

ARTYKUŁ RECENZYJNY

RYSZARD STEMPOWSKI

TEORIA POLITYCZNA W ERZE CYFROWEJ?

Autor recenzowanej książki¹, profesor Uniwersytetu Harvardzkiego, zaczyna od porównania teorii politycznej z filozofią polityczną: „Political philosophy differs from political theory by focusing on the depth of arguments made in and about the political domain, holding at bay not only much of the societal background but also other kinds of inquiry that might bear on the question at hand. Political theory is more broadly oriented and manages its trade-offs differently, tending to integrate more breath at the expense of in-depth investigation of ideas. At this stage, in the process of transferring political thought to the digital domain, it is important to take the more comprehensive view of political theory rather than the narrower one of political philosophy to make sure we transfer the debates in their full richness into this new era” (s. XII).

Czy rzeczywiście tak się te dziedziny różnią od siebie, a zakres teorii politycznej jest szerszy od zakresu filozofii politycznej? Taka opinia odzwierciedla raczej różnicę między wydziałami filozofii i politologii w Stanach Zjednoczonych. To tam jest najwięcej socjologów i filozofów postmodernistycznych i ludzi, którzy łączą, powiedziałbym, *normative reflection with empirical work*, a praktycznie nie ma filozofów analitycznych, którzy zajmują się tylko filozofią. A przecież filozofia, która łączy się z empirycznym badaniami, to wciąż filozofia, a nie coś innego. Ta

RYSZARD STEMPOWSKI, DYREKTOR POLSKIEGO INSTYTUTU SPRAW MIĘDZYNARODOWYCH 1999–2004, WWW.STEMPOWSKI.PL

¹ M. Risse, *Political Theory of the Digital Age*, Cambridge University Press, 2023, 304 s. Egzemplarz recenzyjny od Wydawcy.

sprawa przedmiotowego zakresu tych dyscyplin nie jest jednak specjalnie ważna – lepiej porównywać konkretne dzieła. Zaczynam jednak od tego cytatu z Przedmowy, ponieważ ten impresjonistycznie politologiczny język jest dla książki charakterystyczny. Co do różnicy zaś, sądzę, że jest przeciwnie – filozofia polityczna jest zakresowo szersza od teorii politycznej i dzieli z nią pole badań z naciskiem na to, co ogólne. Teoria polityczna jest raczej dziedziną stosującą kategorie filozoficzne pospołu z własnymi metodami badania przyczyn, przejawów i skutków danego stanu rzeczy.

Rozpatrując wymienione przez siebie różnice między filozofią polityczną i teorią polityczną, Mathias Risse podaje mimochodem definicję procesu, który jego zdaniem określa funkcje tych dyscyplin w naszych czasach, mianowicie przenoszenia myśli politycznej do dziedziny liczb. A przecież tytułowy problem jest nieco inny: „Teoria polityczna w epoce cyfrowej”, czyli jak rozwijać i stosować teorię polityczną, żeby pojmować zdarzenia i inne fakty, które mają czy mogą przybrać postać cyfrową, uwzględniając upowszechnienie stosowania komputerów, powszechny dostęp do internetu, komercyjne i państwowe gromadzenie i przetwarzanie danych oraz – przede wszystkim – rozwój sztucznej inteligencji.

Książka składa się z przedmowy, rozdziału wprowadzającego i dalszych dziesięciu rozdziałów, następnie z epilogu, bibliografii oraz indeksu osobowo-rzeczowego.

Tytuły rozdziałów trafnie informują o ich treści:

1. Wprowadzenie. Cyfrowe Lifeworlds w dziejach ludzkości
2. Ucząc się od amiszów². Filozofia polityczna jako filozofia techniki (*technology*) w stuleciu cyfr
3. Sztuczna inteligencja a przeszłość, teraźniejszość i przyszłość demokracji [zob. niżej]
4. [zob. niżej]
5. Znając i będąc znanym. Badanie epistemicznych uprawnień w światach cyfrowych (*Lifeworlds*)
6. Poza pornografią i dyskredytacją. Obietnice epistemiczne a zgubne skutki techniki *deepfake*

² Amisze – wspólnota protestancka założona w XVII w. przez szwajcarskich menonitów żyjących w Alzacji. W XVIII w. wielka migracja do USA, obecnie ok. 300 tys. osób.

7. Czwarta generacja Praw Człowieka. Prawa epistemiczne w Life 2.0 i Life 3.0

8. O kapitalizmie inwigilacyjnym, władzy instrumentalnej i fizyce społecznej: zapewnianie oświecenia światom cyfrowym (*Digital Liveworlds*)

9. Dane jako fakty społeczne. Sprawiedliwość dystrybucyjna wobec Big Data

10. Bóg, Golem i ubóstwiający Gadżety. Znaczenie życia w świecie cyfrowym

11. Status moralny i uczestnictwo (*membership*) polityczne. Ku teorii politycznej Życia 3.0

Czytelnik z pewnością dostrzeże trzy faktyczne zasadnicze części tekstu:

(1) Streszczenia wszystkich rozdziałów pomieszczone w rozdziale wprowadzającym,

(2) rozdziały 2–11,

(3) krótkie konkluzje poszczególnych rozdziałów w Epilogu.

Jeżeli pojawi się pokusa zestawienia treści części (1) i (3), może zamiast czytania wszystkich rozdziałów, to trzeba ją odrzucić, gdyż dopiero lektura całości ukazuje stanowisko autora.

Artykuł recenzyjny, krótki z natury, nie może referująco ogarnąć wszystkiego. Omówię rozdział 3, który odnosi się do tematu moim zdaniem najważniejszego, rozdział 4, budzący największe wątpliwości, rozdział 9 o sprawiedliwości społecznej i rozdział 11 o przyszłości.

O TREŚCI ROZDZIAŁU 3

Nie mam nic przeciw omawianiu przyszłości jako faktu kulturowego teraźniejszości, metodzie zresztą niedeklarowanej przez autora³. Jednak podział w tytule rozdziału staje pod znakiem zapytania, gdyż oddzielanie przeszłości od teraźniejszości świadczy o takiej teorii politycznej, która chyba nie liczy się z rozwojem metodologii badań historycznych. Jest to język potoczny, oddający potoczną świadomość. Przeszłość także jest kulturowym faktem teraźniejszości i nie zmienia tego fakt, że mówimy o tym z użyciem gramatycznego czasu przeszłego.

³ A. Appadurai, *The Future as Cultural Fact. Essays on the Global Condition*, Verso, 2012. Zob. moją recenzję: „Polski Przegląd Dyplomatyczny” 2017, nr 3, s. 166–170.

Dla przykładu: bitwa pod Płowcami miała miejsce w roku 1331, może nawet więcej niż jedna w owych dniach, ale o tym wyrabiamy sobie zdanie dzięki XX–XXI-wiecznemu odnajdywaniu (archeologia) i oświetlaniu na nowo – i to z wiedzą aktualną – wszystkich znanych dziś (sic) źródeł z tamtej epoki.

Co autor tej książki widzi jako związek sztucznej inteligencji z demokracją, związek wychwytywany i opisywany przez teorię polityczną? W długim rozdziale wstępnym, skupionym na problematyce władzy sprawowanej przez ogół ludności państwa (*people*), autor omawia zagadnienia znane z literatury w sposób z nią zgodny. AI pojawia się w pytaniu o jej znaczenie dla demokracji liberalnej, nam współczesnej. Autor trafnie wiąże demokrację czasów dawniejszych z rozwojem techniki i technologii, a nie z aktualną AI, chociaż początków AI można szukać już w pracach McCullocha i Pittsa⁴, a nawet wcześniejszych – Rashevsky'ego⁵, aktualna zaś przydatność AI dla teoretyków politycznych może polegać tylko na dostępności programu chatGPT openAI.

Autor postuluje wypracowanie sposobu, dzięki któremu AI „spowoduje, że demokracja będzie działać” (*making democracy work*, s. 49). Twierdzi, że nikt tego jeszcze nie proponował. Omawia też specyficznie rzeczową stronę demokracji (*materiality*, materialność rzeczy, przedmiotów, np. urzędzeń do głosowania etc.) w odniesieniu do organizacji wyborów. Tu AI pojawia się jako implikacja narzucająca się czytelnikowi przez domysł. Autor bowiem o tym wyraźnie nie pisze, inaczej niż o kwestii AI jako narzędzia kontrolowania obywateli, nie tylko w powiązaniu z aktem wyborczym. Wymienia Chiny, ale niejeden czytelnik wie o takiej kontroli więcej. Więcej zresztą mówi czytelnikowi nawet sam autor, nakreślając funkcjonowanie AI jako organizatora demokracji przy coraz mniejszym znaczeniu społecznej roli polityka i coraz większym – algorytmów umożliwiających zarządzanie demokracją. Nawet sędzia nie będzie potrzebny, ponieważ AI potrafi lepiej dowody zgromadzić, ocenić i wypracować wyrok, a nawet wprowadzić w życie (?). Autor nazywa to

⁴ W. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, „Bulletin of Mathematical Biophysics” 1943, t. 5, s. 115–137.

⁵ N. Rashevsky, *Physico-mathematical aspects of excitation and conduction in nerves*, „Cold Springs Harbor Symposia on Quantitative Biology. IV. Excitation Phenomena” 1936; *idem*, *Mathematical Biophysics; Physico-Mathematical Foundations of Biology*, University of Chicago Press, 1938.

Grand Democratic AI Utopia i przypomina Orwella. Istotnie, to zakrawa na totalitaryzm „ludu przez lud dla ludu”. Już teraz współczułbym moim wnuczkom w Oksfordzie i wnukom w Warszawie – czyż bowiem moje wyobrażenie ich przyszłości w tej Utopii nie byłoby kulturowym faktem terażniejszości? Jednak do takiej omnipotencji państwa doprowadzić nie powinniśmy – my, ludzie trzeciej dekady XXI w. A dlaczego tak dokładnie i blisko to datuję? Dlatego, że przełom w zakresie AI może nastąpić już wkrótce, tworząc groźną dla ludzkości osobliwość w postaci pojawienia się superinteligencji i nieodwracalnych tego skutków.

Autor zarzuca teoretykom politycznym, że traktują demokrację jak ideał albo ramy instytucjonalne, zamiast rozważać jej *materiality*, czyli przedmioty (*objects*) kształtujące w pewnym stopniu jej stosowanie. „AI zmienia proces podejmowania decyzji i jak działają w nim ludzie” (w prezentacji rozdziału, s. XIX). Nie sądzę jednak, że ma sens nazywanie sojusznikami – i to „naturalnymi sojusznikami” – demokracji i techniki (technologii). Czy taka antropomorfizacja ma sens? Autor odwołuje się do licznych opracowań, żeby ukazać ewentualną funkcję AI, która zna opinie wszystkich i udostępnia te informacje, czyniąc obywateli kompetentnymi uczestnikami wymiany poglądów w procesie sprawowania władzy (*governance*). Monitorowanie umożliwi AI wychwycenie oszustwa i korupcji, identyfikuje fałszywe informacje o tym procesie, a także np. ukazuje przeszkody w docieraniu do punktów głosowania. A kiedy do wyborów ma w ogóle dochodzić? Otóż wtedy, gdy AI stwierdzi spadek zaufania do rządu poniżej pewnego poziomu. Wynik wyborów zależy nie tylko od liczby głosów oddanych na danego polityka z listy, lecz i od stanowiska głosującej osoby w rozmaitych sprawach publicznych, które będzie znane AI. Danych osobowych AI nie udostępni publicznie, ale je wykorzysta. Taki system sprawi jednak, że sama AI stanie się częścią *governance*. Autor twierdzi, że taki kierunek rozwoju jest jednak niebezpieczny. Zamiast uczynić człowieka jako obywatela lepszym, ulepszylibyśmy system AI i wzmocnilibyśmy państwowe możliwości nadużywania posiadanych danych. Orwell jest przywoływany ponownie. Nie należy więc myśleć o jakiejś Grand Democratic AI Utopia – zamiast tego należy planować sposób wykorzystania AI do udoskonalenia liberalnej demokracji reprezentatywnej.

O tym traktuje kolejna część tego rozdziału, zatytułowana „AI and democracy. Public spheres”. Autor powołuje się na Habermasa

i innych, podnosząc problem zmiany w sferze publicznej. Pojawia się zatem problem „czwartej władzy”, prasy i innych środków informacji, łącznie z mediami „społecznościowymi”, ich *deepfakes* itp. Autor postuluje wyeliminowanie zjawisk szkodliwych, chciałby systemu zgodnego z ideałem demokracji liberalnej. Nie podaje jednak nowego sposobu na zbudowanie go przez obywateli. Nietracący nadziei czytelnik przystępuje do czytania ostatnich części tego rozdziału. Autor przypomina niektóre negatywne zjawiska z dziejów koncentracji władzy w kompleksie wojskowo-przemysłowo-(kontr)wywiadowczym USA. Przedstawia też jednak przykład pożytecznego wykorzystania AI w zakresie funkcjonowania demokracji w Australii (eDemocracy project, Open Forum) i USA (California Report Card), s. 68–69. Nie odnosi się do programu chatGPT i podobnych.

W części tego samego rozdziału, zatytułowanej „AI and Democracy. Economic Power”, autor postuluje, żeby okiełznać (*harness*) wielkie korporacje. Chciałby też powściągnąć technologię, co jest postulatem jeszcze mniej realistycznym. Autor wyraża obawy, czy ludzkość będzie w stanie kontrolować swój przyszły świat. Nie jest pierwszym, który tak alarmuje, ale argumentuje przekonująco.

O TREŚCI ROZDZIAŁU 4

Co do samego prawa do prawdy, autor zdaje sobie sprawę z trudności związanych z przyjęciem takiej kategorii. Powołuje się na Hohfelda, który w 1919 r. zdefiniował w odniesieniu do jednostki przywileje, uprawnienia (*claims*), kompetencje (*powers*) i immunitety⁶. ONZ ustanowiła Dzień Prawa do Prawdy, UNESCO ma monitorować publiczny dostęp do informacji (prawdziwej) itd. Te postanowienia odnoszą się oczywiście do specyficznych kontekstów, ale fraza „prawo do prawdy” jest generalizacją. Co powiedziałby o takim prawie Alfred Tarski, człowiek, który zdefiniował prawdę? W rozważaniach Rissego o prawie do prawdy znalazłem tylko jedno użycie terminu „błąd”, i to w opisie stosunku nauczyciela do ucznia (s. 81). A przecież głoszenie nieprawdy jest szczególnie niebezpieczne, gdy jest intencjonalne, tzn. kiedy jest kłamstwem. Błąd nie jest prawdą, ale też nie jest zamierzony, nie jest kłamstwem. Brakuje mi tego aspektu „prawa do prawdy” w czytanych

⁶ W. Hohfeld, *Fundamental Legal Conceptions*, red. W. Cook, Yale University Press, 1919.

rozważaniach. Nadrabia to częściowo obszernie potraktowany wątek o stosunku Nietzschego do prawdy. W końcu autor także zajmuje się argumentem przeciw tezie o istnieniu prawa do prawdy, nie odrzucając jej, lecz „temperując”. Trafnie wskazuje na ekstremistyczne postulowanie prawdy jedynej w wojnach religijnych.

Wszystko to dotyczy życia w obecnej fazie 2.0. Zgodnie z akceptowaną przez autora klasyfikacją ludzkość znajduje się właśnie w fazie 2.0. Nie wiadomo, czy faza 3.0 zaistnieje, a do fazy 1.0 nie ma już powrotu.

Rozważania o prawie do prawdy kończy w tym rozdziale część zacytowana „Public Reason and the Right to Truth”. Autor wielokrotnie powołuje się na Johna Rawlsa, więc przyjmuję, że to przede wszystkim od niego zaczerpnął pojęcie *public reason*. „Rozum publiczny”, jeśli klasycznego Rawlsa dobrze streszczam, przejawia się wtedy, gdy stanowisko w określonej sprawie jest – czy może być – podzielane przez zróżnicowany pod względem przekonań ogół obywateli w wyniku racjonalnego namysłu jednostek taki ogół tworzących. Czy to stanowisko ma wyrażać prawdę? Powiedziałbym, że raczej prawdziwe porozumienie co do wspólnego interesu prawdziwego.

Autor także stawia sprawę realistycznie: „Prawdomówność (*truthfulness*) ma wielkie znaczenie dla funkcjonowania sfery publicznej. *Public reason* może i musi to uznawać. Lepiej chyba jest jednak uznać jego znaczenie, mówiąc, że te sprawy mają wielką wartość i wymagają żywej ochrony prawnej, zamiast wyrażać to znaczenie w kategoriach praw (*rights*)” (s. 94). I dalej: „Znalezienie właściwego miejsca (*right place*) dla prawdy w organizacji ludzkiego życia” (*ibidem*) jest dobrą odpowiedzią na nieszczęścia wynikające z odrzucenia prawdy, pilnie potrzebną w świecie cyfrowym naszej przyszłości. Takie stanowisko zasadniczo rozprasza wątpliwości, które mógł mieć czytelnik rozważań o prawie do prawdy⁷.

O TREŚCI ROZDZIAŁU 9

Problematyka sprawiedliwości społecznej należy do klasycznych przedmiotów badań filozofii politycznej i teorii politycznej. Może to interesować polskiego czytelnika także z powodu art. 2 naszej Konstytucji:

⁷ Błądziłby jednak, kto by głosił, że „historyczna rzetelność jest tylko wtedy, gdy jest zgodna z naszym zbiorowym, wspólnotowym oczekiwaniem”. Myliłby prawdę z rozumem publicznym.

„Rzeczpospolita Polska jest demokratycznym państwem prawnym, urzeczywistniającym zasady sprawiedliwości społecznej”.

Rozumowanie autora jest proste: sprawiedliwość dystrybucyjna odnosi się do podziału dóbr, zwłaszcza dóbr powszechnego użytku. Takim dobrem są także dane (*data* wszelkiego rodzaju). Właścicielami danych są obecnie ci, którzy je gromadzą – wielkie korporacje. Ponieważ jednak dane odnoszą się praktycznie do wszystkich i są w szerokim zakresie przez wszystkich wytwarzane, przeto kontrola nad nimi powinna być równie powszechna. Kontrola w zakresie danych jest zaś – prawnie dokładniej rzecz pojmując – składnikiem własności tych danych. Dane winny być traktowane jako własność wspólna. Planeta Ziemia też jest wspólna, choć niesprawiedliwie eksploatowana – to wynika z całego rozumowania w tej książce. Autor nie prowadzi myśli dalej, a przecież można by, podążając za tym rozumowaniem, zauważyć, że przecież i praca jest powszechna, jej wyniki powinny być więc wspólne, prawda? A jednak nie są, gdyż system opiera się na procesie akumulacji kapitału. I ten proces decyduje teraz także o badaniach AI, o zbieraniu wszelkich danych i ich przetwarzaniu w celu handlowego wykorzystania. Skracając potencjalnie pełne logiczne rozumowanie, można stwierdzić, że przyszłe funkcjonowanie AI zakwestionuje – unicestwi? – zasady funkcjonowania kapitalizmu w obecnym jego kształcie. Sprawiedliwość dystrybucyjna nie ma jednak definicji absolutnej, zależy od pojmowania egalitaryzmu. Dlaczego AI miałyby działać zgodnie z określoną zasadą egalitarnej dystrybucji? Będzie działać zgodnie z wolą projektującego i wdrażającego ją człowieka, jeśli człowiek utrzyma nieustanną kontrolę nad jej rozwojem, zaś takiej pewności nie ma. Zakładając, że teraz adekwatna kontrola istnieje, czy może się ona utrzymać?

O TREŚCI ROZDZIAŁU 11

Rozdział ten odnosi się do przyszłości niepewnej podwójnie – najpierw jako przyszłość, coś z zasady nieznanego, a po wtóre jako przyszłość niepewna co do samego istnienia gatunku ludzkiego, wspomnianego Life 3.0.

Klasyfikacja Life 1.0, 2.0, 3.0 nie została w pełni objaśniona. Życie 2.0 oznacza zdolność jednostek (*entities*, tutaj odnoszonej do ludzi) do zmiany swego położenia kulturowego (uczenie się indywidualne

i zbiorowe), lecz nie swego kształtu fizycznego (s. 3). Jednostki (*entities*) istniejące w Życiu 1.0 nie są w stanie uczyć się i od początku znajdują się w ustalonym stanie pojmowania świata i niezmiennym kształcie fizycznym. Z lektury książki Tegmarka możemy się domyślać, że ten *physical shape* to nie tyle kształt co stan fizyczny⁸. Zastanawia mnie, kiedy zdaniem Mathiasa Rissego ludzkość znajdowała się w stanie Life 1.0. Czy może odnosi się ono do *entities* innych niż ludzie? Myśl ta wynika m.in. z faktu, że Life (1.0 itd.) jest skrótem terminu Lifeworld. A tu znów mamy komplikację, ponieważ pojęcie to pochodzi od Husserla i oznacza świat doświadczany przez człowieka potocznie, zwyczajnie, codziennie. Zatem tłumaczenie Life 1.0 jako Życie 1.0 nie jest wystarczająco jasne. W omawianym rozdziale mamy jednak do czynienia głównie z problematyką Life 3.0. Ta faza życia nie odnosi się do życia w potocznym pojmowaniu, lecz – czytamy u Tegmarka – do stanu wykraczającego poza obecny stan ewolucji życia na Ziemi, do „życia” w hipotetycznym przyszłym świecie AI, która prześcignęła już inteligencję ludzką. Autor recenzowanej książki sądzi, że przyszły komputer mający świadomość – i to praktycznie nieograniczoną – będzie znał kryteria moralne i będzie mógł być podług nich osądzany przez człowieka. Jednocześnie AI będzie mogła mieć zdolność do działania według swojego programu, może sprzecznego ze zdolnością i programem człowieka (ludzkości). Dalsza komplikacja może się pojawić wtedy, gdy AI będzie funkcjonować w sposób nieprzewidywalny przez nią samą i niesterowalny przez człowieka. Istnieją granice uznawania robota za niewolnika dlatego tylko, że został stworzony przez człowieka. Czy dzieci są niewolnikami? Kluczem jest oczywiście tworzenie maszyn posiadających świadomość. Czy takie urządzenia mogą, czy nawet muszą wymknąć się spod kontroli ludzkiej? Na pytanie o teorię polityczną – stale obecne w tej książce – autor odpowiada przykładami prac Heideggera i Mumforda (technologia formuje świat), Wienera (cybernetyka), Verbeeka (moralność rzeczy), Lovelocka (Novacen, era hiperinteligencji).

Autor wprowadza interesujący wątek stosunku człowiek–zwierzę, szukając analogii do ewentualnej relacji między człowiekiem a urządzeniem wyposażonym w pewne funkcje mózgu.

⁸ Autor czerpie jednak z książki Maxa Tegmarka, *Life 3.0. Being Human in the Age of Artificial Intelligence*, Knopf, 2017. Moje rozumienie tej klasyfikacji zakłada m.in., że stosowanie lekarstwa zmienia ten kształt czy stan fizyczny.

Kluczowe jest oczywiście zagadnienie świadomości. Ktoś pytał, czy posiadają ją „owoce morza” żywcem gotowane w restauracji? One do końca usiłują wydostać się z wrzątku. W Życiu 3.0 może już chodzić nie tyle o komputer, co o system AI. Zastanawiam się, czy warto pytać o teorię polityczną, jeżeli hipotezy odnośnie do AI nasuwają pytanie o pewność dalszego trwania ludzkości złożonej z osobników świadomie siebie formujących i regulujących swoje otoczenie przyrodnicze. Autor wysuwa na plan pierwszy problem racjonalności. Czy racjonalność przewyższa inteligencję i samoświadomość człowieka, będąc ich produktem? Autor przypuszcza, że AI nie może być racjonalna, a w działalności politycznej racjonalność jest niezbędna (s. 245). Ale na czym będzie polegała polityka w epoce superinteligencji?

Wiemy, że zwierzęta, choć bez racjonalności (na pewno?), mogą żyć w symbiozie z człowiekiem racjonalnym (np. niewidomy korzysta z pomocy psa). Może więc AI – z definicji inteligentna – da się w roli inteligentnego psa obsadzić. Teoria polityczna objęłaby także sztuczną inteligencję. Rodzi się jednak pytanie, czy urządzenia AI mogłyby osiągnąć status podobny do obywatelskiego, może nawet taki sam? Sam fakt, że takie pytanie pada, może nam uświadomić, jak bardzo nieprzygotowani jesteśmy na nowe w zakresie AI. Przypomina to pytania Hume’a i Kanta, czy racjonalność ustala i utrwała wartości. Hume uważał, że nie. Każdy inteligentny, świadomy i racjonalny osobnik może sobie stawiać każdy cel. Natomiast Kant sądził, że z racjonalności wypływa moralność. Nie każdy cel jest dopuszczalny. Ponieważ jednak człowiek dopuszcza się niemoralnych czynów wobec bliźniego, to najwyraźniej wartości nie są formowane wyłącznie przez racjonalność kantowską. Nie wystarczy też prosta odpowiedź, że przecież istnieją też emocje, istnieją interesy. Interesy zaś bez wartości nie istnieją. Nie jest trafny argument, że interesy wynikają z „bezinteresownej” inteligencji. Inteligencja, świadomość, racjonalność są przejawami relatywnie zintegrowanej umysłowości, osobowości. To zaś wykracza poza problematykę teorii politycznej.

Recenzowana książka raczej pobudza do zadawania pytań niż oferuje odpowiedzi. Wydaje się, że rozwój teorii politycznej odnoszącej się do Life 3.0 nie zależy wyłącznie od filozofów czy teoretyków, lecz także od badaczy i projektantów w zakresie AI. Można stosować teorię polityczną do państwa, ponieważ wiemy, jak przy pomocy rozmaitych teorii, nie tylko politycznej, instytucję państwa badać i wynikami takich badań

dysponujemy na potrzeby teorii politycznej. Rozwijanie (stosowanie) teorii politycznej hipotetycznych instytucji w Życiu 3.0 musi przebiegać w ścisłym związku z adekwatnym kontrolowaniem projektowania urządzeń i podsystemów oraz funkcjonowania systemu AI w trakcie ich wytwarzania.

Mathias Risse opublikował książkę, która – powtarzam – pobudza do stawiania pytań i rozwijania rozmaitych składników teorii politycznej świata epoki sztucznej inteligencji, a to może doprowadzić do sformułowania adekwatnej teorii politycznej.

Podstawowe pytanie odnosi się do współdziałania międzypaństwowego i państw z prywatnymi przedsiębiorcami w zakresie budowania i kontrolowania światowej sieci AI. Taka sieć może powstać niezależnie od zamiarów projektantów urządzeń i programów AI, zatem i woli rządów oraz działających prywatnie przedsiębiorców. Może powstać przynajmniej z dwóch powodów: (1) może zaistnieć program w zakresie AI z właściwościami nieprzewidywanymi na etapie projektowania oraz inicjalnego wdrażania, a pośród takich właściwości mogą się pojawić niebezpieczne dla rodzaju ludzkiego; (2) w obecnym stanie teorii politycznej i stosunków międzypaństwowych i rynkowych (w tym i relacji między producentami towarów i oferentami usług na światowych rynkach i w relacjach z rządami) zaistnienie skutecznego systemu powszechnej kontroli nad pracami w zakresie AI może być niemożliwe i choćby z tego powodu nie zaistnieć wcale. Zakaz badań w zakresie AI też nie jest egzekwowalny, chyba że się to będzie wdrażać siłą. Wtedy jednak groźba ze strony AI może się nagle stać drugorzędna⁹.

Teoretycy polityki mogą opracować rozmaite scenariusze działań, których celem będzie uzyskanie pożądanego stanu, dla konkretnych rządów, organizacji międzyrządowych na czele z ONZ oraz najpotężniejszych przedsiębiorców i najważniejszych ośrodków naukowych. Nie chodzi o jakiś absurdalny zakaz rywalizacji, chodzi o to, żeby przebiegała ona na podstawie zasad współdziałania w celu osiągnięcia przez ludzkość

⁹ Zob. R. Stemplowski, recenzja: Ch. Miller, *Chip War: The Fight for the World's Most Critical Technology*, Simon&Schuster, 2022, w: *idem*, *Lektura uzupełniająca. Notes analityka polityki*, Wydawnictwo Wolski, 2023, także moje zwięzłe omówienia książek: N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2017; P. Norvig, S. Russell, *Artificial Intelligence: A Modern Approach*, 4th ed., Pearson, 2022; S.J. Frantzman, *Drone Wars: Pioneers, Killing Machine, Artificial Intelligence, and the Battle for the Future*, Post Hill Press, 2021, *ibidem*.

wspólnego celu. Nawet zdefiniowanie takiego stanu, cóż dopiero jego wdrożenie, jest problemem. Mowa wszak o zasadach ustrojowych, instrukcjach administracyjnych i wojskowo-policyjnych, hierarchii finansowanych przez państwa tematów badawczych, mowa o regulaminach pracy zarobkowej oraz ogólnych wskazówkach zachowań obywatelskich i o bardzo wielu innych sprawach. Wobec takiego wyzwania, jakie stwarza AI, ludzkość jeszcze nie stawała. Wymaga to niezwłocznego zwiększenia wysiłku teoretyków polityki. Nie twierdzę, że samotnie zbawią oni świat, ani nawet że okażą się w tych wysiłkach najbardziej wpływowi, ale pytam, co może być obecnie dla nich ważniejszego.

28 kwietnia 2023 r.