarazem do wrażeń zmysłowych i stanów psychofizycznych podmiotu, dostarczanie jednostce dodatkowych danych oraz analiz dotyczących probabilistyki zdarzeń, pomoc w realizacji celów w ramach norm moralnych, a ostatecznie uczenie się i wspieranie jednostki w tym wspólnym rozwoju to podstawowe cechy rozważanej sztucznej inteligencji (Bostrom, 2016: 46). Z jednej strony, rozpatrywana w kontekście hipotetycznej, harmonijnej współpracy z człowiekiem, sztuczna inteligencja wydaje się bardzo ludzka, z drugiej — może doprowadzić do poważnych zmian w kondycji człowieka, zacierając wyraźną granicę między maszyną a jej konstruktorem (Wieczorek, 2020: 55–57). Dostęp sztucznej inteligencji do świadomości człowieka prowadziłby do internalizacji dostarczanych przez nią danych, analiz, wrażeń i norm, co prowadziłoby do utraty znaczenia zapośredniczeń w tym, co inne. Połączenie sztucznej inteligencji i ludzkiego umysłu posiadałoby cechy wyróżnionej przez Nicka Bostroma szybkiej i jakościowej superinteligencji. Byłaby ona inteligentniejsza oraz szybsza niż obdarzony naturalnymi zdolnościami umysł człowieka. Sztuczna inteligencja mająca dostęp do danych zewnętrznych, w tym innych jednostek AI, posiadałaby również pewne cechy superinteligencji zbiorowej, której praktyczny wymiar ujawniałby się w inicjatywach mających na celu znalezienie zgodności pomiędzy dobrem osobistym jednostki a dobrem wspólnym (Bostrom, 2016: 88–94).

Ricoeurowska teoria nie tylko umożliwi opis różnorodnych zmian w wyniku asysty AI będącej elementem cyborgizacji, ale jest również odporna na redukcjonizmy, które z jednej strony dyskurs o tożsamości sprowadzają do pytania o niezmiennie jądro trwałości, a z drugiej — zupełnie kwestionują jej znaczenie. Uwzględnia ona również najbardziej paradoksalne i aporetyczne przypadki, odnosząc się do argumentacji Dereka Parfita, który kwestionuje przekonanie, że odpowiedzi na pytanie o istnienie struktury tożsamości można udzielić w każdym przypadku (Parfit, 2012: 241–260). Ricoeur tworzy teorię tożsamości narracyjnej w oparciu o zasadę, która uwzględnia w niej to wszystko, co normalnie działa jako zaburzenie prawidłowości. Teoria ta, posługując się narzędziami literackiej fikcji takimi jak Husserłowskie wariacje wyobrażeniowe oraz wskazując na aspekty dziejowości i czasowości każdej wymyślonej opowieści i naukowej spekulacji, umożliwia analizę kondycji ludzkiej w tych wszystkich paradoksalnych przypadkach i eksperymentach myślowych, do których zaliczyć należy również połączenie ludzkiego mózgu z AI (Ricoeur, 2003: 235–248).

PRZYPADEK, LOS I OPOWIADANIE O SOBIE

Narracja i przywoływana przez Ricoeura za Arystotelesem intryga, czyli fabuła polegająca na ułożeniu faktów, spraw i zdarzeń w ich wymiarze ilościowym i jakościowym świadczącym o jej pięknie, uległaby zmianom w wyniku

wprowadzania AI do życia osobowego jednostki (Arystoteles, 2008: 316). Przemianie uległaby dynamiczna relacja między tym, co zgodne, a tym, co niezgodne. Zgodność jest dla francuskiego filozofa porządkiem i przewidywalnym układem zdarzeń, a niezgodność stanowi efekt różnych przypadków i wydarzeń losowych. Zaburzają one początkowy porządek, ale jednocześnie rozwijają fabułę (Ricoeur, 2003: 234–235). Zapośredniczenie występujące obok współzawodnictwa między tym, co zgodne i niezgodne, Ricoeur nazywa konfiguracją. Jest to domagające się syntezy ukształtowanie fabuły przynależne sztuce kompozycji, która „polega na tym, aby ukazać ową niezgodność jako zgodną” (Ricoeur, 2008a: 70). Za pomocą konfiguracji możliwe staje się wpisanie przypadkowości i wielości wydarzeń w jedność historii oraz nadanie sensu narracji. Jedność ta może być jednocześnie burzona i rozpraszana przez różne przypadki i wydarzenia losowe, które mogą zostać wkomponowane w historię opowieści dzięki wyływającej z niej mocy jednoczenia (Ricoeur, 2003: 235). Przejście życia do jedności jest ujmowane przez intrygę. Nie odrywa się ona od życia, ale nadaje mu znaczenie poprzez opowieść o nim. Historia, w której osoba uczestniczy, jest przedstawiana tak, by wyjaśnić jej kluczowe elementy: dlatego ktoś się zachował w jakiś sposób, zmienił poglądy itp. Dokonuje się to bez przekreślania wcześniejszych wydarzeń i przekonań osoby dzięki modyfikacjom i zmianom w opowieści, które umożliwiają inne spojrzenie na fakty dzięki upływowi czasu i refiguracji. Refiguracja jako relektura umożliwia ponowne przyjrzenie się światu i sobie samemu, pozwalając na reorganizację świadomości czytelnika, który staje się poniekąd czytelnikiem samego siebie.

Opowieść o osobie przeżywającej swoje życie w asyście AI charakteryzowałaby się zmianą relacji pomiędzy tym, co zgodne, a tym, co niezgodne, na korzyść tej pierwszej. Asysta sztucznej inteligencji pozwoliłaby w jakimś ograniczonym zakresie przewidywać podmiotowi interakcje z innymi ludźmi oraz możliwy rozwój różnych sytuacji. W ten sposób ograniczona zostałaby nieprzewidywalność zdarzenia, a możliwość jego zaistnienia zostałaby poddana matematycznym prawom prawdopodobieństwa. Podmiot w asyście AI dysponowałby wiedzą o możliwości zajścia mniej prawdopodobnego zdarzenia. W ten sposób wyeliminowane zostałyby z życia człowieka wszystkie te momenty, gdy zupełny przypadek rozbija w gwałtowny i nieprzewidywalny sposób ludzką pewność. W asyście AI mniej prawdopodobne wydarzenia uwzględnione w matematycznej kalkulacji nie zaskoczyłyby podmiotu, który byłby kierowany do podjęcia najlepszych możliwych w danych okolicznościach działań. W ten sposób podmiot byłby mniej narażony na to, co przypadkowe, nieprzewidywalne i poddane działaniom innych ludzi. Bierność i pasywność podmiotu poddanego działaniom innych ludzi uległyby ograniczeniu na rzecz aktywności i optymalizacji działań w każdej sytuacji. Podmiot w asyście AI zostałby pozbawiony możliwości akceptacji nieprzewidywalnego zdarzenia, przypadku, którego prawdopodobieństwo zajścia byłoby mu znane. Tymczasem według francuskiego filozofa

to właśnie akceptacja przypadkowości w odniesieniu do historycznego losu jest koniecznym elementem dla przezwyciężenia trwogi ujawniającej się jako możliwość pozbawienia dziejów ich sensu (Ricoeur, 1991: 42). Co więcej, tworzy ona dla podmiotu możliwość przekształcenia przypadku w przeznaczenie, które niesie ze sobą nowe sensory wzbogacające dotychczasowe życie człowieka (Ricoeur, 2010: 62–63).

Jeśli decyzja o implementacji AI pozostawałaby poza horyzontem decyzyjnym podmiotu, ale byłaby efektem woli opiekunów prawnych lub systemu społecznego, to należy przywołać myśl Jürgena Habermasa pytającego, czy podmiot nadal pozostawałby autorem swojego życia, twórcą swojej biografii (Habermas, 2003: 33, 71). Zaimplementowany już na początku życia moduł AI — chociażby w celu zharmonizowania i przyspieszenia rozwoju podmiotu na wielu płaszczyznach — zmieniałby dynamikę procesów wychowawczych i socjalizacji. Kształtowane od narodzin życie podmiotu wspomaganego przez sztuczną inteligencję tylko pozornie gwarantowałoby mu — poza rolami postaci, o której się opowiada, i narratora — rolę autora swojej własnej historii. Przyswajane przez podmiot i budujące częściowo jego historię losowość i nieprzewidywalność, z którymi się zmagają, zostałyby zmodyfikowane przez asystę AI. Jednocześnie zawężone zostałyby granice samorefleksyjności osoby nad swoim życiem, podobnie jak proces jego przyswajania w ramach jego przepracowywania.

PAMIĘĆ I RESENTYMENT

Fundamentem tożsamości narracyjnej jest pamięć. To na niej są ugruntowane opowieści i historie o danej wspólnotce, ale też o konkretnej osobie. Naturalne różnice w pamięci skutkują pluralizmem narracji, które czasami się wzajemnie uzupełniają, a innym razem wchodzą z sobą w konflikt. Ten ostatni jest zazwyczaj konsekwencją nadmiaru lub niedoboru pamięci. Owe jej patologie prowadzą do tworzenia różnych narracji dotyczących tych samych wydarzeń i osób. Z jednej strony nadmiar pamięci jest spowodowany niejednokrotnie bolesnymi wydarzeniami i nie pozwala osobie na uwolnienie się od przeszłości oraz skutkuje przymusem powtarzania opracowanym przez Sigmunda Freuda. Blokują krytyczną rolę historii, uniemożliwiając jej mediację. Z drugiej strony niedobór pamięci, szczególnie w modusie biernego zapominania, zniekształca historię i uniemożliwia powiązaną z aktywnym zapominaniem zdolność przebaczenia (Ricoeur, 2006b: 106).

W ramach pracy pamięci Ricoeur rozróżnia pamięć *mnēmē-memoria* i *anamnēsis-remiŋstentia*. Ta pierwsza jest rodzajem wspomnień pojawiających się w pamięci podmiotu mimowolnie, bez aktywnego zaangażowania z jego strony, przynależąc do tożsamości *idem*. Natomiast *anamnēsis* jest pamięcią wymagającą aktywności podmiotu, który poszukuje tego, co przeszłe. Stanowi

ona rodzaj pamięci refleksyjnej oraz pełni funkcję krytyczną w stosunku do mimowolnych wspomnień. *Anamnēsis-reministentia* przynależy do tożsamości *ipse*, szczególnie w tych wszystkich wypadkach, kiedy odniesienie do niej umożliwia podmiotowi podejmowanie i planowanie działań. Ten perspektywny wymiar czasem zresztą zawiera w sobie elementy reminiscencji, tak jak obietnica zawiera w sobie aspekt niezapominania momentu jej złożenia, który odchodzi w przeszłość.

Asysta AI towarzysząca człowiekowi od urodzenia zmieniałaby zależność pomiędzy pamięcią *mnēmē-memoria* i *anamnēsis-reministentia*. Zmiana ta mogłaby następować w wielu kierunkach, pośród których zarysować należy dwa skrajne bieguny. W pierwszym z nich AI stanowiłaby źródło informacji o przeszłości człowieka, zawierając informacje nie tylko o wydarzeniach i wrażeniach zmysłowych, ale również o emocjach towarzyszących podmiotowi. W ten sposób byłaby ona odpornym na zapomnienie zbiorem danych, do którego podmiot w każdej chwili mógłby się odwołać. W drugim przypadku, bardziej odpowiadającym sztucznej inteligencji obdarzonej inicjatywą działania i uczącej się, pamięć podmiotu byłaby przez nią współkształtowana. AI motywowałaby podmiot do przywoływania konkretnych wydarzeń z przeszłości w ramach *anamnēsis-reministentia* lub byłyby one ujawniane przez sztuczną inteligencję w świadomości podmiotu jako wspomnienia przynależące *mnēmē-memoria*. W ten sposób AI mogłaby pomagać podmiotowi w realizacji jego życiowych planów lub współkształtować jego działania i plany, podporządkowując je dobru wspólnemu i godząc je z planami innych (Ricoeur, 2006b: 102–105).

Współkształtowanie pamięci przez sztuczną inteligencję niesłoby poważne konsekwencje dla kształtowania się tożsamości narracyjnej podmiotu. Sztuczna inteligencja przejęłaby inicjatywę w kształtowaniu pamięci podmiotu, zarówno wtedy, kiedy motywowałaby go do pracy pamięci, jak też wtedy, gdy sama przywoływałaby wspomnienia. W ten sposób AI zdobywałaby również inicjatywę kształtowania działań podmiotu i jego zdolności, w tym opowieści jego samego o nim samym. Szczególnie zagrożony byłby ludzki akt wybaczenia, występujący w różnych modusach i prowadzący do przekształcenia pamięci. AI zdolna do wymazywania z pamięci trudnych lub traumatycznych doświadczeń uniemożliwiłaby podmiotowi zmianę zawartego w pamięci odniesienia do świata, która jest częścią jego tożsamości. Opowieść o złych i dobrych czynach zarówno popełnianych, jak i doświadczanych od innych, stanowi jego treść. Współkształtowanie pamięci przez AI niesłoby zagrożenie manipulacją narracji o przeszłości. Zagroziłoby również wolności w poszukiwaniu przez człowieka szczęścia, które jest powiązane z kierowaniem się na cele i ich osiągnięciem. Ich odgórne podporządkowanie przez AI dobru wspólnemu ograniczyłoby możliwości jego osiągnięcia (Wójtowicz, 2022: 13–14). Jednocześnie akt refleksji człowieka nad utraconą zdolnością pracy nad pamięcią niesłoby zagrożenie resentymentem i fatalizmem. Byłyby one nie tylko skutkiem — jak zaznacza

Habermas w analogicznej sytuacji zaprojektowania genomu — utraty władzy nad własną pamięcią, ale również wynikiem niemożliwości przyswojenia własnej historii i brakiem przepracowania błędów w trybie etycznej samorefleksji (Habermas, 2003: 21).

WZAJEMNE UZNANIE I DZIAŁANIE PODMIOTU

Uzupełniając koncepcję tożsamości narracyjnej o teorię wzajemnego i obopólnego uznania, Ricoeur zakreśla możliwości rozwijania pokojowych relacji międzyludzkich nawet w sytuacji, gdy modelowa symetryczność zostanie zakłócona. Rozwijając swoją koncepcję francuski filozof opiera się na modelu Axela Honnetha i wyróżnia za nim trzy wymiary uznania na płaszczyznach miłości, prawa i poważania społecznego. Budowana na płaszczyźnie miłości ufność, prowadząca do nabywania przez podmiot pewności, będzie rezonowała na poziomie wzajemnego uznania na płaszczyźnie prawnej i poważania społecznego, dając oparcie tym wszystkim działaniom, które mimo że skierowane w stronę gwarantowanej instytucjonalnie symetrii, w codzienności niejednokrotnie od niej odbiegają. Ta nabywana od początku życia ufność stanowi podstawę wiary w siebie, szacunku i poważania siebie, które uzdalniają podmiot do nowych działań (Honneth, 2012: 26–28). Poświadczają one tożsamość osoby jako zdolnej do działań i stanowią etyczną podstawę zabezpieczającą podmiot przed oparciem się wyłącznie na strukturze instytucjonalnej, która może zastąpić wewnętrzne motywy i zagrozić pokojowym relacjom, gdy tej podpory zabraknie lub gdy zostanie ona pozbawiona znamion sprawiedliwości.

Ricoeur zwraca uwagę, że ludzkie działanie często zawiera w sobie wiele innych czynności, od bardziej złożonych po najbardziej podstawowe. Są one poddane pewnym nadającym sens całości prawom i zasadom, które francuski filozof nazywa regułą konstytutywną. Jako przykład podaje zarządzanie państwem, które podporządkowuje sobie wiele innych działań połączonych wspólnym sensem związanym z dobrem wspólnym. Ricoeur zauważa, że reguła konstytutywna nie jest zewnętrzna w stosunku do działań, ale nadaje im znaczenie i może być przeniesiona na płaszczyznę aktów mowy, które również są działaniami ustanowionymi przez regułę. Objawiają się one szczególnie w intencjonalnych, illokucyjnych aktach mowy wypowiedzianych w formie obietnic, prośb czy rozkazów (Austin, 1993: 645). Reguła konstytutywna pozwala również na podkreślenie charakteru interakcji działań. Nabierają one wtedy wymiaru społecznego ze względu na ich podejmowanie w odpowiedzi na działania innych. To odniesienie generuje gęstą sieć powiązań będących stosunkami interakcji, które mogą przybierać różne formy, od współpracy, przez współzawodnictwo, po konflikt. Stosunki interakcji zostają uwewnętrznione, gdy z jednej strony pozwalają nam na uniezależnienie się od pewnych innych działań w ramach

nabywanych kompetencji, a z drugiej — zawsze odsyłają do działań osób, od których się uczyliśmy (Ricoeur, 2003: 253–258).

Uwzględnienie asysty AI w teorii działania podmiotu prowadzi do pytania o możliwość oderwania reguły konstytutywnej od niektórych aktów. Mimo że lepsza znajomość powiązań różnych działań wspierałaby jedność konfiguracji poddanej obranemu celowi, to czy nie opierałaby się ona jednak bardziej na AI niż na wewnętrznej regule konstytutywnej danych działań? Wątpliwość ta wybrzmiewa szczególnie w analizie działań o charakterze etycznym. Akty takie jak obiecywanie czy wybaczenie oderwane zostałyby nie tylko od swoich własnych, konstytuujących sens zasad, ale również od złotej reguły porządkującej międzyludzkie relacje w sytuacji asymetrii (Ricoeur, 2003: 445). Pozbawienie działań takich jak obiecywanie czy przebaczenie reguły konstytutywnej prowadziłoby do ograniczenia możliwości zdobywania szacunku do siebie i poważania siebie (Ricoeur, 1995: 29–30; por. Habermas, 2003: 36). Zmianie uległaby również wskazywana przez Ricoeura zależność między etyką i moralnością, w momentach konfliktów i podporządkowania normy dążeniu etycznemu wymagająca mądrości praktycznej — arystotelesowskiej *fronesis* (Ricoeur, 2004: 73–78). Utrata tej ostatniej byłaby konsekwencją roli AI wspierającej podmiot zbiorem norm i nakazów naznaczonych roszczeniem do powszechności, a być może również jakąś formą przymusu. Prowadziłaby ona do automatyzacji rozwoju moralnego świadomości, której działania nie byłyby efektem namysłu etycznego i refleksji. Pod znakiem zapytania stałaby możliwość nie tylko zyskania szacunku do siebie w wyniku podejmowania etycznych działań, ale również poważania siebie zakładającego poważanie innych podobnie jak siebie samego. Wymaga ono zaufania i poświadczenia konstytuujących relacje we wspólnocie zarówno w ramach narzuconych reguł, jak i wspólnych, przyjętych za swoje obyczajów i tradycji (Warmbier, 2018: 297–305).

NIEPEWNOŚĆ I RELACJA Z INNYM

Ricoeur podkreśla, że relacja Ja–Ty ma fundamentalne znaczenie dla wzajemnej konstytucji tożsamości narracyjnej oraz ustanowienia fenomenologii niepewności. Poszukując podłoża dla relacji wzajemnego uznania, francuski filozof ukazuje konieczność symetryczności pomiędzy podmiotami, których wzajemna konstytucja nie znosi bezpośrednio niepewności, ale odślania warunki i możliwości jej przewyżczenia. Ricoeurowski opis wyłania się z dwóch krytyk, pierwszej, dotyczącej idealistycznej teorii Edmunda Husserla konstytucji innego w Ja i przez Ja, oraz drugiej, wskazującej na Levinasowską asymetrię w konstytucji Ja wyłącznie przez innego (Ricoeur, 2004: 148–157). Poza brakiem poszukiwanej przez Ricoeura równowagi tego samego i innego w dialektyce podmiotu asymetrii te niosą zupełnie różne doświadczenia niepewności

związane z innością innego. Ricoeur, analizując zarysowany przez Husserla proces konstytucji podmiotowości innego jako *alter ego*, krytykuje jego całkowitą dostępność i oczywistość dla świadomości. Idealistyczna teoria niemieckiego filozofa zakładająca konstytucję innego w Ja ograniczałaby niepewność związaną z drugim człowiekiem i tworzyłaby niebezpieczną iluzję poznania inności innego. Drugą kwestionowaną przez Ricoeura koncepcją opisującą relację Ja–Ty w konstytucji tożsamości podmiotu jest Levinasowska idea zakładnika mająca uniemożliwić zdobycie panowania nad innym przez Ja. Oddanie całkowitej inicjatywy innemu w konstytuowaniu sobości następuje w wyniku objawienia, epifanii twarzy innego. Jej wezwanie „nie zabijaj” oparte na fundamencie etycznym i zupełnie pozbawione epistemologicznej aktywności podmiotu Ja domaga się odpowiedzi (Levinas, 1998: 74–75). Gdy relacja Ja–Inny zostanie doprowadzona do skrajnej asymetryczności i zakwestionowana zostanie jakkolwiek możliwość komunikacji, przyniesie ona nieskończony ładunek niepewności. Ostatecznie inny może okazać się nie mistrzem sprawiedliwości, ale katem. Levinasowska hiperbola pociąga za sobą nieprzewidywalną hiperbolizację niepewności, pozostawiając pasywny podmiot na łasce innego.

Ricoeur poszukuje symetrycznej relacji Ja–Ty, która nie będzie usuwała niepewności u początku konstytucji, podobnie jak nie będzie jej hiperbolizowała bez możliwości przewyciężenia. Tworzy ona raczej pole współkonstytucji, w którym pokonywanie niepewności jest związane z ruchem angażującym poza sobością również inność innego i własną inność Ja. Niepewność będzie pojawiała się jako element w porządku genezy sensu po obu stronach relacji wraz z fenomenem inności innego. Konfrontacja z niepewnością w zaproponowanej przez Ricoeura relacji Ja–Ty jest wzajemna, ale w swoim horyzoncie możliwości nie przynosi ona pewności rozumianej jako pokonanie wszelkich wątpliwości. W zamian proponuje rodzaj ufności przynależny poświadczeniu, będącej wiarą w innego (Jakubowski, 2017: 182). Poświadczenie stawia innego w relacji osobowej nie tylko jako kogoś, kto jest ośrodkiem działań, ale również jako tego, który uważa mnie za podmiot obdarzony różnymi zdolnościami. Ta wzajemność poświadczeń uwzględnia również przynależną drugiej osobie możliwość doznawania, a więc i doświadczenia niepewności pochodzącej z mojej strony.

Relacje Ja–Ty oraz fenomenologia niepewności uległyby zmianom przy uwzględnieniu asysty AI mającej na celu uzyskanie przez podmiot jak największych pewności i przewidywalności. Doskonalsze i szybsze od ludzkiego pozyskiwanie danych przez AI wraz z ich interpretacją i analizami dostarczonymi człowiekowi zmniejszyłoby możliwość wystąpienia ludzkich błędów w tych obszarach. Mimo że niepewność nie zostałaby całkowicie zniesiona, to byłaby ona powiązana ze znajomością prawdopodobieństwa zdarzeń. Ta matematyzacja niepewności towarzyszącej własnej nieprzejrzystości sięgnęłaby również innego razem z wiedzą na temat jego działań, motywacji i planów. Znajomość ich prawdopodobieństwa i uwzględnienie go w kontekście podejmowanych przed

podmiot decyzji implikowałyby zmianę dążących do symetryczności relacji. Proces przewidywania niepewności zostałby pozbawiony wzajemności przez jednostronną, opartą na statystyce konstytucję innego wraz z jego działaniami i motywacjami w Ja i przez Ja. Asysta AI po obu stronach relacji Ja–Ty tylko pozornie wprowadziłaby równowagę. W takim przypadku podmioty stałyby się dla siebie jeszcze bardziej przewidywalne i przejrzyste, opierając swoje relacje bardziej na zabezpieczonej zewnętrznym prawem umowie niż rozpraszać niepewność obietnicy i zaufaniu (Arendt, 2010: 277). Asysta sztucznej inteligencji po obu stronach relacji znosiłaby instytucję świadka w przypadku obietnicy, sama stanowiąc instytucjonalną gwarancję dochowania słowa. Matematyzacja niepewności i zmiana jej doświadczania nie tylko doprowadziłyby do umaszynowienia i zautomatyzowania międzyludzkich relacji, ale zmieniłyby również proces poznawania drugiego. Aspekt epistemologiczny relacji usunąłby w cień relacje etyczną, która zostałaby zdominowana przez moralną normę wbudowaną w systemy AI.

ZMIANA DOŚWIADCZENIA INNOŚCI

Przebudowa relacji społecznych i porządków uznania oraz modyfikacja doświadczenia niepewności i przypadkowości skutkowałyby zmianami w postrzeganiu przez podmiot inności, która w dialektyce z byciem sobą samym stanowi ramy dla zaproponowanego przez Ricoeura opisu tożsamości osoby. Francuski filozof wskazuje, że inność w dialektyce z „byciem sobą” współkonstruuje osobowy, ucieleśniony podmiot *le soi* (Drwięga, 2022: 158–160). Inność ujawnia się jako fenomen pasywności i bierności w doświadczeniu fenomenologicznym tego-który-jest-sobą-samym (Ricoeur, 2003: 528–529). Nie jest ona przez niego w pełni uchwytywana i poznawalna w akcie refleksji. Inność podważa samodzielną rolę świadomości w konstytucji podmiotu i daje o sobie znać w bardzo wielu różnych doświadczeniach. Jednym z nich jest wymykająca się sobości niepewność, powiązana z rozważaną wcześniej przez Ricoeura w *Philosophie de la volonté* mimowolnością. Ta zaś wskazuje na pierwotnie warunki bycia w świecie podmiotu towarzyszące mu od narodzin aż do śmierci (Ricoeur, 1950: 399–402). Ricoeurowski podmiot jest ograniczony przez to, co mimowolne, i nawet w tym, co dobrowolne, w świetle świadomości, towarzyszy mu jakaś niepewność związana z tymi ograniczeniami. Nie mogą być pewien, czy ciało zawsze wytrzyma narzucony mu wysiłek lub zniesie ból i chorobę, czy jakaś sytuacja nie ujawni traumy a marzenie senne ukrytego w nieświadomym pragnienia, czy spotkanie z innym nie ukaże niechcianych przeze mnie cech charakteru. Ostatecznie podmiot nie może mieć pod całkowitą kontrolą życia, któremu zagraża niepewny w swoim czasie i miejscu moment śmierci. Tym samym nawet zwracająca się ku sobie samej świadomość zostaje pouczona przez

to, co mimowolne, o konieczności skierowania się ku inności zawsze obciążonej niepewnością (Kumorek, 2022: 219–226). Niepewność nie tylko jest związana z innym człowiekiem, ale stanowi również element świata wewnętrznego podmiotu współtworzonego przez przynależną mu inność. Pośród niej francuski filozof wyróżnia fenomeny cielesności i sumienia, z których pierwszy będzie wielokrotnie źródłem niepewności, a drugi jako „zew, który idzie ze mnie a przecież przeze mnie” (Heidegger, 1994: 346), będzie paradoksalnie ujawniał pewność nakazu lub zakazu, jednocześnie pozostawiając niepewność co do pochodzenia tego głosu.

Asysta AI, zmieniając doświadczenie niepewności, przypadkowości i mimowolności, modyfikowałaby rolę inności w konstytucji tożsamości podmiotu. Mimo że podmiot pozostawałby nadal ucieleśniony, to doświadczenie żywej cielesności *Leib* zostałoby poddane większej kontroli, a podmiot zyskałby nowe zdolności umożliwiające ograniczenie wpływu tego wszystkiego, co jest w jego ciele przeżywane jako inne. Cyborgizacja podmiotu już na poziomie samej tylko asysty AI kierowałaby doświadczenie ciała w stronę fizyczności *Körper* i w konsekwencji dominowałaby również postrzeganie cielesności innego człowieka (Husserl, 1982: 141–144). Najbardziej problematyczna jest analiza hipotetycznej zmiany doświadczenia inności sumienia w wyniku asysty AI. Biorąc pod uwagę poświadczającą rolę sumienia, nie można wykluczyć zaniku jego funkcjonowania w wyniku nieustannego uzyskiwania przez podmiot jakiejś formy prawdopodobnej pewności lub po prostu utożsamiania jego głosu z analizami i propozycjami etycznych działań dostarczanych przez AI pełniącą rolę *superego* (Lubowicka, 2000: 79–84). Analiza zmian doświadczanej inności sumienia wymagałaby dokładnego opisu struktur modelu psychoanalitycznego i zbadania, w jakim zakresie to, co nieświadome, byłyby uchwytnie dla AI lub wnioskowane z zachowania i działania podmiotu (Žižek, 2021: 188–189).

Asysta AI zmieniłaby relacje pomiędzy byciem tym samym i byciem sobą człowieka, zestawianych przez Ricoeura we wzorcach charakteru i dotrzymanego słowa. Ricoeurowska analiza tego pierwszego wskazuje na nakładanie się *idem* i *ipse* w tych wszystkich momentach, kiedy podmiot jest rozpoznawany przez innych po danych cechach charakteru oraz gdy on sam siebie również rozpoznaje jako ich posiadacz. Natomiast model dochowanego słowa rozłącza tożsamość *idem* i *ipse*, ponieważ dotrzymanie obietnicy spoczywa na podmiocie, na jego byciu sobą (Ricoeur, 1992: 33–36). Wskazywana przez Habermasa „technicyzacja etycznej samowiedzy gatunkowej” znosząca potrzebę obietnicy i darowania pozbawiłaby częściowo podmiot zdolności do postrzegania siebie w kategoriach *ipse* oraz do rozpoznawania innych osób jako zdolnych do bycia samodzielnymi podmiotami działania (Habermas, 2003: 49). W konsekwencji przebudowa relacji społecznych i postrzeganie innych wyłącznie w kategoriach *idem* podważyłaby rozpoznanie przez innych również samego podmiotu jako zdolnego do działań.

POKUSA SAMOPRZEJRZYSTOŚCI

Chęć wiedzy i opartej na niej zdolności podejmowania decyzji jest jednym z motorów napędowych rozwoju cywilizacji, ale też rozwoju podmiotów konstytuujących te cywilizacje. Poznanie zarówno rzeczywistości, jak i samego siebie zawsze leżało w obszarze zainteresowania człowieka. Niekiedy to zainteresowanie podmiotu samym sobą nie dopuszcza możliwości nieznaności siebie, zakładając samoprzejrzyste kartezjańskie Ego (Kartezjusz, 2004: 34–38). Innym razem niezaakceptowana nieznanost siebie staje się dla podmiotu motywacją do skazanej na niepowodzenie próby jej eliminacji z wewnętrznego życia. A mieszcząca się między tymi skrajnymi biegunami postawa zaakceptowania swojej nieprzejrzystości, kierująca podmiot na długą drogą samorozumienia, jest usłana niepewnością, upadkami oraz utratami wymagającymi przepracowania (Ricoeur, 1960: 24–26; Ricoeur, 2008b: 343–355). Sztuczna inteligencja może stanowić obietnicę modyfikacji tych wymagających czasu i niejednokrotnie uciążliwych z punktu widzenia podmiotu procesów prowadzących do poznania siebie.

Jednym z filarów Ricoeurowskiej teorii tożsamości narracyjnej jest inspirowana myślą filozofów podejrzeń krytyka samoprzejrzystego podmiotu w koncepcjach Kartezjusza i Husserla (Ricoeur, 1985: 238–241). Kartezjańska formuła „myślę, więc jestem” może być według Ricoeura co najwyżej założeniem w filozofii, a nie wnioskiem lub weryfikowalnym faktem. Powołując się na mistrzów podejrzeń, francuski filozof wskazuje, że kartezjańskie „Ja myślące” jest abstrakcyjnym i fikcyjnym podmiotem wyrwanym z kultury i historii. Ricoeur odwraca perspektywę, argumentując za pierwszeństwem bycia podmiotu w świecie, w którym poznanie jest podporządkowanie egzystencji. Podważa zwrot ku idealizmowi Husserla, dostrzegając w nim zagrożenie dominującej roli świadomości opierającej się na percepcji i ignorującej całą historyczność. Ten błąd idealizmu, w którym podmiot konstytuuje w Ja i przez Ja całą rzeczywistość, wraz z innością innego jest przez Ricoeura poddawany wieloaspektowej krytyce (Ricoeur, 2006a: 194–233). Wskazuje on, że interpretacja siebie wychodzi poza wyłączną sferę świadomości, ujmując cały kontekst Heideggerowskiego „bycia-w-świecie” wraz z językiem, historią, kulturą itp. Wymaga ona czasu i działań, w których podmiot poznaje sam siebie, ale jest też narażony na błędy w tym poznaniu oraz zafałszowanie świadomości siebie (Rosner, 1989: 13–20).

Wspomożenie procesorów myślenia przez AI dostarczającą dane wprost do świadomości osoby modyfikowałoby przebieg procesu przekształcania się świadomości w samoświadomość. Zmiana ta następowałaby w procesie częściowego utożsamiania świadomości ze sztuczną inteligencją, która byłaby elementem wspomagającym wyjaśnianie oraz rozumienie rzeczywistości, a być może w pewnym stopniu również decydującym o działaniach podmiotu chociażby dzięki roli jakiejś formy panoptyzmu (Foucault, 2020: 291–337). W takim wypadku początkowe skierowanie świadomości ku czemuś zewnętrznemu, przekraczającemu ją

samą, zostałyby zwrócone ku wnętrzu, ku niej samej, a więc stałyby się bliskie Husserlowskiej, idealistycznej interpretacji. Nawet jeśli zewnętrzne czynniki i zapośredniczenia nie zostałyby zupełnie wyeliminowane, to ich wyjaśnianie zdominowałyby uwewnętrznione procesy oparte nie na percepcji, jaką w *Ideach* i *Medytacjach* proponował Husserl, ale raczej na wewnętrznym strumieniu analiz pochodzący od sztucznej inteligencji, która przetwarzałaby i interpretowałaby dane dostarczane przy pomocy zmysłów, wzbogacając je dodatkowo o informacje na temat otaczającego podmiot świata (Husserl, 1974: 79–127). W ten sposób rola innych, świata kultury czy historii zostałyby w znaczący sposób ograniczona, co nie pozostałoby bez wpływu na rozumienie samego siebie przez podmiot wyposażony w AI. Świadomość, w intencjonalnym ruchu ku danym dostarczanym przez sztuczną inteligencję, kierowałaby się w stronę samej siebie, zyskując rangę podobną do Husserlowskiej samowiedzy. Podmiot nie interpretowałby siebie za pośrednictwem rozumienia i interpretacji swoich działań, przeszłości i wytworów, ale tak jak chciał Husserl w *Ideach* i *Medytacjach*, samorozumienie następowaloby wprost (Husserl, 1974: 347–352). Umiejscowienie procesu rozumienia siebie wyłącznie w samym sobie zmieniałoby formowanie się tożsamości narracyjnej podmiotu, która z asystą AI zmierzałaby do kartezjańskiej, samoprzejrzystej myśli stanowiącej samą siebie (Žižek, 2021: 223). Jednak to, co bezpośrednio dane w samoprezentacji, byłoby dostępne sztucznej inteligencji, która zyskując cechy superinteligencji zbiorowej mogłaby je przekazywać dalej, innym podmiotom (Diederich, 2021: 66–68). Podmiot w procesie przekształcania się w superinteligencję traciłby jednocześnie inicjatywę w zwracaniu się ku sobie, a treści świadomości dotyczące nie tylko spostrzeżeń zmysłowych, ale również uczuć, woli i wyobrażeń znajdowałyby się pod coraz większą kontrolą sztucznej inteligencji. Człowiek w asyście AI stawałby się pasywny, przy założeniu, że będzie jeszcze można w ogóle stwierdzić odrębność ludzkiego podmiotu, a nie tylko ludzki komponent superinteligencji.

PODSUMOWANIE

Przeprowadzone analizy wskazują, że asysta sztucznej inteligencji w celu wsparcia optymalnych działań podmiotu doprowadziłaby do radykalnej zmiany jego doświadczeń oraz procesu kształtowania tożsamości narracyjnej. Podmiot wyposażony w asystę AI znalazłby się w dużo bardziej przewidywalnym dla siebie świecie, w którym nie byłoby miejsca na zupełny przypadek. Zachowania własne oraz innych ludzi byłyby wówczas znane podmiotowi wraz z prawdopodobieństwem ich wystąpienia. Zmiana doświadczanej niepewności, przypadkowości i pamięci zmodyfikowałaby możliwości odnoszenia się podmiotu w aktach refleksji do własnego życia. Człowiek w asyście AI zostałby pozbawiony możliwości przepracowania błędów, których nie popełniłby, a eliminacja z życia różnych

zmagania i walki o uznanie odrywałaby go od przeżywania życia jako jego autor. Życie to byłoby również uboższe o pewne działania, których podmiot, poddany moralnym normom pochodzącym od AI, nie musiałby podejmować i ponosić za nie odpowiedzialności. Takie działania jak obiecywanie, przebaczenie czy darowanie straciłyby na znaczeniu w społeczeństwie, w którym sztuczna inteligencja pełniłaby rolę instytucjonalnego gwaranta. Asysta AI prowadziłaby do oderwania od tych działań wewnętrznej względem nich reguły konstytutywnej. Wraz z przebudową porządków uznania i uniezależnieniem się od innych podmiot kształtowałyby swoje zdolności w mniejszym stopniu w oparciu na byciu rozpoznany i uznanym, a bardziej na umiejętnościach bezpośrednio uzyskanych dzięki sztucznej inteligencji.

Rozważanie zmian wynikający z asysty AI w kontekście Ricoeurowskiej koncepcji tożsamości narracyjnej ukazuje również modyfikację relacji pomiędzy klasycznymi wzorcami bycia sobą i bycia tym samym na korzyść tego pierwszego. Równoczesna utrata znaczenia inności spowodowałaby, że człowiek postrzegałby i rozpoznawał siebie w bezpośrednim zwrocie ku sobie. Modyfikacja doświadczeń niepewności poddanej przez AI zasadom prawdopodobieństwa pozbawiłaby dotychczasowej roli zaufanie, troskę i szacunek, które pojawiają się na horyzoncie długiej drogi wzajemnych uznań. Tym samym podważona zostałaby rola inności w konstytucji podmiotowego *soi*. Samoprzejrzysty, neokartezjański podmiot w swojej konstytucji odnosiłby się do siebie bez potrzeby zapośredniczeń w inności innego, sumienia, a nawet wydziedziczonej z inności własnej cielesności. Paradoksalnie podmiot stawałby się bierny w tym zwrocie ku sobie, tracąc inicjatywę na rzecz sztucznej inteligencji.

BIBLIOGRAFIA

- Arendt, H. (2010). *Kondycja ludzka*. (Przeł. A. Łagodzka). Warszawa: Aletheia.
- Arystoteles. (2008). *Poetyka*. (Przeł. H. Podbielski). Warszawa: Wydawnictwo Naukowe PWN.
- Austin, J.L. (1993). *Mówienie i poznawanie. Rozprawy i wykłady filozoficzne*. (Przeł. B. Chwedończuk). Warszawa: Wydawnictwo Naukowe PWN.
- Bostrom, N. (2016). *Superinteligencja. Scenariusze, strategie, zagrożenia*. (Przeł. D. Konowrocka-Sawa). Gliwice: Helion.
- Diederich, J. (2021). *The psychology of artificial superintelligence*. Cham: Springer.
- Drwięga, M. (2020). *Odpowiedzialna wolność*. Kraków: Księgarnia Akademicka.
- Foucault, M. (2020). *Nadzorować i karać. Narodziny więzienia*. (Przeł. T. Komendant). Warszawa: Aletheia.
- Habermas, J. (2003). *Przyszłość natury ludzkiej. Czy zmierzamy do eugeniki liberalnej?*. (Przeł. M. Łukasiewicz). Warszawa: Wydawnictwo Naukowe Scholar.
- Heidegger, M. (1994). *Bycie i czas*. (Przeł. B. Baran). Warszawa: Wydawnictwo Naukowe PWN.
- Honneth, A. (2012). *Walka o uznanie. Moralna gramatyka konfliktów społecznych*. (Przeł. J. Duraj). Kraków: Zakład Wydawniczy Nomos.
- Husserl, E. (1974). *Idee czystej fenomenologii i fenomenologicznej filozofii. Księga druga*. (Przeł. D. Gierulanka). Warszawa: Państwowe Wydawnictwo Naukowe.

- Husserl, E. (1982). *Medytacje Kartezjańskie*. (Przeł. A. Wajs). Warszawa: Państwowe Wydawnictwo Naukowe.
- Jakubowski, J. (2017). *Skończoność egzystencjalna. Studium nad filozofią Paula Ricoeura*. Bydgoszcz: Epigram.
- Kartezjusz. (2004). *Medytacje o filozofii pierwszej*. (Przeł. J. Hartman). Kraków: Wydawnictwo Zielona Sowa.
- Kumorek, M. (2022). Rola wzorców tożsamości osobowej w Ricoeurowskiej hermeneutyce podmiotu. *Roczniki Filozoficzne*, 70(3), 205–230. DOI: 10.18290/rf22703.9.
- Levinas, E. (1998). *Całość i nieskończoność. Esej o zewnętrżności*. (Przeł. M. Kowalska). Warszawa: Wydawnictwo Naukowe PWN.
- Locke, J. (2011). *Rozważania dotyczące rozumu ludzkiego. Księga I i II*. (Przeł. W.M. Kozłowski). Warszawa: Hachette.
- Lubowicka, G. (2000). *Sumienie jako poświadczenie. Idea podmiotowości w filozofii Paula Ricoeura*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Parfit, D. (2012). *Racje i osoby*. (Przeł. W.M. Hensel & M. Warchala). Warszawa: Wydawnictwo Naukowe PWN.
- Przegalińska, A.K. (2016). *Istoty wirtualne. Jak fenomenologia zmieniła sztuczną inteligencję*. Kraków: Universitas.
- Ricoeur, P. (1950). *Philosophie de la volonté* (t. 1: *Le volontaire et l'involontaire*). Paris: Aubier.
- Ricoeur, P. (1960). *Philosophie de la volonté* (t. 2: *Finitude et culpabilité (parte prima): l'homme faillible*). Paris: Aubier.
- Ricoeur, P. (1985). O pewnej filozoficznej interpretacji Freuda (s. 241–259). (Przeł. H. Igalson). W: P. Ricoeur. *Egzystencja i hermeneutyka. Rozprawy o metodzie*. Warszawa: Pax.
- Ricoeur, P. (1991). Trwoga rzeczywista i złudna (s. 31–50). (Przeł. P. Kamiński). W: P. Ricoeur. *Podług nadziei. Odczyty, szkice, studia*. Warszawa: Pax.
- Ricoeur, P. (1992). *Filozofia osoby, wykłady z filozofii współczesnej im. ks. prof. Konstantego Miłchalskiego CM*. (Przeł. M. Frankiewicz). Kraków: Wydawnictwo PAT.
- Ricoeur, P. (1995). *Le juste*. Paris: Le Seuil.
- Ricoeur, P. (2003). *O sobie samym jako innym*. (Przeł. B. Chelstowski). Warszawa: Wydawnictwo Naukowe PWN.
- Ricoeur, P. (2004). *Drogi rozpoznania*. (Przeł. J. Margański). Kraków: Znak.
- Ricoeur, P. (2006a). Fenomenologia i hermeneutyka: wychodząc od Husserla... (s. 194–233). (Przeł. M. Drwięga). W: J. Migasiński & I. Lorenc (Red.). *Fenomenologia francuska. Rozpoznania — interpretacje — rozwinięcia*. Warszawa: Wydawnictwo IFiS PAN.
- Ricoeur, P. (2006b). *Pamięć, historia, zapomnienie*. (Przeł. J. Margański). Kraków: Universitas.
- Ricoeur, P. (2008a). *Czas i opowieść* (t. 1: *Intryga i historyczna opowieść*). (Przeł. M. Frankiewicz). Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Ricoeur, P. (2008b). *Czas i opowieść* (t. 3: *Czas opowiadany*). (Przeł. U. Zbrzeźniak). Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Ricoeur, P. (2010). *Miłość i sprawiedliwość*. (Przeł. M. Drwięga). Kraków: Universitas.
- Rosner, K. (1989). Paul Ricoeur — filozoficzne źródła jego hermeneutyki (s. 5–60). W: P. Ricoeur. *Język, tekst, interpretacja*. (Przeł. K. Rosner). Warszawa: Państwowy Instytut Wydawniczy.
- Warmbier, A. (2018). *Tożsamość, narracja i hermeneutyka siebie. Paula Ricoeura filozofia człowieka*. Kraków: Universitas.
- Wieczorek, K. (2020). Orszak wdzięcznych marionetek, czyli maszyna jako urzekający potwór. *Logos i Ethos*, 52, 49–77. DOI: 10.15633/lie.3577.
- Wójtowicz, M. (2022). Transhumanistyczna koncepcja szczęścia postczłowieka — krytyczna analiza wybranych aspektów. *Teologia i Moralność*, 31(1), 9–18. DOI: 10.14746/TIM.2022.31.1.1.
- Žižek, S. (2021). *Hegel i mózg podłączony*. (Przeł. M. Kropiwnicki). Warszawa: Wydawnictwo Krytyki Politycznej.