

ŁUKASZ IWASIŃSKI



## BIG DATA A PROBLEM REPREZENTACJI POZNAWCZEJ

**ABSTRACT.** Łukasz Iwasiński, *Big data a problem reprezentacji poznawczej* [Big Data and the problem of mental representation] edited by Andrzej Wawrzynowicz, „Człowiek i Społeczeństwo” vol. LIII: *Paradygmaty metafizyki* [Paradigms of metaphysics], Poznań 2022, pp. 241–261, Adam Mickiewicz University. ISSN 0239–3271, <https://doi.org/10.14746/cis.2022.53.14>.

In this article, the author reflects on Big Data analytics in the context of the problem of cognitive representation. There are many voices declaring that the era of Big Data has brought a radical breakthrough in human cognitive abilities. Some – especially in the world of business and marketing, and to a lesser extent in the field of science – argue that for the first time we can reach a clean, objective picture of reality and keep track of its changes. The article is a critical commentary to this thesis. In Big Data analytics, cognitive activities are assessed not from the point of view of their compliance with reality, but the possibility of achieving set goals. Big Data mining can be, and often is, an important tool for reality control and forecasting – which does not mean it can discover objective truth and create accurate representations of reality.

**Keywords:** cognitive representation, Big Data, anti-realism, constructivism

Łukasz Iwasiński, Uniwersytet Warszawski, Wydział Dziennikarstwa, Informacji i Bibliologii, 00-927 Warszawa, ul. Nowy Świat 69, e-mail: [lukiwas@gmail.com](mailto:lukiwas@gmail.com), ORCID: <https://orcid.org/0000-0003-2126-7735>.

### Reprezentacja poznawcza

Problem reprezentacji poznawczej jest jednym z centralnych zagadnień nowożytnej refleksji filozoficznej. Ma on fundamentalne znaczenie nie tylko dla epistemologii, ale także zagadnień z dziedziny naukoznawstwa, filozofii oraz socjologii wiedzy i nauki. Jak pisze Marek Sikora, pojęcie reprezentacji poznawczej wyłoniło się w wyniku sporu między prezentacjonizmem

a reprezentacjonizmem. Orędownicy pierwszego stanowiska uznawali, że „podmiot w poznaniu postrzega obiekty rzeczywiste”, a drugiego, że „przedmiotem poznania nie są obiekty rzeczywiste, lecz ich obrazy lub skutki, tj. pewne reprezentacje tych obiektów” (Sikora, 2007: 9). W obrębie samego reprezentacjonizmu od początku rysowały się konflikty, przede wszystkim między realizmem a idealizmem subiektywnym. Zwolennicy pierwszego kierunku twierdzili, że przedmiotem poznania są reprezentacje realnych bytów, niezależne od poznającego podmiotu, a drugiego – że reprezentacje poznawanych bytów nie są niezależne od podmiotu i aktu poznania, że w procesie poznania nie istnieje nic zewnętrznego wobec samego procesu poznania. Z czasem rozwinęto wiele bardziej zniuansowanych wersji ujęć problemu reprezentacji. Główne kontrowersje we współczesnej refleksji nad reprezentacją poznawczą to, zdaniem Sikory (2007: 10–12): (1) spór między realizmem i antyrealizmem, będący kontynuacją, ale też pogłębieniem, wspomnianego konfliktu między realizmem a idealizmem subiektywnym, oraz (2) spór między realizmem a konstruktywizmem. W pierwszym przypadku chodzi o „status poznawczy wytworów poznania naukowego”: realiści przyjmują, że owe wytwory są konstruktami, lecz reprezentują realny świat, odzwierciedlają jego kształt, strukturę, antyrealiści natomiast uznają, że wytwory poznania status prawdziwości otrzymują nie dlatego, że są zgodne z rzeczywistością, ale dlatego, że są zgodne z przyjętą ideą prawdziwości. Spór drugi koncentruje się nie tyle na wytworach poznania, ile na procesie poznania: realiści uznają, że poznawana rzeczywistość jest zewnętrzną wobec poznającego podmiotu i poprzedza akt poznania, a konstruktywiści utrzymują, że podmiotu i przedmiotu poznania nie da się rozdzielić. Uogólniając, można przyjąć, że na jednym biegunie debaty znajduje się pogląd realistyczny, zakładający, że wytwory poznania są obrazem prawdziwego świata, a ów świat jest niezależny od podmiotu poznania, który w procesie poznania go odkrywa. Co za tym idzie, status prawdziwości twierdzeń o świecie zależy wyłącznie od jego, poprzedzającej proces poznania, natury. Na drugim biegunie są stanowiska antyrealistyczne i konstruktywistyczne, uznające, że wytwory poznania są traktowane jako prawdziwe na podstawie określonych rozstrzygnięć o charakterze społecznym, będących produktem negocjacji, w które uwikłane są gry interesów, określone wartości i ideologie, kształt instytucji społecznych itd. Podmiot współkształtuje tu przedmiot swego poznania (por. Zybertowicz, 1995: 58–65)<sup>1</sup>.

---

<sup>1</sup> Taka klasyfikacja prezentuje stanowiska skrajne. Rozwój dyskusji dotyczącej reprezentacji poznawczej prowadzi do tego, że przedstawione tu podziały u wielu autorów,

Realizm w myśleniu zdroworozsądkowym, potocznym, może wydawać się oczywisty i jest domyślnie przyjmowany przez dużą część środowiska naukowego, które wierzy, że wykonując podporządkowane odpowiedniej, rygorystycznej metodologii operacje, zdoła odkryć obiektywne fakty oraz jednoznacznie zrekonstruować prawidłowości i właściwości świata, a tym samym stworzyć jego adekwatny obraz (a przynajmniej, w toku rozwoju nauki, coraz bardziej się do tego celu przybliżać). Współczesna filozofia oraz socjologia wiedzy i nauki skłaniają się jednak w stronę antyrealizmu/konstruktywizmu, bądź też problematyzują i komplikują stanowisko realistyczne.

### Źródła krytyki realizmu

Zasadnicze znaczenie dla sproblematyzowania realizmu miała koncepcja Immanuela Kanta, jego „przełom kopernikański” (1957: 26–33; 1960: 65–67). Stworzył on grunt dla poglądów negujących niezależność przedmiotu i podmiotu poznania. W przeważającej liczbie koncepcji realizmu przedmiot poznania jest wtórny wobec samego procesu poznania i poznającego podmiotu. Z myśli filozofa z Królewca wypływa wniosek, że dostępna poznającemu odbiorcy wizja rzeczywistości jest projekcją rzutowaną na aprioryczne kategorie umysłu, które ową rzeczywistość (przedmiot poznania) porządkują i konstytuują. Idee Kanta istotnie przyczyniły się do rozwoju sceptycyzmu w zakresie możliwości dotarcia do niezależnego od człowieka kształtu rzeczywistości. Sikora wprost nazywa Kanta pionierem myśli konstruktywistycznej (Sikora, 2007: 22). Także Jerzy Szacki (2006: 122) u Kanta szuka źródeł krytyki realizmu – od niego wywodzi problem, „czy rzeczywistość społeczna [choć możemy spytać o rzeczywistość w ogóle – dop. Ł.I.] jest nam dana jako zbiór gotowych «rzeczy» czy też jest tak lub inaczej «konstruowana»”. Z kolei Jan Woleński (2005: 468) interpretuje Kanta jako epistemologicznego antyrealistę (a zarazem realistę ontologicznego).

Na przełomie XIX i XX w. konwencjoniści (Henri Poincaré, Pierre Duhem, Édouard LeRoy)<sup>2</sup> argumentowali, że dane empiryczne nie są wystarczające dla ustalenia prawdziwości sądu. Nie istnieją nagie fakty, dane

---

w wielu punktach, ulegają zatarciu; komplikuje się też kwestia samych definicji (zob. Krajewski, 1995; Zybortowicz, 1995; Czarnocka, 2003; Majcherek, 2004).

<sup>2</sup> W nawiasach podaję najbardziej paradygmatycznych przedstawicieli (ewentualnie ośrodki naukowe) danego nurtu.

niezawierające uprzednich założeń o charakterze czysto umownym. Same sposoby gromadzenia danych, konstrukcji wskaźników, narzędzi pomiarowych i innych składowych procesu poznania stanowią konwencje, mające istotny wpływ na kształt owego sądu (zob. Kołakowski, 1966: 145–165; Heller, 2009: 30–31). Zgodnie z Popperowskim postpozytywizmem z lat trzydziestych XX w. obserwacje zawsze są interpretacjami w świetle teorii (Popper, 1977: 91; jakkolwiek jego stanowisko w tej kwestii nie było konsekwentne, zob. Chmielewski, 1995: 174–175; Majcherek, 2004: 56–57). Dziedzictwem postpozytywizmu jest też radykalizowanie konwencjonalistycznego stanowiska o niedookreśleniu/niezdeteminowaniu twierdzeń przez dane (Willard Van Orman Quine), w myśl którego z danego zbioru danych można wyprowadzić różne interpretacje, zbudować na ich podstawie różną wiedzę i tym samym różne reprezentacje świata (zob. Chmielewski, 1997: 211–218; Sikora, 2007: 99–100).

Silny asumpt do krytyki realizmu dały w latach sześćdziesiątych ubiegłego stulecia poststrukturalizm i dekonstrukcjonizm. W poststrukturalistycznej koncepcji poznania Michela Foucaulta wiedza o świecie i tym samym jego reprezentacja są produktem ukształtowanego społecznie dyskursu, będącego narzędziem władzy (zob. Szacki, 2006: 902–908; Bińczyk, 2017). W poststrukturalizmie/dekonstrukcjonizmie Jacques’a Derridy rzeczywistość została sprowadzona do „tekstu”, któremu nadać można różne sensy – są one efektem gry znaczeń, nieograniczonej żadnym zamkniętym systemem (zob. Chmielewski, 1997: 194–204; Elliott, 2011: 136–139). Z kolei autorzy z kręgu brytyjskich studiów kulturowych (szkoła Birmingham) uznali, że rzeczywistość nie jest neutralnym przedmiotem badań, lecz środowiskiem, w którym ludzie tworzą i walczą o swoje wartości oraz znaczenia. Nigdy nie mamy dostępu do rzeczywistości samej w sobie, ale jedynie do jej kulturowo wytworzonej, naznaczonej ideologią reprezentacji (zob. Dziamski, 2016: 58–60). Na gruncie hermeneutyki (Hans-Georg Gadamer) rozumienie dokonuje się w horyzoncie przed-sądów wynikających ze sposobu naszego bycia w świecie i uwikłane jest w koło hermeneutyczne (zob. Dehnel, 2006: 162–168; Tuchańska, 2012: 131–132). Neopragmatyzm (Richard Rorty) przyjmował, że poznanie rzeczywistości jest formą jej „używania”. Przedmioty poznania nie mają żadnego sensu poza praktyką dyskursywną podjętą dla określonego celu (zob. Tuchańska, 2012: 189–290; Turner, 2008: 695–696). Krańcowy wyraz krytyka realizmu znalazła w nieklasycznej (konstruktywistycznej) socjologii wiedzy i wyrastających z niej nurtów określanych mianem socjologii wiedzy naukowej (Sociology of Scientific Knowledge), studiów nad nauką i techniką (Science and Technology Studies)

czy społecznych badań nad nauką (Social Studies of Science). Za pioniera konstruktywistycznej socjologii wiedzy naukowej uznać można Ludwika Flecka i jego pracę *Powstanie i rozwój faktu naukowego* z 1935 r. (napisaną oryginalnie w języku niemieckim, polski przekład: Fleck, 1986). Przytaczana przez Olę Amsterdamską (1992: 137) Mary Hesse twierdzi, że ów nurt ujmuje wiedzę jako „zupełnie nie-empiryczną, a tylko społeczną”. Tu wiedza staje się kategorią wyłącznie socjologiczną; przyjmuje się, że poznanie nie ma swojej własnej, immanentnej logiki, niezależnej od zewnętrznych, społeczno-kulturowych czynników. Takie ujęcie, jak się wydaje, odnieść można do najbardziej radykalnych nurtów konstruktywistycznej socjologii wiedzy i nauki. Amsterdamska wskazuje trzy jej główne kierunki: mocny program szkoły edynburskiej (David Bloor, Barry Barnes), mikrosocjologiczny konstruktywizm lub mikrokonstruktywizm (Karin Knorr-Cetina) i makrokonstruktywizm lub socjologię translacji (Bruno Latour)<sup>3</sup>.

### Przełom epistemologiczny?

Pojęcie big data nie ma ogólnie przyjętej definicji. Nie ma ścisłego kryterium, które pozwoliłoby odróżnić zbiór danych „tradycyjnych” od big data. W koncepcji tej nie chodzi tylko o wielkość wolumenu danych. Definiuje się ją raczej przez możliwości, jakie zbiór danych stwarza, i sposób postępowania z nimi. Można przyjąć, że big data to duże zbiory danych (co najmniej peta-, a nawet eksabajty), charakteryzujące się zróżnicowaniem (różne typy, formaty: dane liczbowe, tekstowe, graficzne, audio, wideo i inne) i dużą dynamiką przepływu (dane są wytwarzane i rejestrowane w sposób ciągły oraz przetwarzane w czasie rzeczywistym), do przetwarzania których wymagane są bardziej zaawansowane narzędzia niż tradycyjne relacyjne bazy danych i arkusze kalkulacyjne. Te elementy akcentuje jedno z pierwszych i najbardziej rozpowszechnionych ujęć big data, tzw. 3V (*volume, variety, velocity*).

Do świata analityków big data rekrutują się osoby o różnym intelektualnym zapleczu i przygotowaniu teoriopoznawczym (Żulicki, 2020). W tym środowisku popularne jest jednak przekonanie, że możliwość przetwarzania dużych zbiorów danych przyniosło znaczący przełom w metodologii badań,

---

<sup>3</sup> Wspomniane nurty nie wyczerpują wszystkich stanowisk w socjologii wiedzy i nauki, które można uznać za konstruktywistyczne – co zaznacza sama autorka (Amsterdamska, 1992: 139). W ramach tej subdyscypliny istnieją również stanowiska bardziej umiarkowane.

a nawet – przełom w epistemologii (Cukier i Mayer-Schönberger, 2014; Kitchin, 2014; Stevens, Wehrens i de Bont, 2018).

Tony Hey, Stewart Tansley i Kristin Tolle (2009), powołując się na Jima Graya, proponują ujęcie dziejów nauki w cztery następujące po sobie paradygmaty. Zmiana paradygmatu wynika tu z pojawienia się nowych technik analizy danych. Pierwszy paradygmat (empiryczny) oparty był na bezpośredniej obserwacji i opisie zjawisk. Drugi (teoretyczny) wyłonił się na początku nowożytności i bazował na tworzeniu modeli i generalizacji oraz rozwijaniu na ich podstawie teorii. Trzeci (obliczeniowy) związany był z komputeryzacją i opierał się na symulacjach złożonych zjawisk. Obowiązujący dziś, czwarty paradygmat zasadza się na eksploracji dużych zbiorów danych<sup>4</sup>. Rob Kitchin (2014) identyfikuje w ramach tego paradygmatu dwa metodologiczne stanowiska: „nowy empiryzm” („new empiricism”) oraz „naukę opartą na danych” („data-driven science”). Drugie z tych stanowisk zgodne jest z koncepcją Graya – nie odrzuca się w jego ramach dotychczas stosowanej metodologii nauk, nie neguje przydatności teorii w procesie poznania, lecz stawia się przede wszystkim na indukcję, dąży do oddolnego tworzenia wiedzy o świecie poprzez przetwarzanie dużych zbiorów danych, a nie badanie odgórnych hipotez. Dysponowanie i możliwość analizowania wielkich zbiorów danych z najróżniejszych dziedzin pozwala na ujawnianie ich złożonych związków, które wcześniej, przed umowną erą big data, nie były możliwe do uchwycenia – w tym przede wszystkim tkwi *novum* „nauki opartej na danych”<sup>5</sup> (por. Frické, 2015). Nie zakłada się tu, że dane samoistnie reprezentują prawdę o świecie, lecz że po odpowiednim ich przetworzeniu i zinterpretowaniu, skonstruowaniu modeli i symulacji, pomogą odpowiedzieć na określone pytania. Nie oczekuje się także, że tworzona w ten sposób wiedza jest pewna i ostateczna.

„Nowy empiryzm” jest bardziej radykalny. Jego zwolennicy deklarują samowystarczalność danych w procesie poznania, w czym upatrują zupełnie nowego sposobu tworzenia wiedzy o świecie. Polegać ma on – mówiąc w pewnym uproszczeniu – na zastąpieniu myślenia przyczynowo-skutkowego

---

<sup>4</sup> To niewątpliwie nadmiernie uproszczona wizja, można jednak przyjąć, że w pewnym przybliżeniu i uogólnieniu wskazuje dominujące w danych czasie stanowisko poznawcze i związaną z nim metodologię. Sikora (2007: 37) w zgodzie z tą klasyfikacją wskazuje na przykład, że nowożytna nauka rozwinęła się dzięki przejściu od metodologii indukcyjnej do idealizacyjnej. Dyskusja z tak ujętą historią nauki wykracza poza ramy tego artykułu.

<sup>5</sup> Zaznaczmy jednak, że zorientowane na empirię i indukcję, stroniące od budowania odgórnych hipotez, metody badań praktykowane są od dawna w wielu obszarach nauki – jak choćby teoria ugruntowana w naukach społecznych (zob. Konecki, 2000).

wskazywaniem korelacji i wzorów w ogromnych zbiorach danych; śledząc na bieżąco zachodzące między nimi relacje, można kreować obraz rzeczywistości tu i teraz. Czyni to zbędnym odwoływanie się do teorii, stawienie hipotez i tworzenie modeli. Podejście to głosi, że analizy big data pozwalają na bezpośredni wgląd w obiektywną rzeczywistość, jej doskonałe odwzorowanie i obserwowanie jej zmian w czasie rzeczywistym. Debatę nad tak pojętym „nowym empiryzmem” w dużej mierze sprowokował artykuł Chrisa Andersona *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete* z 2008 r. W świecie biznesu oraz popularnonaukowej publicystyki tezy te stały się popularne.

Autorzy sytuujący się w nurcie „nowego empiryzmu” wierzą, że ich analizy wreszcie pozwolą w pełni „przemówić danym” (Cukier i Mayer-Schönberger, 2014: 35); wskazują, że przetwarzanie big data daje „pełen obraz” przedmiotu analizy (Oracle, 2015), „wgląd w życie w całej jego złożoności” (Rieder i Simon, 2017: 5); piszą o „eksploracji rzeczywistości” („reality mining”) (Eagle i Pentland, 2006), uciekają się nawet do Putnamowskiej metafory „punktu widzenia Boga” („God’s eye view of ourselves”) (Pentland, 2009: 80). Opis z okładki książki *Wszyscy kłamią. Big data, nowe dane i wszystko, co Internet może nam powiedzieć o tym, kim naprawdę jesteśmy* Setha Stephensa-Davidowitza głosi, że dane „wreszcie pokazują nagą prawdę”. Syllabus prowadzonego na Uniwersytecie Warszawskim kursu „Nowe źródło informacji. Big Data” deklaruje, że narzędzia do rafinacji big data „mogą zapewnić prawdziwy obraz przeszłości, obraz w czasie rzeczywistym i prognozy” (UW, 2021). W tym zdaniu zawarte jest zasadnicze z punktu widzenia praktycznej przydatności analiz big data oczekiwanie, że możliwe stanie się nie tylko tworzenie wiernego obrazu rzeczywistości, ale także przewidywanie jej przyszłych stanów – zarówno w dziedzinach przyrodniczych, jak i na obszarze zachowań społecznych (por. Barábasi, 2002). Takie ambicje przejawiało nowożytnie przyrodoznawstwo od swego zarania. Jego dążeniem było – jak pisał Edmund Husserl (1993: 12–13) – „indukowanie z wszystkich faktycznie danych zjawisk przyszłe możliwości i prawdopodobieństwa w takim zakresie i z taką dokładnością, jakie przekraczają wszelką związaną z naocznością empirię”. O ile jednak nowożytna nauka usiłowała odkryć abstrakcyjne, uniwersalne, ogólne prawa przyrody, to „nowi empiryści” twierdzą, że można dotrzeć do istoty rzeczywistości, stworzyć jej doskonałe, czyste odbicie, bez potrzeby generowania uogólnień, tworzenia teorii.

Jeśli przyjmiemy, że dokładność mapowania rzeczywistości i tworzenia prognoz jest funkcją ilości danych i przy odpowiednim ich nagromadzeniu

będziemy w stanie odtworzyć w procesie poznania wiernie świat i w sposób bezbłędny wskazywać przyszłe wydarzenia, to mamy do czynienia ze swoistą wersją Demona Laplace’a – wiarą, że danetyzacja<sup>6</sup> i matematyzacja rzeczywistości – potencjalnie – pozwolą zidentyfikować wszelkie istniejące w świecie obiekty, wzorce relacji między nimi oraz trendy i nieomylnie prognozować przyszłe stadia tegoż świata. Jednakże, o ile istnieje spór, czy w ogóle możliwy jest jednoznaczny, obiektywny opis rzeczywistości, o tyle – opierając się na najpowszechniejszych filozoficznych interpretacjach współczesnej matematyki i fizyki – nie budzi większych kontrowersji konstatacja, że Laplace’ański determinizm jest z gruntu utopią (Majcherek, 2004: 15–49; Życiński, 1993: 222; 2018: 133–134, 197–202).

Kitchin (2014) stanowisko „nowych empirystów” streścił za pomocą następujących stwierdzeń:

- analizy big data pozwalają na uchwycenie całej badanej populacji, a nie tylko jej próbki;
- analizy big data nie wymagają odwoływania się do teorii, tworzenia modeli czy przyjmowania wstępnych hipotez;
- dzięki zastosowaniu obiektywnych technik analizy dane są wolne od ludzkiej stronniczości, a wszelkie odkryte wzorce i relacje odpowiadają rzeczywistości; jako takie dane mówią same za siebie;
- sens danych jest zawarty w nich samych, jest niezależny od kontekstu i wiedzy dziedzinowej, dlatego dane mogą być interpretowane przez każdego, kto potrafi przetwarzać je za pomocą operacji statystycznych.

Autor wskazuje na fałsz tych twierdzeń. Po pierwsze, analizy big data zawsze opierają się na jakiejś części zasobu potencjalnie dostępnych danych, a ich wyniki stanowią jeden z możliwych obrazów danego fragmentu rzeczywistości; zależne są od wyboru i jakości danych, użytego sposobu ich przetwarzania oraz technologii. Po drugie, rzekomo doskonale indukcyjne odwzorowywanie świata za pomocą big data nie dzieje się w teoretycznej próżni – dane są zbierane przez instrumenty zbudowane na podstawie określonych teorii, zgodnie z metodologią opartą na jakiejś teorii i przetwarzane według procedur bazujących również na teoriach czy prawach nauki. Analizy big data nie byłyby możliwe bez zaangażowania wyrafinowanej teorii matematycznej. Po trzecie, fakt, że dane nie mówią same za siebie, wynika wprost z dwóch poprzednich punktów. Po czwarte, wyniki analiz big data nie

---

<sup>6</sup> Danetyzację rozumiem jako fenomen kwantyfikowania i przekładania wszelkich elementów rzeczywistości na dane, by móc poddać je agregowaniu i algorytmizacji; więcej zob. Iwasiński, 2016.



są wprost zrozumiałe i jednoznaczne, lecz wymagają interpretacji (pojawia się np. zjawisko apofenii, czyli wskazywania związków tam, gdzie faktycznie ich nie ma). Ograniczenie się do wskazania statystycznych relacji między danymi to nadmierny redukcjonizm. To nie operacje matematyczne, ale określony dyskurs – oparty na jakiejś teorii, zakorzeniony w pewnym kontekście – wyposaża dane w znaczenie. Dane muszą więc zostać zinterpretowane – w świetle określonej formalnej teorii, ale też przekonań, wartości, a nawet interesów badacza.

Nowi empiryści zdają się uznawać, że dane po prostu „są”, a ich pozyskiwanie jest procesem przezroczystym. Do pozyskania danych niezbędny jest wszakże pomiar – a ten, jeśli dokonywany jest za pomocą jakiegoś instrumentu, wymaga przekazania informacji za pomocą sygnału i jest zawsze teoretycznie obciążony. Jak pisze Paweł Zeidler (2010: 7): „Opis procesu przekazywania informacji za pomocą sygnału, zgodnie z klasyczną teorią komunikacji C. Shannona i W. Weavera, zakłada znajomość teorii źródła wysyłającego sygnał, teorii ośrodka, przez który on przebiega, a który może być źródłem szumów, oraz teorii odbiorcy – rejestratora sygnału, tj. aparatu percepcyjnego człowieka lub instrumentu pomiarowego”.

By dane były użyteczne, muszą zostać odpowiednio przetworzone. Ian Hacking<sup>7</sup> (za: Leciejewski, 2013: 33–34) wymienia następujące operacje, jakim poddawane są zgromadzone dane: oszacowanie (polegające na obliczeniu prawdopodobieństw błędów), redukcję (stosowaną zwykle dla uproszczenia modelu, uczynienia go bardziej czytelnym), analizę (odpowiednie modelowanie – w zależności od celu) i interpretację (w świetle jakiejś teorii). Sądzę, że w pewnych przypadkach znaczenie ma jeszcze prezentacja danych, która też nierzadko wymaga przyjęcia określonych teoretycznych ustaleń. Dopiero po zrealizowaniu tych kroków z danych wydobyć można znaczenie.

Przetwarzanie sygnałów pobranych przez urządzenia pomiarowe często wiąże się z potrzebą zmiany sygnału analogowego (o charakterze ciągłym) na cyfrowy (o charakterze dyskretnym) – tylko taki jest zrozumiały dla komputera i może być poddany dalszym procedurom analitycznym. Aby tego dokonać, sygnał należy poddać próbkowaniu, kwantowaniu i kodowaniu do postaci binarnej. Każdy rodzaj przetwornika analogowo-cyfrowego ma określoną rozdzielczość, częstotliwość i czas przetwarzania, każdy generuje

---

<sup>7</sup> Hacking pisze o obróbce danych w praktyce eksperymentalnej, ale procesom, o których wspomina, poddawane są zebrane w trakcie pomiaru dane w ogóle, by stały się nośnikiem jakiejś treści i były użyteczne.

określone błędy, a dobranie odpowiedniego wymaga znacznej wiedzy o charakterystyce sygnału wejściowego.

Te oraz kolejne elementy układu zaangażowanego w przetwarzanie danych – różnego rodzaju interfejsy, komputery, oprogramowanie, metody i techniki analizy danych – są ściśle zależne od teorii i nieneutralne, a więc od ich wyboru i konfiguracji może zależeć wynik obróbki i analizy danych (por. Leciejewski, 2013: 44–61, 68–75). W tym sensie żaden system pobierający i przetwarzający dane nie jest neutralny. Jak sądzę, istnieje również obawa, że w realiach tzw. przetwarzania bez granic (*ubiquitous computing*), którego istotą jest przetwarzanie w czasie rzeczywistym ogromnych ilości danych z wielu źródeł, przede wszystkim z sieci czujników i współpracujących ze sobą inteligentnych urządzeń (zgodnie z koncepcją Internet of Things / Internet of Everything), małe błędy i opóźnienia w poszczególnych elementach układów przetwarzania danych mogą ulegać niekontrolowanemu zwielokrotnieniu.

### **Wizja danych w nie-realistycznym (antyrealistycznym/konstruktywistycznym) modelu reprezentacji poznawczej**

Przyjmuję, że dane to zaobserwowane (przez człowieka lub urządzenie) sygnały bądź znaki. Mogą one mieć formę odczytu z urządzenia pomiarowego, liczby albo innego symbolu, fotografii, dźwięku czy innego sygnału.

Z punktu widzenia problematyki tego artykułu szczególnie istotne jest stwierdzenie, że nie istnieją w pełni surowe dane; żadne dane nie są wyposażone w samoistny sens, nie posiadają niezależnej od kontekstu, z góry określonej esencji – i jako takie nie mogą być źródłem jedyne właściwego, prawdziwego obrazu rzeczywistości. Co za tym idzie, nie sposób oddzielić czystych danych od interpretacji. Analitycy big data traktują zwykle dane jako obiektywne jednostkowe informacje, zaobserwowane i zarejestrowane (przez człowieka albo przyrząd). Takie stanowisko wyraźnie odróżnia dane (które mają charakter empiryczny) od interpretacji (która ma charakter spekulatywny). Dane są faktami, które należy uznać, polemizować można z ich interpretacjami. Pozostaje to w zgodzie ze zdroworozsądkowym założeniem, że z faktami się nie dyskutuje.

Jak pokazałem wyżej – to założenie antyrealistyczne/konstruktywistyczne nurty refleksji kwestionują. Jest ono we współczesnej filozofii nauki podważane powszechnie, nawet wśród filozofów nauki identyfikujących się

raczej z realizmem. Na przykład Michał Heller stwierdza: „filozofia nauki przebyła długą drogę w dążeniu do zrozumienia natury faktów [...]. Wiadomo dziś z pewnością, że nie istnieją «nagie fakty». [...] Między składową faktyczną naszej wiedzy a jej składową teoretyczną występuje sprzężenie zwrotne” (Heller, 2009: 42–43). Co więcej, od procesu interpretacji zależy to, czy jakiś aspekt rzeczywistości w ogóle zostanie włączony w obręb analizowanych danych. „Dane muszą zostać uznane za dane, by zaistnieć jako takie, a owo uznanie opiera się na jakiejś interpretacji” (Gitelman i Jackson, 2013: 3).

Innymi słowy, dane są zawsze, w jakimś stopniu, konstruowane. Przedmiot poznania zostaje uznany za daną, jeśli proces jego konstrukcji nie jest problematyzowany. Warto odwołać się do metafory „czarnej skrzynki”, obecnej w socjologii wiedzy od wczesnych lat siedemdziesiątych XX w., ale rozpropagowanej przez Brunona Latoura dekadę później. Dla tego autora czarną skrzynką może być każdy wskaźnik, narzędzie zbierające dane, inne urządzenie techniczne, ale też twierdzenia i teorie czy procedury statystyczne. Stają się one czarnymi skrzynkami, gdy zakończy się proces ich konstrukcji – zostaną uznane za pewne i nieproblematyczne, a ich społeczne komponenty zapomniane (za: Sojak, 2004: 238–240).

Michael Mulkey i Nigel Gilbert sugerują, że o uznaniu jakiegoś elementu rzeczywistości za daną decydować może także rodzaj dyskursu, jaki uprawiają badacze, a mianowicie posługiwanie się „empirycznym repertuarem retorycznym”. Charakteryzuje się on tym, że „mówiący przedstawia swoje działania oraz przekonania jako całkowicie neutralne, pozwalające przemówić oczywistości zjawisk empirycznych” (za: Sojak, 2004: 227); jest bezosobowy, sformalizowany, często hermetyczny, prezentuje działania jako nieproblematyczne i zrutynizowane, unika eksponowania kontrowersji, decyzje podejmowane przez ludzi składa na karb urządzeń czy procedur. O wyborze tego repertuaru decydować mogą czynniki społeczne, pragmatyczne, niezwiązane z wewnętrzną logiką samego badania (za: Sojak, 2004: 227–228). Mechanizm ukrywania czynnika ludzkiego w procesach tworzenia wiedzy, który badali w latach osiemdziesiątych XX w. Steven Shapin i Simon Schaffer (za: Sojak, 2004: 240), dziś, w dobie wszechobecnych algorytmów zarządzających najróżniejszymi sferami życia społecznego, wydaje się szczególnie ważki i bardziej niż wtedy aktualny.

Jak w praktyce wygląda konstrukcja danych i obrazowanej za ich pomocą rzeczywistości? Przywołam tu argumenty Andrzeja Zybertowicza (1995: 86–87) na rzecz stanowiska konstruktywistycznego, uzupełniając je o uwagi innych autorów i starając się zilustrować je prostymi, czytelnymi przykładami.

Po pierwsze, poznanie zawsze uwarunkowane jest zasobami pojęciowymi i wartościami wytworzonymi w danej kulturze. Weźmy następującą informację – liczbę osób chorych w danej społeczności. Jak się okazuje, wcale nie jest ona jednoznaczna, bo zależy od przyjętej definicji choroby (czyli właśnie zasobów pojęciowych i wartości). Na przykład kilka dekad temu do tego zbioru zaliczani byliby homoseksualiści, a po roku 1972 – według wytycznych American Psychiatric Association – już nie. Tym samym obiektem w innym kontekście kulturowym przypisane więc zostaną inne sensy (co zmienia obraz rzeczywistości). Często sensy przypisywane są jakimś obiektom bądź konstruktom na podstawie wskaźników czy też indeksów, które są w oczywisty sposób konstruowane społecznie (istnieją wszak różne sposoby – oparte na inaczej skonstruowanych wskaźnikach – np. liczenia bezrobocia czy inflacji). Pogłębione rozważania o konstruowaniu miar w ekonomii i ich społecznym wymiarze podejmuje Franciszek Chwałczyk (2018). Włodzimierz Okrasa (2017: 164) ujmuje samą kwantyfikację (obejmującą konwencję, pomiar i komunikację) jako zjawisko socjologiczne.

Po drugie, Zybertowicz przywołuje tezę Duhema–Quine’a. Głosi ona, że dane same w sobie nie determinują teorii (dodajmy: poznania w ogóle). Opierając się na samych danych, nie da się sformułować żadnej sensownej wypowiedzi, każda taka wypowiedź musi odwoływać się do wiedzy wykraczającej poza owe dane. By móc zrobić z danych jakiś poznawczy użytek, musimy je „zanieczyścić” jakąś wiedzą pochodzącą spoza nich. Dysponując tym samym zbiorem danych, jesteśmy zatem w stanie sformułować empirycznie równoważne, ale sprzeczne ze sobą interpretacje (Zybertowicz, 1995: 90–91). Egzemplifikacją tezy Duhema–Quine’a, wskazującą, że dane same w sobie nie determinują ustaleń badaczy, i akcentującą ściśle społeczny aspekt tworzenia danych, są badania reprezentantów Bath School nad kontrowersjami naukowymi. Wykazały one, że istnieje pewien zakres dowolności interpretacyjnej tych samych danych laboratoryjnych. Gdy pojawia się kontrowersja dotycząca tego, które wyjaśnienie przyjąć, za ich rozstrzygnięcie odpowiadają często czynniki społeczne, a nie te wynikające z samej natury badanego przedmiotu (zob. Afeltowicz, 2012: 76–77; Sojak, 2004: 225).

Po trzecie, powiada Zybertowicz, dane zawsze są w sposób selektywny wypreparowywane z jakiegoś fragmentu rzeczywistości. Dobór analizowanych w określonym badaniu danych ma charakter arbitralny – pewne zostają ujawnione, inne pominięte, co wpływa na wnioski i wytworzoną na ich podstawie reprezentację świata. Jeśli podamy, że odsetek kierowców ukaranych za wykroczenie drogowe w jakimś okresie wzrósł z 2% do 5% ogółu ich populacji, to można powstać wrażenie, że zaczęli oni jeździć

mniej ostrożnie. Ale jeśli uzupełnimy tę informację o fakt, że w tym czasie potroiła się liczba patroli na drogach albo że bardzo zaostrzono przepisy, to rzuci to inne światło na to zagadnienie. Na tę kwestię można spojrzeć także z perspektywy problemu niezdeterminowania wiedzy przez dane, którego jednym z ujęć jest teza Duhema–Quine’a. Określony zbiór danych w świetle wiedzy pochodzącej z innych danych czy też innych dyskursów zmienia swe znaczenie – ten proces reinterpretacji potencjalnie może być nieskończony. W perspektywie antyrealistycznej/konstruktywistycznej każdy dyskurs nadający sens danemu wycinkowi rzeczywistości społecznej jest wszak jedynie mniej lub bardziej chwiejną i doraźną konstrukcją. W radykalnej interpretacji owe sensory można zawsze przewartościować, odwołując się właśnie do innego dyskursu – interpretacja nigdy nie będzie obiektywna i ostateczna.

Dodajmy, że krytyka statystyki z perspektywy konstruktywistycznej mówi nawet o „wojnach statystycznych”, czyli posługiwaniu się statystyką w celu przeforsowania określonego obrazu rzeczywistości i formułowania roszczeń (Miś, 2017: 82).

### **„Nowy empiryzm” – perswazyjny dyskurs rynku usług big data**

Dyskurs „nowego empiryzmu” zwykle uprawiany (i promowany) jest przez biznes oferujący usługi związane z przetwarzaniem dużych zbiorów danych i inne środowiska posługujące się analityką danych do celów instrumentalnych, komercyjnych, zorientowanych na zysk. Narracja nowych empirystów oddziałuje jednak na wyobraźnię społeczną, tworzy swoiste imaginaria (William Housley (2015) i Ben Williamson (2015) piszą o „Big Data imaginary”) i – co szczególnie istotne – wpływa na świadomość decydentów politycznych. Gernoth Rieder i Judith Simon (2016; 2017) zwracają uwagę na wiarę w obiektywność analiz big data w zarządzaniu, administracji i polityce publicznej (m.in. w obszarze służby zdrowia, edukacji, bezpieczeństwa publicznego, rynku pracy). Określają ten fenomen mianem „datatrust”. Dlatego nie sposób zaprzeczyć, że dyskurs „nowego empiryzmu” ma istotne społeczne konsekwencje. W tym – czasem ślepy – zaufaniu do wyrażonych liczbowo uzasadnień wspomniani autorzy, nawiązując do Theodore’a M. Portera (1995), widzą skrajny wyraz typowej dla nowoczesności wiary w bezosobowe, abstrakcyjne, zmatematyzowane systemy analityczne, przyznawania takim właśnie systemom legitymacji do ostatecznego rozstrzygnięcia, co jest prawdą i jaka jest rzeczywistość.

Z kolei zwolennicy „nauki opartej na danych” częściej rekrutują się ze środowisk akademickich (Kitchin, 2014). „Nauka oparta na danych” jest świadoma uproszczeń „nowego empiryzmu”.

Sami „nowi empiryści” prawdopodobnie widzą siebie jako realistów. Wskazuje na to ich przekonanie – wyrażane w przywoływanych wyżej cytatach – że wyniki analiz big data są odbiciem realnego świata, zachodzących w nim relacji i jego struktury; że możliwe jest wierne odwzorowanie rzeczywistości za pomocą analiz big data, osiągnięcie obiektywnego wglądu, dotarcie do – klasycznie, tzn. korespondencyjnie, rozumianej – prawdy. Dyskurs ten jest, jak się zdaje, nieświadomy, bądź też świadomie ignoruje wspomniane wyżej antyrealistyczne/konstruktywistyczne nurty namysłu nad poznaniem. Można powiedzieć, że stanowi on rodzaj propagandy scjentystycznej (Woźniak, 2000). Deklaracje „nowych empirystów” należy traktować jako retoryczny środek stosowany przez branżę *data science*. Mają na celu przekonać do dostarczanych przez nią analitycznych, zwłaszcza zorientowanych na predykcję, narzędzi. Dlatego ta realistyczna retoryka znalazła entuzjastów, zwłaszcza w obszarze zarządzania sprzedażą i marketingiem (choć, jak już zaznaczyłem, oddziałuje ona także na administrację i politykę publiczną). Jest prosta i ma duży potencjał perswazyjny – operuje przekazem, który można sprowadzić do następujących haseł: dzięki naszemu oprogramowaniu poznasz rzeczywistość, zobaczysz, jak realnie wygląda świat; odkryjemy go dzięki danym, a nie spekulacjom; bazujemy na faktach, a nie domysłach. Jako dowody wskazuje się, że ilość posiadanych danych i umiejętność odnajdywania relacji (zwłaszcza korelacji) między nimi wprost przekłada się na zwiększenie sprzedaży i wskaźników finansowych (Kitchin, 2014: 4). Takie argumenty nie świadczą jednak o tym, że tworzona w wyniku analiz big data wiedza reprezentuje obiektywny świat. Argument efektywnościowy nie wyklucza też takiego stanu rzeczy – w istocie nie mówi nic o prawdziwości czy realności wytworów poznania, ale jedynie o zdolności do osiągania dzięki nim wyznaczonych celów. Wbrew promowanej retoryce „nowy empiryzm” należałoby więc traktować jako rodzaj instrumentalizmu, który klasyfikowany jest jako stanowisko antyrealistyczne bądź konstruktywistyczne (zob. Woleński, 2005: 138; Heller, 2009: 51). Klasyczny empiryzm nigdy nie aspirował zresztą do poznania ostatecznej prawdy o świecie, o świecie istniejącym poza umysłem, ani do wiedzy pewnej – wyrósł wszak ze sprzeciwu wobec teorii, które takie pretensje rościły (Popkin i Stroll, 1994: 346, 356, 393–394). Stąd myśl ojców nowożytnego empiryzmu sytuuje się w obszarze antyrealizmu – w szczególności George’a Berkeley’a i Davide’a Hume’a, jedynie

stanowisko Johna Locka można traktować jako rodzaj realizmu pośredniego (krytycznego) (Woleński, 2005: 394).

## Big data a transcendentarno-pragmatyczna forma reprezentacji wiedzy

Analizy big data zawsze opierają się na jakichś matematycznych modelach – a każdy model upraszcza rzeczywistość. Jak pisze Cathy O’Neil (2017: 46–48): „Żaden model nie może uwzględnić wszystkich komplikacji prawdziwego świata [...]. Tworząc model, dokonujemy wyboru tego, co wydaje się nam wystarczająco ważne, by to uwzględnić [...]. Luki informacyjne modelu odzwierciedlają oceny oraz priorytety ich twórców [...]. Modele, które z założenia mają być bezstronne, odzwierciedlają z góry założone cele oraz osobiste przekonania [...]. Modele są opiniami opisanymi językiem matematyki [...]. Zawsze musimy zadać sobie pytanie, nie tylko o to, kto stworzył dany model, lecz również co ta osoba lub przedsiębiorstwo starają się osiągnąć”. Z kolei Sikora (2007: 197–198) zauważa, że „wszystkie modele mają charakter kontekstowy” i są konstruowane ze względu na określone cele i wartości.

W pewnym sensie analizy big data można traktować jako rodzaj eksperymentów. Jak pisze Hacking: „Eksperymentować to tyle, co kreować, oczyszczać i stabilizować zjawiska” (za: Sikora, 2007: 135). Natomiast służąca ich realizacji techniczno-ludzka infrastruktura stanowi rodzaj laboratorium. W tak rozumianym laboratorium zachodzi ciąg czynności służących wytworzeniu określonej wiedzy. Czynności te mają wymiar nie tylko techniczny, ale i społeczny, kulturowy, polityczny – obecne na każdym etapie produkcji wiedzy: od zbierania danych, przez ich przetwarzanie, po interpretację. To uwikłanie widać choćby w procesie budowania algorytmów przetwarzających dane, których celem jest prognozowanie zachowań ludzi czy przyszłych stanów rzeczywistości (Iwasiński, 2020). Matematyk David Sumpter argumentuje, że nie istnieją algorytmy wolne od ideologicznych nachyleń. Każdy algorytm, z jakiegoś punktu widzenia, jest niesprawiedliwy, dyskryminuje jakąś grupę, która poddawana jest za jego pomocą analizie. A to, której grupy będzie to dotyczyć, nie jest kwestią matematyki, tylko aksjologii – przekonań i poczucia sprawiedliwości twórcy algorytmu. „Niesprawiedliwość przypomina jedną z tych gier whack-a-mole, które można spotkać na festynach. W takiej grze kret pojawia się w różnych miejscach. Uderzasz go młotem w jednym miejscu, a on wyskakuje w innym [...]. Nie istnieje równanie na sprawiedliwość. Sprawiedliwość jest czymś

ludzki” (Sumpter, 2019: 83–84; por. O’Neil, 2017: 126–149; szerzej o aksjologicznym wymiarze nauki zob. Lekka-Kowalik, 2008).

W kontekście wszystkich powyższych rozważań nie sposób bronić tezy, że analizy big data dostarczają nam obiektywnej wiedzy i trafnie odzwierciedlają rzeczywistość. Uprawnione wydaje się co najwyżej stwierdzenie, że pozwalają skuteczniej, w porównaniu z klasycznymi metodami badań i przetwarzania danych, zinstrumentalizować wiedzę, innymi słowy – pozwalają na dokładniejsze odpowiedzi na konkretne pytania, pod warunkiem że owe pytania zostaną właściwie przełożone na matematyczny model i procedurę analityczną<sup>8</sup>. W tej optyce konceptualizacja wytworów poznania ma charakter instrumentalnie użyteczny, a pytanie o ich odniesienie przedmiotowe „traci sens nadany mu przez klasycznych realistów” (Sikora, 2007: 21). Tak pojmowane poznanie zakorzenione jest w filozofii pragmatycznej, która działania poznawcze ocenia nie z punktu widzenia ich zgodności z rzeczywistością, ale praktycznych skutków, jakie niosą. Opartą na przedstawionym tu rozumowaniu koncepcję reprezentacji poznawczej Sikora nazywa transcendentualno-pragmatyczną. Można ująć ją jako określanie zależności występujących między wytworami poznania a tym, jak i ze względu na jakie cele doszło do ich powstania (Sikora, 2007: 195). Sikora wyszczególnia jej dwa typy: pojęciowy i laboratoryjny. W pierwszym kładzie nacisk na konwencje językowe i fakt, że wybór sieci pojęciowej odnoszącej się do poddawanego analizie świata ma charakter przede wszystkim pragmatyczny. Nawet jeśli dane do analizy zbierane są automatycznie, to konstrukcja tych danych i sposób ich przetwarzania opierają się na określonej operacjonalizacji pojęć. „Mówienie o faktach bez uwzględnienia specyfiki języka, w którym fakty te wyrażamy, to orzekanie o niczym” – powiada Hilary Putnam (za: Sikora, 2007: 121). W typie laboratoryjnym główny wpływ na wytwory poznania mają

---

<sup>8</sup> O’Neil stosuje gradację modeli, uznając je za mniej lub bardziej sprawiedliwe i dokonujące predykcji gorzej lub lepiej. Jako przykład modelu „zdrowego” wskazuje ten stosowany w rozgrywkach baseballowych. Jak pisze – jest on przejrzysty, zasilany ogromną ilością danych zaczerpniętych bezpośrednio z rozgrywek, a nie zmiennymi pośrednimi, ciągle aktualizowany, jego założenia i wnioski są czytelne, cel jasno wyartykułowany. Nie oznacza to jednak, że jest on w pełni obiektywny – jedynie, że ze względu na opisaną wyżej charakterystykę, pozwala lepiej niż inne modele przedstawione w pracy tej autorki osiągać cel, dla którego został stworzony. Nie uważam wszakże, by taki model można było uznać za czyste odzwierciedlenie rzeczywistości – choć O’Neil (2017: 44) stwierdza, że jego sukces bazuje na tym, iż „pozostaje w stałym kontakcie z rzeczywistością, którą stara się zrozumieć lub przewidzieć”.



praktyki, interakcje i otoczenie społeczne. Ta perspektywa typowa jest dla radykalnych wersji konstruktywistycznej socjologii wiedzy naukowej i ma tendencję do popadania w nadmierny socjologiczny redukcjonizm. Na przykład mikrokonstruktywizm Knorr-Cetiny proces, w wyniku którego ustalony zostaje określony sposób przetwarzania danych, analizuje przez pryzmat społecznych interakcji i „strategii konwersacyjnych” członków grupy badawczej – w takiej perspektywie wiedza staje się wyłącznie pochodną konkretnej sytuacji społecznej (za: Amsterdamska, 1992: 148).

Na zakończenie warto pokazać, że przyjęte przeze mnie rozstrzygnięcia dotyczące statusu epistemologicznego oraz reprezentacji analiz big data bywają ujmowane na inne sposoby. W artykule *Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis* (Stevens, Wehrens i de Bont, 2018) autorzy wyszczególnili pięć typów idealnych dyskursów na temat big data: „modernistyczny”, „instrumentalny”, „pragmatyczny”, „naukowy” i „krytyczno-interpretatywny”. Zaklasyfikowali je do różnych kategorii, uwzględniając różne kryteria, wśród których znajdował się status epistemologiczny. Trzy pierwsze, w ich interpretacji, pojmują dane realistycznie, jako byty zewnętrzne wobec poznającego podmiotu, wyposażone w samoistny sens. Przyjmują one, że wnioskowanie wynika bezpośrednio z danych. W przypadku dyskursu „modernistycznego” uznaje się, że analizy big data oferują wiedzę w pełni wiarygodną, w przypadku dyskursu „instrumentalnego” – potencjalnie również, ale wiarygodność jest uzależniona od technologii, w przypadku dyskursu „pragmatycznego” dane traktuje się jako wiarygodne w określonych sytuacjach. W dyskursach „naukowym” oraz „krytyczno-interpretatywnym” dane nie mówią same za siebie: w tym pierwszym wymagają wsparcia teorii, a wiarygodność analiz może być zadowalająca, jeśli spełni się rygorystyczne kryteria metodologiczne, w drugim – interpretacja danych wymaga krytycznego myślenia, a wiarygodność analiz big data jest niska.

### Uwaga końcowa

Z powyższych rozważań nie należy wyciągać wniosku, że neguję wartość i korzyści płynące z analiz big data. Uważam, że mogą być i bywają ważnym, być może najważniejszym i najskuteczniejszym dziś, narzędziem kontroli rzeczywistości i prognozowania (co nie jest równoznaczne ze zdolnością do odkrywania obiektywnej prawdy i tworzenia trafnych reprezentacji rzeczywistości). Nie sugeruję też, że sens danych można dowolnie

relatywizować. Trzeba jednak mieć świadomość, że wybrane elementy rzeczywistości nabierają statusu danych i takiego, a nie innego znaczenia na gruncie określonych założeń, zarówno o charakterze metodologicznym, technicznym czy technologicznym, jak i społecznym oraz kulturowym – odnoszącym się do wartości i interesów uwikłanych w proces analizy społeczności. Nie jestem też zwolennikiem podzielanej przez radykalne nurty konstruktywistycznej socjologii wiedzy i nauki socjologicznego redukcjonizmu i abstrahowania od wewnętrznej, samoistnej logiki procesów poznawczych. Uważam jednak, że społeczny komponent produkcji wiedzy jest w niej (niemal) zawsze obecny, ale w różnych dziedzinach i konkretnych problemach w różnym stopniu.

## Literatura

- Afeltowicz, Ł. (2012). *Modele, artefakty, kolektywy. Praktyka badawcza w perspektywie współczesnych studiów nad nauką*. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- Amsterdamska, O. (1992). Odmiany konstruktywizmu w socjologii nauki. W: J. Niżnik (red.), *Pogranicza epistemologii* (ss. 136–154). Warszawa: IFIS PAN.
- Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired* 16(7).
- Barábasi, A.L. (2002). *Linked: The New Science of Networks*. Cambridge: Perseus Press.
- Bińczyk, E. (2017). Michel Foucault, nieklasyczna socjologia wiedzy oraz perypetie władzy-wiedzy. W: T.M. Korczyński (red.), *Współczesne problemy socjologii wiedzy: w 80-lecie „Ideologii i utopii” Karla Mannheim’a* (ss. 67–82). Warszawa: Warszawskie Wydawnictwo Socjologiczne.
- Chmielewski, A. (1995). *Filozofia Poppera. Analiza krytyczna*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Chmielewski, A. (1997). *Niewspółmierność, nieprzekładalność, konflikt*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Chwałczyk, F. (2018). *Miary jako modele pośredniczące między gospodarką a ekonomią*, <https://www.filozofia-ekonomii.pl/publikacje-psfe/working-papers> (dostęp: 23.09.2021).
- Czarnocka, M. (2003). O realizmie we współczesnej filozofii nauki. *Przegląd Filozoficzny – Nowa Seria*, 48(4), 5–20.
- Dehnel, P. (2006). *Dekonstrukcja – rozumienie – interpretacja: studia z filozofii współczesnej i nie tylko*. Kraków: Universitas.
- Dziamiński, G. (2016). *Kulturoznawstwo, czyli wprowadzenie do kultury ponowoczesnej*. Gdańsk: Wydawnictwo Naukowe Katedra.
- Eagle, N., Pentland, A.S. (2006). Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4), 255–268.

- Elliott, A. (2011). *Współczesna teoria społeczna*. Warszawa: Wydawnictwo Naukowe PWN.
- Fleck, L. (1986). *Powstanie i rozwój faktu naukowego: wprowadzenie do nauki o stylu myślowym i kolektywie myślowym*. Lublin: Wydawnictwo Lubelskie.
- Frické, M. (2015). Big data and its epistemology. *Journal of the Association for Information Science and Technology*, 66(4), 651–661.
- Gitelman, L., Jackson, V. (2013). Introduction: Raw data is an oxymoron. W: L. Gitelman (red.), *Raw Data is an Oxymoron* (ss. 1–15). Cambridge, MA: MIT Press.
- Heller, M. (2009). *Filozofia nauki. Wprowadzenie*. Kraków: Petrus.
- Hey, T., Tansley, S., Tolle, K.M. (2009). *Jim Gray on eScience: A transformed scientific method*, <http://itre.cis.upenn.edu/myl/JimGrayOnE-Science.pdf> (dostęp: 23.09.2021).
- Housley, W. (2015). *Focus: The Emerging Contours of Data Science*. *Discovery Society*, August 3, <https://archive.discoverysociety.org/2015/08/03/focus-the-emerging-contours-of-data-science/> (dostęp: 23.09.2021).
- Husserl, E. (1993). *Kryzys europejskiego człowieczeństwa a filozofia*. Warszawa: Aletheia.
- Iwasiński, Ł. (2016). Społeczne zagrożenia danetyzacji rzeczywistości. W: B. Sosińska-Kalata (red.), *Nauka o informacji w okresie zmian. Informatologia i humanistyka cyfrowa* (ss. 135–146). Warszawa: SBP.
- Iwasiński, Ł. (2020). Social implications of algorithmic bias. W: B. Sosińska-Kalata (red.), *Nauka o informacji w okresie zmian. Rewolucja cyfrowa – dziś i jutro: infrastruktura, usługi, użytkownicy* (ss. 25–35). Warszawa: SBP.
- Kant, I. (1957). *Krytyka czystego rozumu*, t. 1, tłum. R. Ingarden. Warszawa: Państwowe Wydawnictwo Naukowe.
- Kant, I. (1960). *Prolegomena do wszelkiej przyszłej metafizyki, która będzie mogła wystąpić jako nauka*, tłum. B. Bornstein, na nowo oprac. J. Suchorzewska. Warszawa: Państwowe Wydawnictwo Naukowe.
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big data & society*, 1(1), <https://journals.sagepub.com/doi/full/10.1177/2053951714528481> (dostęp: 23.09.2021).
- Kołąkowski, L. (1966). *Filozofia pozytywistyczna: od Hume'a do Koła Wiedeńskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Konecki, K. (2000). *Studia z metodologii badań jakościowych. Teoria ugruntowana*. Warszawa: Wydawnictwo Naukowe PWN.
- Krajewski, W. (1995). „Reprezentacjonizm” i realizm. *Filozofia Nauki*, 3(3), 79–83.
- Leciejewski, S. (2013). *Cyfrowa rewolucja w badaniach eksperymentalnych. Studium metodologiczno-filozoficzne*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.
- Lekka-Kowalik, A. (2008). *Odkrywanie aksjologicznego wymiaru nauki*. Lublin: Wydawnictwo KUL.
- Majcherek, J. (2004). *Źródła relatywizmu w nauce i kulturze XX wieku: od teorii względności do postmodernizmu*. Kraków: Universitas.
- Mayer-Schönberger, V., Cukier, K. (2014). *Big data: rewolucja, która zmieni nasze myślenie, pracę i życie*. Warszawa: MT Biznes.
- Miś, L. (2007). *Problemy społeczne: teoria, metodologia, badania*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- O’Neil, C. (2017). *Broń matematycznej zagłady: jak algorytmy zwiększają nierówności i zagrażają demokracji*. Warszawa: Wydawnictwo Naukowe PWN.

- Okrasa, W. (2017). Społeczne wyznaczniki procesu statystycznego i status statystyki publicznej z perspektywy socjologii statystyki. W: T.M. Korczyński (red.), *Współczesne problemy socjologii wiedzy: w 80-lecie „Ideologii i utopii” Karla Mannheim’a* (ss. 163–188). Warszawa: Warszawskie Wydawnictwo Socjologiczne.
- Oracle (2015). *Assessing the role of Big Data in tackling financial crime and compliance management*, <https://www.oracle.com/a/ocom/docs/industries/financial-services/fs-big-data-fccm-wp.pdf> (dostęp: 23.09.2021).
- Pentland, A. (2009). Reality mining of mobile communications: Toward a new deal on data. W: S. Dutta, S. Mia (red.), *The Global Information Technology Report 2008–2009* (ss. 75–80). Geneva: World Economic Forum.
- Popkin, R.H., Stroll, A. (1994). *Filozofia*. Poznań: Zysk i S-ka Wydawnictwo.
- Popper, K.R. (1977). *Logika odkrycia naukowego*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Porter, T.M. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- Rieder, G., Simon, J. (2016). Datatrust: Or, the political quest for numerical evidence and the epistemologies of Big Data. *Big Data & Society*, 3(1).
- Rieder, G., Simon, J. (2017). Big data: A new empiricism and its epistemic and socio-political consequences. *Berechenbarkeit der Welt?*, 85–105.
- Sikora, M. (2007). *Problem reprezentacji poznawczej w nowożytniej i współczesnej refleksji filozoficznej*. Poznań: Wydawnictwo Naukowe Instytutu Filozofii UAM.
- Sojak, R. (2004). *Paradoks antropologiczny*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Stephens-Davidowitz, S. (2019). *Wszyscy kłamią. Big data, nowe dane i wszystko, co Internet może nam powiedzieć o tym, kim naprawdę jesteśmy*. Kraków: Wydawnictwo Literackie.
- Stevens, M., Wehrens, R., de Bont, A. (2018). Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society*, 5(2), 1–21.
- Sumpter, D. (2019). *Osaczeni przez liczby*. Warszawa: Copernicus Center Press.
- Szacki, J. (2006). *Historia Myśli Socjologicznej*. Warszawa: Wydawnictwo Naukowe PWN.
- Tuchańska, B. (2012). *Dlaczego prawda? Prawda jako wartość w sztuce, nauce i codzienności*. Warszawa: Poltext.
- Turner, J.H. (2008). *Struktura teorii socjologicznej*. Warszawa: Wydawnictwo Naukowe PWN.
- UW (2021). [https://usosweb.uw.edu.pl/kontroler.php?\\_action=katalog2/przedmioty/pokazPrzedmiot&prz\\_kod=1600-SZD-WM-NZI&callback=g\\_78db2ce2](https://usosweb.uw.edu.pl/kontroler.php?_action=katalog2/przedmioty/pokazPrzedmiot&prz_kod=1600-SZD-WM-NZI&callback=g_78db2ce2) (dostęp: 23.09.2021).
- Williamson, B. (2015). *Smarter Learning Software: Education and the Big Data Imaginary*. <http://dspace.stir.ac.uk/handle/1893/22743> (dostęp: 23.09.2021).
- Woleński, J. (2005). *Epistemologia: poznanie, prawda, wiedza, realizm*. Warszawa: Wydawnictwo Naukowe PWN.
- Woźniak, T. (2000). *Propaganda scjentystyczna. Funkcje społeczne przekazów popularno-naukowych*. Warszawa: IFIS PAN.
- Zeidler, P. (2010). Czy można zaobserwować orbitale? O problemie obserwowalności i rzeczywistości przedmiotów teoretycznych. *Filozofia Nauki*, 18(4/72), 5–22.

- Zybertowicz, A. (1995). *Przemoc i poznanie: studium z nie-klasycznej socjologii wiedzy*. Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.
- Żulicki, R. (2020). *Data science w Polsce. Etnografia społecznego świata*. <https://dspace.uni.lodz.pl/handle/11089/35209> (dostęp: 23.09.2021).
- Życiński, J. (1993). *Granice racjonalności. Eseje z filozofii nauki*. Warszawa: Wydawnictwo Naukowe PWN.
- Życiński, J. (2018). *Świat matematyki i jej materialnych cieni*. Warszawa: Copernicus Center Press.