

O wpływie dokonań w dziedzinie sztucznej inteligencji na współczesny światopogląd

1.

Jedną z propozycji filozoficzno-światopoglądowych, która odpowiada na współczesne potrzeby i wyzwania społeczeństwa informacyjnego, a szerzej cywilizacji informacyjnej, jest rozwijana przez Pawła Stacewicza i Witolda Marciszewskiego, od roku 2011, koncepcja *światopoglądu informatycznego*. Zgodnie z określeniem tego pierwszego, jest to: „...pewien typ poglądów przedfilozoficznych, które są ugruntowane naukowo, a wspierają się na założeniu, że bardzo istotną rolę w opisie świata i relacji człowiek-świat odgrywają pojęcia informatyczne, jak dane, kod, algorytm czy obliczalność”. Koncepcja ta ma status abstrakcyjnego konstruktów, który może kształtować, na zasadzie *teoretycznego wzorca*, realne światopoglądy poszczególnych ludzi.

(Marciszewski, Stacewicz 2011), (Stacewicz 2015), (Stacewicz 2016), (Stacewicz 2017)

2.

Określony wyżej konstrukt obejmuje dwie warstwy: 1) warstwę opisu świata, czyli *światooglądu* – budowanego za pomocą aparatury pojęciowej informatyki teoretycznej, 2) właściwą warstwę *światopoglądową* – złożoną z przekonań, do których prowadzą pytania formułowane w ramach informatycznego obrazu świata.

Do tez konstytuujących *światoogląd* należą: 1a) Każdy byt mieści w sobie pewną *informatyczną zawartość*, która w połączeniu z substratem materialnym decyduje o jego specyfice (budowie, funkcjach, oddziaływaniach z otoczeniem etc...), 1b) Organizmy, umysły i cywilizacyjne artefakty (np. organizacje) są *systemami przetwarzającymi informacje*, które można modelować za pomocą systemów informatycznych, 1c) Problemy, przed którymi staje ludzki umysł – zarówno w wymiarze indywidualnym, jak i zbiorowym (cywilizacyjnym) – mają zróżnicowaną *złożoność*, którą można określić za pomocą metod informatycznych.

Wśród pytań prowadzących do określonych *typów poglądów* (warstwa 2) można wyróżnić następujące: 2a) Czy zjawiska występujące w świecie fizycznym, m.in. w ludzkim mózgu, mają charakter cyfrowy, czy analogowy?; 2b) Czy ludzki umysł/mózg jest układem *silniejszym obliczeniowo* od komputerów cyfrowych?; 3b) Czy doskonalone przez umysł strategie rozwiązywania problemów, w tym obliczeniowe, są w stanie nadążyć za *rosnącą złożonością* problemów w świecie?

(Stacewicz 2016), (Stacewicz 2019a, 2019b), (Stacewicz 2020a, 2020b)

3.

Na współczesny światopogląd, w tym informatyczny, mają niewątpliwy wpływ dokonania w zakresie *sztucznej inteligencji* (SI) – czyli tego obszaru badań informatycznych, które zmierzają do algorytmizacji i automatyzacji procesów poznawczych człowieka (takich jak percepcja, wnioskowanie, operowanie językiem, uczenie się). Dokonania te – poprzez swój praktyczny sukces, wykraczający poza modelowanie i realizację czynności umysłowych – wzmacniają informatyczny obraz świata; ponadto zaś prowadzą do nowego rodzaju pytań o umysł, świat i wytwarzane przez umysł artefakty.

(Russel, Norvig 2020), (Stacewicz 2010), (Stacewicz, Włodarczyk 2011), (Stacewicz 2021)

4.

Na współczesny światopogląd wpływają szczególnie mocno następujące pytania: 1) Czy procesy stojące za ludzką inteligencją (a szerzej myśleniem) mają charakter *algorytmiczny*? – gdyby tak było, byłibyśmy w stanie skonstruować byty inteligentne/myślące, co prowadziłoby do szeregu problemów etycznych, 2) Czy można stworzyć systemy SI, wykazujące *autonomię* na poziomie rozwiązywania problemów i działania?, 3) Czy systemy SI, w tym autonomiczne, mogą stanowić *zagrożenie* dla naszej cywilizacji (egzystencjalne, społeczne, gospodarcze...)?, 4) W jakim zakresie systemy SI mogą stać się skutecznym *narzędziem kontroli* decyzyjnych elit nad obywatelem?

W toku dyskusji nad referatem sformułowałem pogląd, że największe znaczenie światopoglądowe ma zagadnienie hipotetycznej *autonomii maszyn* (również w kontekście możliwych zagrożeń dla ludzkości). Analizując to zagadnienie, należy rozróżnić autonomię na poziomie metod osiągania pewnych celów, autonomię na poziomie celów oraz autonomię na poziomie kryteriów wyboru celów. Współczesne systemy SI wykazują autonomię najniższego stopnia, co znaczy, że są one w stanie dochodzić do oryginalnych, nieznanych człowiekowi, rozwiązań zadawanych im problemów (np. w drodze obliczeń ewolucyjnych). Jeśli chodzi o dwa wyższe poziomy autonomii, sprawa ich osiągnięcia przez systemy SI pozostaje otwarta.

(Stacewicz 2018), (Stacewicz 2019a)

5.

Stawiane w trakcie referatu i dyskusji kwestie są dyskutowane we współprowadzonym przez Pawła Stacewicza blogu dyskusyjnym *Cafe Aleph* (www.marciszewski.eu), który obejmuje ponad 120 tematycznych dyskusji i ponad 1200 obszernych komentarzy blogowych. Komentarze te mogą stanowić ciekawy materiał empiryczny dla przeprowadzanych w ramach projektu badań nad zmianami światopoglądowymi, którym sprzyjają dokonania w zakresie informatyki, w tym sztucznej inteligencji.

(blog akademicki *Cafe Aleph*), (Stacewicz 2021)

Źródła:

Marciszewski W., Stacewicz P., (2011), *Umysł-Komputer-Świat. O zagadce umysłu z informatycznego punktu widzenia*, Warszawa, Akademicka Oficyna Wydawnicza EXIT.

Russel S., Norvig P., (2020), *Artificial Intelligence: A Modern Approach*, Berkeley.

Stacewicz P., (2010), *Umysł a modele maszyn uczących się. Współczesne badania informatyczne w oczach filozofa*, Warszawa, Akademicka Oficyna Wydawnicza EXIT.

Stacewicz P., Włodarczyk A., (2011), *O modelowaniu informatycznym, ze szczególnym odniesieniem do badań nad sztuczną inteligencją*, *Zagadnienia Naukoznawstwa*, 4, s. 481-500.

Stacewicz P. (red.), (2015), *Informatyka a filozofia. Od informatyki i jej zastosowań do światopoglądu informatycznego*, Warszawa, Oficyna Wydawnicza Politechniki Warszawskiej.

Stacewicz P., (2016), *Informational Worldview. Scientific Foundations and Philosophical Perspectives*, *Studies in Logic, Grammar and Rhetoric*, 2016, nr 47(60), s. 35-47.

Stacewicz P., (2017), *Światopogląd informatyczny. Optymistyczna czy pesymistyczna wizja świata*, [w:] „Spory o wartości: aspekty filozoficzne i administracyjno-prawne”, Zubelewicz Jan (red.), Warszawa, Oficyna Wydawnicza Politechniki Warszawskiej, s.73-88.

Stacewicz P. (red.), (2017), *Różne oblicza informacji*, Warszawa, Oficyna Wydawnicza Politechniki Warszawskiej.

Stacewicz P., (2018), *O możliwości stworzenia maszyn autonomicznych*, [w:] „Studia z filozofii informatyki”, Stacewicz Paweł, Sołoduha Krzysztof (red.), 2018, Warszawa, Wojskowa Akademia Techniczna im. Jarosława Dąbrowskiego, s. 145-171.

Stacewicz P., (2019a), *From Computer Science to the Informational Worldview. Philosophical Interpretations of Some Computer Science Concepts*, *Foundations of Computing & Decision Sciences*, 2019, 44(1), s. 27-43.

Stacewicz P., (2019b), *The Informational Worldview and Conceptual Apparatus* [w:] *Interdisciplinary Investigations into the Lvov-Warsaw School*, Drabarek A., Woleński J., Radzki M. M. (red.), 2019, Szwajcaria, Palgrave Macmillan, s. 259-280.

Stacewicz P., (2020a), *Analogicity in Computer Science. Methodological Analysis*, *Studies in Logic, Grammar and Rhetoric*, 2020, 63(1), s. 69-86.

Stacewicz P., (2020b), *O znaczeniu pojęć informatycznych dla filozofii na przykładzie rozróżnienia między cyfrowością i analogowością*, *Filozofia i Nauka. Studia filozoficzne i interdyscyplinarne*, 2020, 8(1), s. 213-233.



Projekt finansowany z budżetu państwa w ramach programu Ministra Edukacji i Nauki pod nazwą: „Nauka dla społeczeństwa”
Tytuł projektu: *Architektura i infrastruktura Wielkiego Resetu – studium transformacji ku cywilizacji informacyjnej*

Stacewicz P., (2021), *Filozofia, informatyka, sztuczna inteligencja... Wybrane głosy z bloga akademickiego Cafe Aleph*, [w:] *Ewolucja cywilizacyjnej roli i społecznego odbioru nauki*, Lubacz Józef (red.), vol. LXXIV, Warszawa, Wydawnictwo SGGW, Instytut Problemów Współczesnej Cywilizacji im Marka Dietricha, s. 97-109.