

mgr Remigiusz Żulicki

Katedra Socjologii Kultury, Instytut Socjologii, Wydział Ekonomiczno-Socjologiczny, Uniwersytet Łódzki

ORCID: 0000-0003-2624-2422

e-mail: remigiuszżulicki@uni.lodz.pl

Pułapki myślowe data-driven. Krytyka (nie tylko) metodologiczna

The mind traps of data-driven. (Not only) methodological critique

Podejście data-driven jest silnym trendem we współczesnym marketingu i innych dziedzinach życia. Z pozycji krytyki metodologicznej autor omawia pułapki myślowe data-driven. Wskazuje, że wbrew pozorom, dane nie mówią same za siebie, dane nie są „dane”, tylko zbierane i generowane, więc nieobiektywne. Większa ilość danych nie musi oznaczać większej wartości. Stosowane narzędzia kształtują badaną rzeczywistość. Wiara w dane jest wiarą, niemożliwą do zasadniczo logicznego uzasadnienia. Panuje nie konieczność, ale moda na korzystanie z danych i najnowocześniejszych metod obliczeniowych. Autor wskazuje powiązania założeń epistemologicznych z etycznymi oraz konsekwencjami społecznymi oraz proponuje alternatywy.

Słowa kluczowe

data-driven, data science, sztuczna inteligencja, etnografia, metodologia

The data-driven approach is a strong trend in modern marketing and other areas of life. From the position of methodological criticism, I discuss the data-driven mind traps. I indicate that contrary to popular beliefs: data does not speak for themselves; data is not given, but collected and generated, so biased; a larger amount of data does not necessarily mean a higher value; the tools used shape the examined reality; faith in data is a belief that is impossible to justify logically; there is no imperative but trend for the use of data and the most modern quantitative methods. I point to the links between epistemological and ethical assumptions as well as social consequences. I suggest alternatives.

Keywords

data-driven, data science, artificial intelligence, ethnography, methodology

JEL: C18

Wstęp

Truizmem jest stwierdzenie, że data-driven to silny trend we współczesnym marketingu. Robert Kozłowski wskazał, że kierunki zmian w marketingu to m.in. „nacisk na pomiar i ocenę efektywności działań” oraz „dominacja rozwiązań (...) opartych na Internecie i narzędziach cyfrowych” (2016, s. 105). Panuje „nieodparte (*overwhelming*) pragnienie bycia data-driven” — 100% badanych szefów marketingu „wierzy, że aby odnieść sukces, firmy muszą podejmować decyzje marketingowe, bazując na danych” (Kumar i in., 2013, s. 330–331). Dane na temat klientów są więc traktowane jako zasób, który można przekuć na wyniki finansowe: „Celem data-driven marketing jest optymalizacja procesów i komunikacji marketingowej (...). Kiedy poznamy odpowiedź (...), będziemy mogli bez przeszkód prowadzić nie tylko efektowne, ale przede wszystkim

efektywne kampanie online, pozwalające na dotarcie z przekazem do spersonalizowanego użytkownika, co w efekcie przełoży się na wzrost sprzedaży (Jeske, 2018).

Publikowane są — także w Polsce — artykuły naukowe poświęcone problematyce data-driven (Golik-Górecka, 2017; 2018; Kumar i in., 2013; Łapińska, 2019; Micheaux i Bosio, 2018; Mulvenna, Norwood i Büchner, 1998; Nadler i McGuigan, 2018; Provost i Fawcett, 2013; Wang i Qu, 2008), książki (Artun i Levin, 2015; Błazewicz, 2016; Harris, 2012; Jeffery, 2010; Mazurek, 2018; Nicolaus i in., 2016; PTBRiO, 2017), poradniki (Barton i Court, 2013; Chudziak i Zalewska, 2018; Czech, 2018; Evans, 2016; Grygierek, 2017; Halper, 2018; Hippold, 2018; Karasek, 2018; Mah, 2019; Mojżesz-Zimonczyk, 2017; Moneta, 2018), organizowane są warsztaty i konferencje (Divan, 2018; Parvinen, Kaptein i Pöyry, 2012; PTBRiO, 2019).

Data-driven (dosłownie: napędzanie danymi), czyli bazowanie na danych w podejmowaniu decyzji biznesowych (Provost i Fawcett, 2013, s. 53), jest przeciwstawiane podejściu opartemu na intuicji lub opinii, czyli „widzimisię” (Mojżesz-Zimonczyk, 2017). Decydowanie na podstawie intuicji jest przedstawiane jako gorsze (czyli przede wszystkim mniej efektywne), a także przestarzałe i bezsensowne (ponieważ analiza danych staje się coraz bardziej przystępna, z uwagi na łatwiejszy i tańszy dostęp do danych, moce obliczeniowe i narzędzia analityczne).

Podejście to, łączące się, łączone i mylone (Provost i Fawcett, 2013) z pojęciami: big data, sztuczna inteligencja (*artificial intelligence*, AI), data science (DS), uczenie maszynowe (*machine learning*, ML), jest obecnie modne nie tylko w marketingu, ale w wielu dziedzinach biznesu, nauki, i sfery publicznej. Z tego powodu, z pozycji krytyki metodologicznej wobec entuzjastycznego zaufania do danych i analiz ilościowych, przedstawię charakterystyczne dla data-driven pułapki myślowe i zaproponuję alternatywy. Artykuł jest głównie studium literatury, zawiera elementy własnych badań etnograficznych DS w Polsce¹.

Warto być data-driven

Jako główne benefity podejścia data-driven wskazywano na możliwość pomiaru rezultatów działań marketingowych, a w konsekwencji na poprawę ich efektywności w stosunku do kosztów. Zazwyczaj także korzysta się z danych, by monitorować wzorce zachowań klientów — mowa o sklepach internetowych (*on-line retailers*) — w celu wykrywania zdarzeń, np. odejścia klienta i reagowania na nie (Mulvenna i in., 1998, s. 35). Zbadano wpływ stosowania podejścia data-driven na wydajność (*performance*) firm. Stwierdzono wydajność wyższą o 5–6% w firmach stosujących podejście data-driven niż w firmach niestosujących tego podejścia, przy kontroli innych zmiennych (Brynjolfsson, Hitt i Kim, 2011).

Z pomocą analizy danych dąży się do odejścia od segmentacji klientów do personalizacji komunikatu/produktu. Segmentacja to podejście, gdzie relacja przebiegała na linii firma-wielu klientów w możliwie jednorodnej grupie. Personalizacja to relacja firma-pojedynczy klient, co ma być charakterystyczne dla marketingu z użyciem big data: „Klient jest traktowany podmiotowo jako niezależna jednostka o specyficznych potrzebach i preferencjach. Firma pragnie poznać tę charakterystykę poprzez rozmowy, utrzymywanie relacji oraz poprzez zbieranie i analizę śladów cyfrowych, jakie klient pozostawia w ramach różnych interakcji” (Surma, 2017, s. 48–49).

Firmy podkreślają korzyści płynące z takiej komunikacji „one-to-one”, mówiąc m.in. o coraz bardziej wymagającym kliencie z pokolenia Z (Jeske, 2018). Dzięki personalizacji oferta rynkowa będzie jakoby coraz lepiej odzwierciedlać ludzkie preferencje, a konsumenci będą mogli poznać i zaspokoić swoje głębokie czy nieuświadomiane potrzeby (Iwański, 2017, s. 125).

Korzystanie z danych doprowadzić ma do tego, że: „będziemy tworzyć firmy bystrzejsze i reagujące szybciej niż ludzie, którzy te firmy prowadzą” (Minelli, Dhiraj i Chambers, 2013, s. XVII). Analiza danych ma pomóc firmom w wytyczaniu celów, a także dostarczać propozycji sposobów osiągnięcia wybranych celów.

Popularność zdobywa także niejasne pojęcie AI. Rozumiane jest ono jako szeroka gama technologii: od bazujących na zaprogramowanych regułach przez te korzystające ze statystyki czy ML do najnowszych technik uczenia głębokiego (*deeplearning*). Systemy oparte na AI wykonują zadania, korzystając z danych i obliczeń, np. rozpoznawanie mowy, tłumaczenie, rozpoznawanie obrazów, prognozowanie, podejmowanie decyzji (*determination*) (Crawford i in., 2018, s. 44). W marketingu mówi się w tym kontekście o automatyzacji działań, używane jest też pojęcie MarTech: „Inwestycja w nowe technologie oraz ich adaptacja stają się koniecznością, a im szybciej przedsiębiorstwa z nich skorzystają, tym większą przewagę konkurencyjną zdobędą. Powszechny dostęp do ogromnych zasobów danych da przewagę tym działom marketingu, które je odpowiednio przetworzą przy użyciu technologii i będą stale aktualizować swoją wiedzę w zakresie dostępności nowych rozwiązań MarTech” (Łapińska, 2019).

Statystyka ma być najważniejszym źródłem przewagi konkurencyjnej, „plastikiem przyszłości” (Schmidt, Rosenberg i Eagle, 2014, s. 214–215), zaś nowe wcielenie statystyka/analitka danych — data scientist — najbardziej pożądanym, atrakcyjnym zawodem XXI wieku (Davenport i Patil, 2012).

Dane nie mówią same za siebie

Entuzjazm wobec możliwości i sukcesów praktycznego wykorzystywania danych zaowocował deklaracjami o nadejściu epoki petabajtów (*The Petabyte Age*). Streszczają to słowa „liczby/dane mówią same za siebie”: „Jest to świat, w którym ogromne ilości danych i matematyki stosowanej zastępują każde inne narzędzie, które można wykorzystać. Każdą teorię ludzkiego zachowania, od językoznawstwa po socjologię. Zapomnijmy o taksonomii, ontologii i psychologii. Kto wie, dlaczego ludzie robią to, co robią? Chodzi o to, że oni to robią i możemy śle-

dzić i mierzyć to z niespotykaną wiernością. **Przy wystarczającej ilości danych, liczby mówią same za siebie** [podkreślenie RZ]” (Anderson, 2008).

Krytycy wskazują, że takie „aroganckie” podejście uznaje za nieważne każdą teorię z każdej dziedziny nauki i marginalizuje inne sposoby zdobywania wiedzy (Boyd i Crawford, 2012, s. 666). Savage i Halford piszą jasno: „Liczby nie mówią same za siebie. To my umożliwiamy te wypowiedzi — za pomocą metod, które stosujemy i interpretacji, które czynimy” (2017, s. 11). Entuzjaści mówią o końcu epoki ekspertów dysponujących teoretyczną wiedzą, ponieważ w myśl charakterystycznej dla Epoki Petabajtów zasady „co, zamiast dlaczego” interesujące są jedynie relacje między zmiennymi, a nie teorie te relacje wyjaśniające.

Wskazałem (Zulicki, 2017, s. 194), że — zdaniem entuzjastów — wiedza dziedzinowa jest więcej niż zbędna. Ona po prostu przeszkadza. Typowy przykład to pokazana w filmie pt. *Moneyball* historia drużyny baseballowej Oakland Athletics. Wcześniej słaby zespół zdobył pierwsze miejsce w lidze, uzyskując 20 zwycięstw pod rząd. Trener poległ na ilościowej analizie danych, odrzucając wiedzę, doświadczenie i „widzimisię” emerytowanych zawodników i trenerów (Cukier i Mayer-Schönberger, 2014).

Podsumowując: „Najważniejszym efektem big data będzie to, że decyzje oparte na danych ulepszą jakość ocen dokonywanych przez ludzi lub sprawią, że całkowicie stracą one na znaczeniu. (...) Ekspert czy specjalista w danej dziedzinie straci część swojego znaczenia na rzeczy statystyka czy analityka danych, którzy są nieskrępowani starymi metodami rozwiązywania problemów i pozwalają przemawiać danym” (Cukier i Mayer-Schönberger, 2014, s. 185).

Boyd i Crawford nie zgadzają się z podobnymi tezami, ale wskazują, że big data/DS zmieniają sposób, w jaki myślimy o badaniach w ogóle. To zmiana epistemologiczna i etyczna. Przeramowaniu podlegają kluczowe pytania o konstytucję wiedzy, proces badania, obcowanie badacza z informacjami czy naturę rzeczywistości. Akty pomiaru kształtują mierzony świat (Boyd i Crawford, 2012, s. 665), zatem nie są to „tylko” narzędzia.

Krytykę podjął Rob Kitchin (2014), syntetyzując wiele wcześniejszych prac. Twierdzi on, że big data/DS mają wartość poznawczą dla wielu obszarów nauki, ale nie bez odwołania do teorii dziedzinowych. Mówienie o zmianie paradygmatów, końcu ery ekspertów i tym podobne entuzjastyczne stwierdzenia są bezzasadne, a także niebezpieczne w sensie poznawczym. Wskazał on na cztery epistemologiczne obietnice big data, odnosząc się do każdej z nich (Kitchin, 2014, s. 4–6):

- **Pozwala ująć całość problemu (w myśl strategii N = all) i zapewnić pełne wsparcie decyzjom** — zawsze badana jest próba, a nie populacja, chociażby z uwagi na ramy czasowe;

przy nieznanymi jej obciążeniach wyciąganie wniosków o całej populacji prowadzi do błędów; dane nie są czystą reprezentacją wycinka rzeczywistości — zbierane są zawsze z pewnego punktu widzenia; pomiary są społecznie konstruowane (silną i nieprzekraczalną ramę tworzą decyzje o tym, co zapisywać i przechowywać).

- **Do uzyskania wartościowych wyników nie są potrzebne teorie ani stawianie hipotez** — big data także nie wzięło się znikąd, zatem nie jest wolne od założeń filozoficznych i ontologicznych; reprezentacjami założeń są technologie do pracy z danymi; iluzją jest niestawianie pytań i hipotez — zostały postawione wcześniej i gdzie indziej, niż sądzą entuzjaści.
- **Dane mówią same za siebie, nieobciążone zbędą teorią, więc wyniki analiz są znaczące i zgodne z prawdą o świecie** — za wynikami analiz stoją zarówno zastosowane technologie, aparat matematyczny, jak i wiedza potoczna analityków; wyniki same w sobie nie mają znaczenia, więc analiza danych zawsze odbywa się wewnątrz pewnej ramy interpretacyjnej (świadomej lub nie), obciążającej uzyskiwane rezultaty.
- **Wyniki analiz, niezależnie od przedmiotu, może interpretować każdy posiadający rozzeznanie w statystyce** — ignorowanie teorii substancjalnych, szczególnie w przypadku, gdy badani są ludzie, prowadzi do bardzo ograniczonych wniosków; koncentracja na szukaniu ilościowych zależności w zbiorze danych prowadzi do wniosków powierzchownych, trywialnych bądź bezsensownych, będących skutkiem wykrycia związków pozornych.

Kitchin widzi dwie potencjalne ścieżki rozwoju współczesnej nauki. Tę zdecydowanie bardziej pożądaną nazywa nauką opartą na danych (*data-driven science*) — ma być to przeformułowanie sposobu jej uprawiania (propozycję tę przedstawię dalej). Inny scenariusz to skrajny empiryzm, czyli dane mówią same za siebie (*data can speak for themselves free of theory*) jako zasada główna (Kitchin, 2014, s. 10).

Także matematycy argumentowali za brakiem możliwości uprawiania nauki wyłącznie na podstawie „surowych danych” (Hosin i Vulpiani, 2018). Autorzy odwołali się do słów Andersona o liczbach mówiących same za siebie. Nazywają je manifestem ideologicznym danocentrycznego (*datacentric*) entuzjazmu i wskazują na dwie jego cechy. Po pierwsze, wygoda — za przekonaniem o mówiących za siebie liczbach stoi obietnica uzyskania prostej odpowiedzi na pytanie dotyczące bardzo skomplikowanej materii. Po drugie, odrzucenie przestarzałych poglądów — ostrożne i pracochłonne przyrodnicze podejście do zdobywania wiedzy o świecie jest jakoby nieefektywne, ograniczające i hamujące rozwój nowoczesnych technologii, zatem powinno być odrzucone (Hosni i Vulpiani, 2018, s. 121).

Dane nie są dane

„Nie może umykać nam fakt, że nie są one [dane] neutralne i nie wyrażają konieczności, lecz są mniej lub bardziej arbitralną konstrukcją stworzoną przez ludzi” (Iwasiński, 2017, s. 126). Dane są zawsze zebrane lub wygenerowane, nigdy nie są one nam dane (od czasownika „dać”). Nie można zatem traktować danych jako neutralnej, przejrzystej, pewnej, autonomicznej, prawdziwej reprezentacji rzeczywistości (Bowker, 2005; Gitelman i Jackson, 2013).

Określenie *raw data* użyte jest nawet w kursie Google Analytics dla początkujących (Google Akademia Analytics, 2019). „Surowe dane (*raw data*) to jednocześnie oksymoron i zły pomysł. Dane należy troskliwie przyrządzać (*should be cooked with care*)” (Bowker, 2005, s. 184), czyli podchodzić krytycznie do założeń stojących za mechanizmem zbierania lub generowania danych. Wszyscy analitycy interpretują dane, czego pierwszym krokiem jest uznanie czegoś za dane (Boyd i Crawford, 2011, s. 5). Pretensje do obiektywizmu i dokładności tylko dlatego, że dane istnieją i jest ich dużo, są zatem niezasadnione i mylące.

Szczególnie w przypadku pozyskania dużej ilości danych pierwszym etapem umożliwiającym dalszą analizę jest ich przygotowanie, tzw. czyszczenie danych (*data cleaning/preprocessing*). To fundamentalna część pracy *data scientisty*, określane zwyczajowo jako pracochłonna, zajmująca około 80% czasu realizacji projektów (CrowdFlower, 2017, s. 5; Kaggle, 2017; Wickham, 2014, s. 1). Etap ten jest jednak kolejną interpretacją danych (Boyd i Crawford, 2011, s. 5).

Więcej danych = większa wartość

Big data, epoka petabajtów to określenia akcentujące wielkość danych. To, że wielkość nie świadczy o wartości, np. nie uprawnia do uogólniania wyników z próby na populację, jest uznane w metodologii badań społecznych od dziesięcioleci (Cain i Finch, 1981). Znany jest przykład pierwszego sondażu przedwyborczego Gallupa. Przyniósł on w latach 30. XX w. lepsze rezultaty niż badanie na większej próbie, gdzie respondentami byli wyłącznie prenumeratorzy „Literary Digest” (Babbie, 2003). Ponadto dane zbierane automatycznie w dużych ilościach mogą zawierać liczne błędy i braki (Pink i in., 2018). Przekonanie, że więcej danych oznacza lepsze analizy, to zatem kwestia etosu (Boyd i Crawford, 2011, s. 6). Może to prowadzić do błędu zwanego apofenią (*apophenia*) — dostrzegania nieistniejących wzorców, które pojawiają się w danych przypadkowo, tylko ze względu na ich ilość (Boyd i Crawford, 2011, s. 2).

Na to samo wskazują osoby zajmujące się DS. Odkrycie związku między zmiennymi może odzwierciedlać sposób funkcjonowania świata, jak i być przypadkowe: „Liczba istotnych relacji między elementami zbioru danych (...) jest o całe rzędy wielkości mniejsza [niż relacji pozornych]. Nie różni też tak szybko jak ilość dostępnych informacji: ilość prawdy na świecie nie zmieniła się tak bardzo od czasu wynalezienia Internetu, a nawet prasy drukarskiej. Większość danych to zwykły **szum** [podkreślenie oryginalne], podobnie jak większość wszechświata stanowi pusta przestrzeń” (Silver, 2014, s. 234–235).

Zatem małe dane mogą z powodzeniem prowadzić do bardziej wartościowych rezultatów niż dane wielkie (Bornakke i Due, 2018; Faraway i Augustin, 2018; Jifa i Lingling, 2014; Kitchin i Carrigan, 2014; PTBRiO, 2017). Należy jednak pamiętać, że stosowane w DS metody ML nie generalizują w oparciu o próbę reprezentatywną. Model wyuczony metodą prób i błędów na jednym zbiorze testuje się na innym zbiorze danych (Larose, 2005). Badacze społeczni wskazują, że do zastosowań biznesowych korzystanie z danych behawioralnych, niedeklaracyjnych — czyli analizy typu big data/DS — może dać rezultaty lepsze z punktu widzenia praktyki rynkowej niż sondaże na próbach reprezentatywnych, głównie dzięki większej liczbie bardziej szczegółowych danych (Savage i Burrows, 2007).

Nie znaczy to jednak, że zawsze wraz z ilością danych wzrasta wartość ich analizy. Choć od strony technicznej podkreśla się rolę ilości danych w tzw. uczeniu modeli ML, to w DS znany jest problem obciążenia (bias) wyników działania modelu z uwagi na specyfikę zbioru danych (USACM, 2017; Chouldechova, 2017; Dunson, 2018; Shah i in., 2015).

Tylko narzędzia vs. aż narzędzia

Mówi się o dwóch kulturach modelowania zjawisk czy dwóch podejściach do nauki w ogóle (Breiman, 2001; Ceri, 2018; Manhart, 1996). Przeciwwstawia się sobie naukę opartą na teoriach (*formal science model-driven*) nauce opartej na danych (*data science data-driven*). Pierwsza ma polegać na wyjściu od teorii, postawieniu hipotez i testowaniu ich za pomocą danych zebranych w badaniu, druga zaś na analizie danych bez wcześniejszych pytań i założeń (przy czym jedno podejście nie wyklucza drugiego) (Ceri, 2018, s. 69). Tym samym big data/DS to niewątpliwie inne niż naukowe narzędzia poznawcze, lecz nie „tylko” narzędzia, ale „aż” narzędzia. Przypomina to tezę Marshalla McLuhana: „kształtujemy nasze narzędzia, a następnie one kształtują nas (*We shape our tools and thereafter they shape us*)” (Culkin, 1967, s. 70).

Narzędzia badawcze i analityczne zmieniają sposób myślenia o badaniach społecznych i o teoriach społecznych (Boyd i Crawford, 2011, s. 3). Ten nowy sposób myślenia określano krytycznie jako „uwodzicielski pseudopozytywizm” (Dalton, Taylor i Thatcher, 2016, s. 6) czy łagodniej „wywodzący się z idei pozytywistycznej” (Kitchin, 2014, s. 7), a także „odnowiony naturalizm”, w którym społeczeństwo traktowane jest jak fenomeny przyrody — lawiny śnieżne, ruchy ławic ryb, interakcje komórek (Törnberg i Törnberg, 2018, s. 2). Wnioskowanie o ludziach wyłącznie w optyce mierzenia ich zachowania może mieć nie tylko konsekwencje poznawcze (jak błędne czy powierzchowne wnioski), ale także etyczne, do czego wrócę w podsumowaniu.

Wiarą w dane

W entuzjastycznej pracy pt. *Big Data. Rewolucja, która zmieni nasze myślenie, pracę i życie* ukuto termin danetyzacja (*datafication*), który oznacza zmianę w dane różnych rodzajów ludzkiej aktywności (Cukier i Mayer-Schönberger, 2014). Biznes i rządy, korzystając z danych i metadanych serwisów, np. wyszukiwarki Google, portali Facebook, Twitter, LinkedIn, Tumblra, iTunes, komunikatorów Skype, WhatsApp, serwisu YouTube i darmowych platform poczty elektronicznej jak Gmail czy Hotmail, mają dostęp do zapisu wycinka ludzkich czynności, które jeszcze 20 lat temu były rejestrowane w małej skali bądź wcale. „Dane i metadane, nie tak dawno traktowane jako bezwartościowy produkt uboczny serwisów internetowych, zostały stopniowo zmienione w cenny zasób, który można wydobywać, wzbogacać i wykorzystywać ponownie w dochodowych produktach” (van Dijck, 2014, s. 199).

Danetyzacja jest „dążeniem do zawężania obszarów, które nie podlegają ewidencji” (Iwasiński, 2016, s. 137). Iwasiński przywołuje słowa Lwa Manowicha (2012, s. 335), który uważa, że efektem danetyzacji będzie zamiana świata w jedną wielką bazę danych. Pisano o „świadomości big data”, tzn. „przekonaniu, że istnieje mierzalny komponent wszystkiego” i że da się „przekształcić niezliczone wymiary rzeczywistości w dane” (Cukier i Mayer-Schönberger, 2014, s. 132). Zdaniem José van Dijck danetyzacja jest przejawem ideologii dataizmu (*dataism*). Dataizm zasadza się na wierze w obiektywność kwantyfikacji i zaufaniu do agentów zbierających dane oraz przekonaniu, że dane są najważniejszym komponentem poznania rzeczywistości (van Dijck, 2014, s. 198). Yuval N. Harari wskazuje na dataistyczny pogląd, że każdy układ czy zjawisko sprowadza się do przetwarzania danych. W konsekwencji „dataiści są sceptyczni wobec ludzkiej wiedzy czy mądrości, skłaniając się do zaufania w big

data i algorytmy komputerowe” (Harari, 2017, s. 213–214).

Dataizm czy ideologia danetyzacji nie cechuje wyłącznie analityków. Zaczniemy od spostrzeżenia, że użytkownik usług Google czy portalu Facebook nie jest ich klientem, choć firmom tym przeświecają jakoby nastawione na użytkownika misje „uporządkowania światowych zasobów informacji tak, by stały się powszechnie dostępne i użyteczne dla każdego” (Google, 2019) i „wspierania ludzi w budowaniu wspólnoty i łączeniu świata ze sobą” (Facebook, 2019). W ich modelu biznesowym użytkownicy nie płacą za usługi, bo nie są klientami, tylko produktami. Produktami, które firmy sprzedają reklamodawcom i innym podmiotom zainteresowanym uwagą użytkowników czy danymi o nich (Surma, 2017, s. 66–67, 70; Szymielewicz i Iwańska, 2019).

Podkreślam, że nie ma tu mowy o działaniu wbrew użytkownikom. Prowadzi to do konstatacji, że ideologia danetyzacji jest, jak stwierdziła van Dijck, oparta na powszechnych normach społecznych. W sondażu mniej niż 1/3 użytkowników byłaby skłonna płacić za owe usługi, gdyby firma je dostarczająca miała przestać korzystać z ich danych na potrzeby marketingowe. Tym samym, ideologia danetyzacji jest najczęściej przyjmowana bezkrytycznie przez wszystkie zainteresowane strony (van Dijck, 2014, s. 200–201), co nie znaczy, że staje się ona czymś innym, niż ideologią. Przypisywane W.E. Demingowi słowa „wierzymy w Boga, inni muszą pokazać dane (*In God we trust. All others bring data*)” (Gutierrez, 2014, s. 102) rozumiem zatem jako „wierzymy w dane”.

Moda na dane — zwrot obliczeniowy

Skuteczność omawianych tu pułapek data-driven sprawia, że mamy do czynienia z tzw. zwrotem obliczeniowym (*computational turn*) dotyczącym myślenia i prowadzenia badań. Dzieje się tak nie tylko w biznesie. Akademia i rządy także dążą do wydobywania z danych maksymalnej użyteczności (Boyd i Crawford, 2011, s. 13). Zarówno rządy, biznes, jak i nauka traktują dane jak Świętego Graala, prawdziwe odzwierciedlenie ludzkiego zachowania zebrane za pomocą przezroczystych termometrów, zatem najbardziej wiarygodne, a co za tym idzie użyteczne źródło informacji (van Dijck, 2014, s. 199). Mówi się o „logice cywilizacji numerycznej” — istnieje to, co jest kwantyfikowalne, a musi być takie przynajmniej, by możliwa była wycena w pieniądzu (Krzysztofek, 2012, s. 225–226). Zwrot obliczeniowy porównuje się do zmiany społecznej w konsekwencji produkcji masowej Forda. Tak jak taśmowy system produkcji samochodów zmienił rozumienie pracy, związku człowieka z pracą i społe-

czeństwo w szerokim sensie w XX w., tak w wieku XXI zmiany społecznej ma dokonać zwrot ku obliczeniowemu rozumieniu świata. Dokonuje się on za pomocą dużej ilości danych i narzędzi do pracy z nimi (Boyd i Crawford, 2011, s. 3; 2012, s. 665).

J. Surma (2017, s. 23) manifestem tzw. obliczeniowych nauk społecznych nazwał pracę Davida Lazera z zespołem (2009). Zauważono, że nowe możliwości zbierania i analizy danych dotyczących ludzi są już przedmiotem pracy badawczej firm (wymieniono Google i Yahoo) i agencji rządowych USA (wskazano National Security Agency, NSA²), zatem jeżeli akademia nie chce pozostawić innym podmiotom tego rodzaju prac, musi działać (Lazer i in., 2009, s. 721).

W pracach obliczeniowych nauk społecznych widoczny jest entuzjazm wobec stosowanej metodologii i epistemologii. Pojawiały się omówione wyżej pretensje do prawdziwości, obiektywności, pewności, słuszności i użyteczności uzyskiwanych wyników. Pisano o „eksploracji rzeczywistości” (Eagle i Pentland, 2006), „sygnałach prawdy” jako „twardych miarach” zachowań społecznych (Arena, Pentland i Price, 2010; Pentland i Heibeck, 2008), analizach o bezprecedensowym zasięgu, głębi i skali (Lazer i in., 2009, s. 722), inżynierii społecznej dla lepszego świata (Eagle i Greene, 2014), uzyskaniu „punktu widzenia Boga na nas samych (*God's eye view of ourselves*)” (Pentland, 2009, s. 80). Fizyk Albert-László Barabási wyrażał nadzieję odkrycia ścisłych, matematycznych praw opisujących i pozwalających prognozować ludzkie zachowanie dzięki analizie obiektywnych danych (Krzysztofek, 2011, s. 126).

Cześć obliczeniowych nauk społecznych poddała krytyce własne założenia i metody. Pisano o sprzeciwie wobec „danych mówiących samych za siebie” (Chang, Kauffman i Kwon, 2014) czy o pułapce pozornego obiektywizmu (Eder, 2014). Tak czy inaczej, tego rodzaju podejście jest obecnie bardzo modne i postrzegane jako konieczne w różnych dziedzinach życia.

Z punktu widzenia data scientists

Na podstawie badań własnych sędzę, że osoby zajmujące się zawodowo DS postrzegają oczekiwania klientów (tzw. biznesu) wobec rozwiązań opartych na danych jako wygórowane lub naiwne. Ich zdaniem klienci z niewielkim zapleczem wiedzy technicznej (informatycznej i statystycznej), wiedzeni modą napędzaną obietnicami marketingu produktów i usług analitycznych, widzą DS jako magię. Magię, za pomocą której rozwiązany zostanie dowolny problem biznesowy. Data scientist przeciwnie, widzi wdrożenie rozwiązania opartego na danych jako majsterkowanie, na które składa się

szereg działań technicznych, prawnych, organizacyjnych, analitycznych. W poniższym fragmencie zwracam uwagę na zwrot „dane w Excelu”. Excel w środowisku DS jest postrzegany jako najprostsze, bardzo ograniczone narzędzie analityczne. Zdecydowanie nie jest to narzędzie do DS. Szczególnie przechowywanie danych w Excelu — o czym tu mowa — traktowane jest jako nieprofesjonalne.

Rozmówca: (...) *Coś co może też jest trochę śmieszne i też potrafi przeszkadzać jest są z kolei wydumane yy i rozbuchane oczekiwania, biznesu tak? Się naczytają nastuchają się tam na konferencjach że tam ten data mining i data science w ogóle jak Midas przyjdzie dotknie rękę i to wszystko się rozwiąże, natomiast bardzo rzadko na tych konferencjach mówi się o tym że **owszem**³, tak może być, ale żeby to się wydarzyło, to i to trzeba wykonać. I to jest proces, i to są lata, i to są duże sumy, wchodzą, i tej świadomości często klienci nie mają. Mówią: słuchajcie to jest problem, macie tutaj dane, w Excelu coś tam, i w ogóle za trzy miesiące niech to będzie wszystko automatycznie, i niech jeszcze odbiera telefony tak? (...) Wysokie oczekiwania, bardzo, tak jak mówię samo w byłoby super, natomiast nie ma tej drugiej, bardzo rzadko jest świadomość (...) [wywiad 6]⁴.*

Cześć klientów fetyszyzuje najmodniejsze terminy: DS, ML, AI. Modne narzędzia i metody są przez klientów utożsamiane z uzyskaniem magicznego rozwiązania problemu biznesowego. Brak użycia przez data scientista w komunikacji z klientami tych fetyszyzowanych słów kluczy może skutkować brakiem zaufania do jego profesjonalizmu. Inna część klientów „nie ma kłapek na oczach” i nie zwraca uwagi na narzędzia/metody (akceptuje niemodną statystykę), tylko na rozwiązanie problemu.

Rozmówca: (...) *właściwie w tej chwili doszło do tego, że jak się nie napisze że jest się data scientistem i jak się nie napisze że się używa machine learningu i sztucznej inteligencji toooo ludzie **nie wierzą** że się potrafi usprawnić ich biznes przy wykorzystaniu danych (...). Badacz: Ale czy to znaczy, że teraz klienci oczekują takiej nazwy, data science? Rozmówca: Ooo, to jest mi trudno powiedzieć. Wydaje mi się, że tak, booo to jest takie słowo klucz, albo taki termin klucz że aha, data science jest to taki gość, że mmmm, ma jakieś tam takie skomplikowane narzędzia i tymi narzędziami coś tam sobie **kręci**, wychodzą z tego modele i te modele automagicznie usprawniają biznes. Najlepiej jednak, żeby to była sztuczna inteligencja, i **część** klientów tego oczekuje. Żeby była sztuczna inteligencja, żeby było żeby była informacja o machine learning: „tak, my usprawnimy twój biznes z wykorzystaniem AI”. Natomiast ci klienci, którzy wiedzą, że chodzi o podejmowanie decyzji w oparciu o dane i to nie do końca ma znaczenie, jak wykorzystamy te dane, czy tam jest ta sztuczna inteligencja czy tam jest **regresja linio-***



wa, a może zrobimy zwykłą tabelkę, ci klienci, którzy to wiedzą, no tooo [pauza] nie mają kłapek na oczach i nie sądzą że jak im zrobimy AI, to załatwi ich wszystkie problemy [wywiad 5].

Na początku 2019 r. w środowisku DS popularny stał się żart wyjaśniający różnicę między ML a AI, co nazywam różnicą pomiędzy majsterkowaniem a magią: „Jeżeli jest napisane w Pythonie, to prawdopodobnie ML; jeżeli w PowerPoincie, to prawdopodobnie AI” (Alekseichenko, 2019a; 2019b). Żart ten pokazuje, że środowisko techniczne odcina się od nietechnicznej narracji o produktach i usługach analitycznych. Dla nich AI to coś, co istnieje tylko jako termin marketingowy na slajdach prezentacji. W opozycji techniczny termin ML istnieje jako metody możliwe do wykonania w Pythonie — najpopularniejszym w DS języku programowania (Gagné, Karmanov i Hudson, 2018; Owczarek, 2018; Piatetsky, 2017; Strong, 2018).

Środowisko DS nie kwestionuje sensowności podejścia data-driven w biznesie, raczej wierzy w dane i stara się to podejście rozpowszechniać. Zainteresowanie biznesu podejściem data-driven może być jednak postrzegane w środowisku jako powierzchowne. Dane bywają fetyszem pozwalającym np. na pozornie obiektywną perswazję. Moja rozmówczyni zaznaczyła, że spotyka się z niechęcią biznesu do poświęcenia czasu i uwagi na pogłębione analizy.

Rozmówca: (...) *żeby analityka była w sercu, jakby samej koncepcji że każde działania są mierzone i trzeba je oceniać i można badać jak wpływają bezpośrednio już na doświadczenia pojedynczego użytkownika. Chciałabym, właśnie staram się zaszczepiać to też w innych ludziach z firmy z innych działów, więc okazało się że ta praca wewnątrz firmy też zajmuje mi sporo czasu i sporo część mojego dnia w pracy. Badacz: Rozumiem, czyli takie nastawienie na analitykę starasz się przekonać innych w swojej firmie, tak? Do nastawiania na na analitykę? Rozmówca: Tak, i z jednej strony jest to łatwe bo wszystko jest teraz w data driven i każdy chce być data driven, a z drugiej strony to jednak jest ciężka praca, żeby chwilę czasu swojego poświęcić jeszcze na wejście dodatkowego narzędzia i sprawdzanie jakiś liczb, zastanowienie się dlaczego czasem one się nie zgadzają z innym narzędziem, więc ja rozumiem że jest jakaś taka niechęć, żeby to robić, ale wydaje mi się że warto, warto. Sprawdzać Badacz: Ale to też yyy mówisz że z jednej strony ludzie chcą, tak? Że jest jakaś taka moda na data driven czy jak to się tam nazywa. Rozmówca: Jak najbardziej chcą bo to też lepiej wygląda jeżeli, mamy działania i wspieramy efekt działań danymi a nie tylko mówimy że coś zrobiliśmy, wkleiliśmy te kody na stronie, śledzą, wysyłamy te reklamy, a klient nie widzi u siebie jakby znaczącego wzrostu przychodów, więc dobrze sobie czasem **poprzec** że na przykład dany kanał, właśnie za który płacimy może nie*

*powoduje że ludzie od razu wchodzi z niego i kupują ale zostaje gdzieś ludziom w pamięci i dopiero po jakiś czasie wracają do tego. Więc ta analityka też się czasem przydaje w tłumaczeniu różnych spraw. (...) W agencjach marketingowych, jakby oni wszyscy bardzo **chcą** być data driven i robić zaawansowane rzeczy ale [pauza] ale tu się liczy **czas**, że jak czegoś nie zrobisz w kilka godzin to już, nie chce się, po co marnować czas na to. Więc trzeba bardzo dużo przekonywać ludzi, że optaca się czasem dać analitykowi 2 tygodnie spokoju żeby w spokoju pomyśleć co zrobić [wywiad 17].*

Twierdzę, że środowisko DS w niewielkim stopniu — poza wiarą w dane — wpada w omawiane tutaj pułapki myślowe podejścia data-driven. Raczej wpadają w nie klienci, czyli osoby korzystające z produktów i usług analitycznych, ponieważ te pułapki są właśnie na nie zastawiane. Strategie marketingowe dotyczące „magicznych” produktów i usług analitycznych bazują z jednej strony właśnie na owych pułapkach, z drugiej na obietnicach dotyczących poprawy wyników finansowych, szybkości, wygody i efektywności uzyskanej dzięki wdrożeniu takich rozwiązań (por. Peltarion, 2018). Dodatkowo brak wiedzy technicznej i presja czasowa praktyki biznesowej sprzyjają skuteczności omawianych pułapek.

Propozycja: (ilość + jakość + teoria) × krytyka

Co się stanie, gdyby połączyć wielkie dane (*big data*) i grube dane (*thick data*)? Autorką określenia „grube dane” jest Tricia Wang, socjolożka stosująca metody etnograficzne w badaniach marketingowych (Wang, 2013; 2016). Oznacza ono materiały zbierane za pomocą metod i technik badań jakościowych. Łączenie wglądów (*insights*), jakie dają analiza danych ilościowych typu big data i analiza materiałów jakościowych, ma prowadzić do osadzenia wyników analizy ilościowej w kontekście, zatem ułatwić ich wyjaśnianie i rozumienie (Bornakke i Due, 2018, s. 1–4).

Autorzy podają m.in. przykład badania Wendy F. Hsu, która zajmowała się niezależną sceną muzyki rockowej osób pochodzących z Azji w USA. Hsu używała technik jakościowych: wywiadów, obserwacji, w tym obserwacji mediów społecznościowych i stron WWW. Zauważyła na MySpace, że zespoły w USA mają fanów np. z Azji. Pobrała dane (*web scraping*) o lokalizacji fanów z MySpace, co pozwoliło poszerzyć wnioski i pokazać zasięg geograficzny społeczności fanów badanej muzyki (Hsu, 2014).

Także w popularnonaukowym badaniu dotyczącym przekraczania granic intymności kobiet bawiących się w nocnych klubach w Brazylii wyraźnie widać, że do-

piero osadzenie analizy ilościowej w kontekście prowadzi do zrozumienia problemu. Badaczki ubrane w sukienkę wyposażoną w niewidoczne sensory dotknięć uczestniczyły w imprezach, niejawnie nagrywano także materiał audio i wideo. Choć w raporcie wideo podkreślona jest rola sukienki z sensorami, to nie wykresy obrazujące ilość dotknięć, ale nagrane interakcje badaczek z mężczyznami pozwalają zrozumieć problem, o czym świadczą zaprezentowane w materiale reakcje odbiorców badania (Ogilvy, 2018).

O komplementarności badań etnograficznych i DS przekonuje także Heather Ford, która we współpracy z data scientistami badała powstawanie artykułów na Wikipedii (Ford, 2014; 2016). Powstała już pierwsza monografia dotycząca omawianego zagadnienia. Przy współpracy data scientistów i badaczy społecznych zawarto tam porady metodologiczne i praktyczne oraz przykłady badań (Knox i Nafus, 2018).

Pisze się o potrzebie łączenia etnografii z metodami typu big data na zasadzie triangulacji metod. „Bogactwo analityki danych [ilościowych] zwiększa, a nie zmniejsza zapotrzebowanie na uzupełnianie ich o wyniki badań etnograficznych” (Jemieliński, 2018, s. 17). Także Kitchin w omawianym artykule postulował ostrożne korzystanie z możliwości, jakie daje big data/DS, obok tradycyjnego w naukach społecznych podejścia do badań ilościowych i jakościowych. Mówił on o potencjalnej strategii, która miałaby łączyć w sobie indukcję, dedukcję i abdukcję. Chodzi o naukę opartą na danych (*data-driven science*) w tym sensie, że dzięki eksploracyjnej analizie dużych lub nieustrukturyzowanych danych można generować pytania i hipotezy dla kolejnych etapów badania osadzonego w ramach teorii dziedzinowej (Kitchin, 2014, s. 5–6). Inni autorzy, także zainspirowani m.in. omawianym artykułem Kitchina, mówią o „symfonicznych naukach społecznych”, postulują — podobnie jak on — abdukcyjne podejście, upatrując w nim przyszłość tych nauk (Halford i Savage, 2017, s. 1145).

Uważam, że tego rodzaju podejście do badania ludzi (łącznie tradycję społecznych badań ilościowych, jakościowych i tzw. big data) możliwe jest zarówno w nauce, jak i w dziedzinach praktycznych, np. marketingu. Niemniej musi, moim zdaniem, być ono krytyczne, czyli świadome wyżej omówionych zagadnień, które nazywam pułapkami myślowymi. Powtarzam, że:

- dane nie mówią same za siebie,
- dane nie są „dane”, tylko zbierane i generowane,
- większa ilość danych nie musi oznaczać większej wartości (finansowej ani poznawczej),
- stosowane narzędzia kształtują badany wycinek rzeczywistości,
- wiara w dane jest wiarą niemożliwą do zasadniczo logicznego uzasadnienia,
- panuje moda na korzystanie z danych i najnowocześniejszych metod obliczeniowych.

Podsumowanie

Pomimo wpadania w omawiane pułapki, podejście data-driven bez wątpienia może przynosić pożądane efekty poprawy wyników finansowych. Wiąże się to jednak z konsekwencjami społecznymi. Wbrew opiniom laików i entuzjastów dane oraz ich analiza nie są obiektywne, bezstronne czy sprawiedliwe — są narzędziem w ludzkich rękach, a „modele są opiniami opisanymi w języku matematyki” (O’Neil, 2017, s. 46–48). Pułapki tu opisywane sprawiają, że pod pozorem matematycznej racjonalności, efektywności i bycia na bieżąco z najnowszymi technologiami dochodzi np. do wzmacniania i utrwalania nierówności płci czy rasowych (Crawford, West i Whittaker, 2019; Crawford i in., 2018; Grush, 2015; Julia i in., 2016; Merity, 2017; Richardson, Schultz i Crawford, 2019) oraz rozszerzania i wzmacniania praktyk inwigilacji (Botsman, 2017; Crawford i in., 2018, s. 12–18; Symeonidis, Tsormpatzoudi i Preneel, 2015; Zuboff, 2015). Ponieważ magia danych czy tzw. AI może zachwycać i onieśmielać, szczególnie osoby nietechniczne, to oprócz lektury krytycznych prac zachęcam do uczenia się statystyki, metodologii badań, programowania i ML⁵, żeby odczarować te dziedziny.

Pseudopozytywizm podejścia data-driven bazuje, w mojej ocenie, na traktowaniu ludzi w kategoriach zachowania (*behaviour*). Zachowanie jest przedmiotem pomiarów i akcji oraz źródłem dochodu. W naukach prawnych już pod koniec lat 90. XX w. dostrzegano zbliżony problem pozbawiania ludzi indywidualności w wyniku działania systemów opartych na analizie danych (Vedder, 1999). Stwierdzono też, że ponieważ analitycy pracują z danymi, to definiują się jako mający styczność z matematyką, a nie z ludźmi (Metcalf i Crawford, 2016, s. 3). Jednak szczególnie w przypadku marketingu mierzy się i wykonuje akcje na ludziach. Czasami człowieczeństwa tych „przedmiotów” nie dostrzega także tzw. biznes, co widać w poniższym fragmencie.

Badacz: *A czy dobrze zrozumiałem, że uważasz, że ta branża marketingowa jest, jest trochę nieetyczna?* Rozmówca: *Tak.* Badacz: *Tak? A co w niej jest takiego?* Rozmówca: *Może nie bezpośrednio i nie wszyscy, ale czasami jak się idzie na konferencję branżową, i **śłuchasz** tego, **śłuchasz** tego jak się ludzie wypowiadają, na prezentacjach, i często łapiasz się na tym, że oni mówią o użytkownikach jakby to jakiś **osobny byt**. To nie są ci **ludzie** co my, to nie są ludzie tylko [pauza] użytkownicy, **lead**y. [śmiech] *Że my tam wysyłamy na nich działania, i jakoś nikt nie wspomni o tym, że my **też** jesteśmy klientami (...)* [wywiad 17].*

Postuluję znane w socjologii traktowanie tego, co ludzie robią (jako jednostki i jako zbiorowości), w kategoriach działania (*action*), a nie zachowania. Jest to podejście podmiotowe, humanistyczne, akcentują-

ce ludzką sprawczość (*agency*) oraz symetrię badających i badanych⁶. Sądzę, że stoi ono w zgodzie z proponowanym wyżej połączeniem społecznych badań ilościowych, jakościowych i tzw. big data.

Nie zależy mi na obronie pozycji socjologów stosujących podejście jakościowe (do których należę).

Skoro w tym roku wolny rynek w Polsce kończy 30 lat i wypada 25. rocznica zamieszczenia pierwszego banneru reklamowego w Internecie (Szymielewicz i Iwańska, 2019, s. 4), to zależy mi na tym, by biznes był nie data-driven, ale knowledge-driven i przede wszystkim human-driven.

Przypisy

¹ Badania etnograficzne zrealizowano dzięki finansowaniu ze środków Katedry Socjologii Kultury, Instytut Socjologii, Wydział Ekonomiczno-Socjologiczny Uniwersytetu Łódzkiego. Na dzień 04.05.2019 autor przeprowadził 21 wywiadów swobodnych ukierunkowanych (łącznie 21 godz. 42 min) i 39 obserwacji uczestniczących (łącznie około 215 godz.). Badania te są podstawą empiryczną rozprawy doktorskiej pt. *Data science w Polsce. Etnografia społecznego świata*. Autor posługuje się metodą etnografii wielostanowiskowej (Marcus, 1995), a jako ramę teoretyczną przyjmuje teorię światów społecznych (Strauss, 1978), ze szczególnym naciskiem na koncepcję społecznych światów/aren i analizę sytuacyjną (Clarke, 1991; 2005; 2017). Rozmówcy dobierani są do wywiadów celowo, zgodnie ze strategią próbkowania teoretycznego (*theoretical sampling*) (Glaser i Strauss, 1967). Autor składa podziękowania wszystkim osobom uczestniczącym w badaniach.

² W 2013 r. Edward Snowden ujawnił tajne dokumenty dotyczące zbierania cyfrowych danych o ludziach właśnie przez NSA (van Dijck, 2014).

³ W cytowanych fragmentach wywiadów etnograficznych pogrubieniem zaznaczam słowa wypowiedziane z naciskiem. Transkrypcje są dosłowne (verbatim).

⁴ Ten i następane wywiady cytuję z wspomnianych w przypisie pierwszym badań własnych.

⁵ Dostępnych jest wiele podręczników i kursów, także bezpłatnych i w języku polskim (Biecek, 2015b; 2015a; Dasgupta, 2018; DataCamp, 2018; Wickham i Grolemund, 2017).

⁶ Charakteryzuje ono paradygmat interpretacyjny (por. Hałas, 1987, s. 19–20; Konecki, 2000, s. 17–23; Krzemiński, 1986, s. 7–12; Teddlie i Tashakkori, 2009, s. 86).

Bibliografia

- Alekseichenko, V. (2019a). *10 mitów o sztucznej inteligencji*. Pozyskano z <http://biznesmysli.pl/10-mitow-o-sztucznej-inteligencji/>
- Alekseichenko, V. (2019b). *The difference between AI vs ML*. Pozyskano z <https://www.linkedin.com/feed/update/urn:li:activity:6501030890754314240>
- Anderson, C. (2008). *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*. Pozyskano z <https://www.wired.com/2008/06/pb-theory/>
- Arena, M. J., Pentland, A. i Price, D. (2010). Honest Signals — Hard Measures for Social Behavior. *Organization Development Journal*, 28(3), 11–20. Pozyskano z <http://www.questia.com/library/journal/1P3-2127733261/honest-signals-hard-measures-for-social-behavior>
- Artun, O., i Levin, D. (2015). *Predictive Marketing: Easy Ways Every Marketer Can Use Customer Analytics and Big Data*. New York: Wiley.
- Babbie, E. (2003). *Badania społeczne w praktyce*. Warszawa: Wydawnictwo Naukowe PWN.
- Barton, D. i Court, D. (2013). *Three keys to building a data-driven strategy*. Pozyskano z <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/three-keys-to-building-a-data-driven-strategy>
- Biecek, P. (2015a). *Pogromcy Danych. Przetwarzanie danych w programie R*. Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego. Pobrano z <http://pogromcydanych.icm.edu.pl/>
- Biecek, P. (2015b). *Pogromcy Danych. Wizualizacja oraz modelowanie danych*. Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego. Pobrano z <http://pogromcydanych.icm.edu.pl/>
- Błażewicz, G. (2016). *Revolucja z marketing automation: jak wykorzystać potencjał Big Data*. Warszawa: Wydawnictwo Naukowe PWN.
- Bornakke, T. i Due, B. L. (2018). Big-Thick Blending: A method for mixing analytical insights from big and thick data sources. *Big Data & Society*, 5(1), 1–16. <https://doi.org/10.1177/2053951718765026>
- Botsman, R. (2017). *Big data meets Big Brother as China moves to rate its citizens*. Pozyskano z <http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion>
- Bowker, G. C. (2005). *Memory Practices in the Sciences*. Cambridge: MIT Press.
- Boyd, D. i Crawford, K. (2011). Six Provocations for Big Data. *SSRN Electronic Journal*, 1–17. <https://doi.org/10.2139/ssrn.1926431>
- Boyd, D. i Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, 16(3), 199–231. <https://doi.org/10.2307/2676681>
- Brynjolfsson, E., Hitt, L. M. i Kim, H. H. (2011). Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1819486>
- Cain, M. i Finch, J. (1981). Towards a Rehabilitation of Data. W: *Practice and Progress: British Sociology 1950–1980*. London: George Allen and Unwin.
- Ceri, S. (2018). On the role of statistics in the era of big data: A computer science perspective. *Statistics & Probability Letters*, 136, 68–72. <https://doi.org/10.1016/j.spl.2018.02.019>
- Chang, R. M., Kauffman, R. J. i Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems*, 63, 67–80. <https://doi.org/10.1016/j.dss.2013.08.008>
- Chouldechova, A. (2017). Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big Data*, 5(2), 153–163. <https://doi.org/10.1089/big.2016.0047>
- Chudziak, J. i Zalewska, K. (2018). *Trendy, które zmieniają brand marketing*. Pozyskano z <https://nowymarketing.pl/a/18202,trendy-ktore-zmieniaja-brand-marketing>
- Clarke, A. E. (1991). Social Words / Arenas Theory as Organizational Theory. W *Social Organization and Social Process. Essays in Honor of Anselm Strauss* (s. 119–158). New York: Aldine de Gruyter.
- Clarke, A. E. (2005). *Situational Analysis. Grounded Theory After the Postmodern Turn*. London: Sage.

- Clarke, A. E., Friese, C. i Washburn, R. S. (2017). *Situational Analysis: Grounded Theory After the Interpretive Turn*. Los Angeles: Sage.
- Crawford, K., West, S. M. i Whittaker, M. (2019). *Discriminating Systems: Gender, Race and Power in AI*. New York. Pozyskano z <https://ainowinstitute.org/discriminatingystems.pdf>
- CrowdFlower. (2017). *2017 Data Scientist Report*. Pozyskano z https://visit.crowdfunder.com/rs/416-ZBE-142/images/data-scientist-report-dec.pdf?mkt_tok=eyJpJjoiWXPkNk5EQmtNaJkTkdWayIsInQiOiJpY29MV2JjU1RlRhbGh6OUVWcmUwUWpibXJ3cG5pSlFrNjUwUWUwT2hna1VOOU5hGd2tMU3ZEWmhoTVVvVHRXNWVhMFM4eTI1dDJBWWRzVoTVlnRjFkQj14ekNmT
- Cukier, K. i Mayer-Schönberger, V. (2014). *Big Data. Rewolucja, która zmieni nasze myślenie, pracę i życie*. Warszawa: MT Biznes.
- Culkin, J. M. (1967). A Shoolman's Guide to Marshall McLuhan. *The Saturday Review*, 51–53; 70–72.
- Czech, M. (2018). *6 najczęstszych błędów podczas implementacji chatbota i jak ich uniknąć*. Pozyskano z <https://nowymarketing.pl/a/20278,6-najczestszych-bledow-podczas-implementacji-chatbota-i-jak-ich-uniknac>
- Dalton, C. M., Taylor, L. i Thatcher, J. (2016). Critical Data Studies: A dialog on data and space. *Big Data & Society*, 3(1). <https://doi.org/10.1177/2053951716648346>
- Dasgupta, S. (2018). *Machine Learning for Data Science*. Pozyskano z <https://www.edx.org/course/machine-learning-fundamentals-uc-san-diegox-dse220x>
- DataCamp. (2018). *DataCamp. The Smartest Way To Learn Data Science Online*. Pozyskano z <https://www.datacamp.com/>
- Davenport, T. H. i Patil, D. J. (2012). Data scientist: The sexiest job of the 21st century. *Harvard Business Review*. <https://doi.org/10.1109/MITP.2016.41>
- Dijck van, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance and Society*, 12(2), 197–208. <https://doi.org/10.24908/ss.v12i2.4776>
- Divan, M. J. (2018). *Data-driven decision making*. 2017 International Conference on Infocom Technologies and Unmanned Systems: Trends and Future Directions. *ICTUS 2017*, 50–56. <https://doi.org/10.1109/ICTUS.2017.8285973>
- Dunson, D. B. (2018). Statistics in the big data era: Failures of the machine. *Statistics and Probability Letters*, 136, 4–9. <https://doi.org/10.1016/j.spl.2018.02.028>
- Eagle, N. i Greene, K. (2014). *Reality Mining: Using Big Data to Engineer a Better World*. Cambridge, London: The MIT Press. Pozyskano z <http://www.jstor.org/stable/j.ctt9qf8q3>
- Eagle, N. i Pentland, A. (2006). Reality mining: Sensing complex social systems. *Journal Personal and Ubiquitous Computing*, 10(4), 255–268. <https://doi.org/10.1007/s00779-005-0046-3>
- Eder, M. (2014). Metody ścisłe w literaturoznawstwie i pułapki pozornego obiektywizmu — przykład stylometrii. *Teksty Drugie*, (2), 90–105.
- Evans, N. (2016). *Five Steps to Better Campaigns in Data-Driven Marketing*. Pozyskano z <https://www.cmo.com/features/articles/2016/3/25/five-steps-to-better-campaigns-in-datadriven-marketing.html#gs.5xr9fw>
- Facebook. (2019). *About*. Pozyskano z <https://www.facebook.com/pg/facebook/about/>
- Faraway, J. J. i Augustin, N. H. (2018). When small data beats big data. *Statistics & Probability Letters*, 136(May), 142–145. <https://doi.org/10.1016/j.spl.2018.02.031>
- Ford, H. (2014). Big Data and Small: Collaborations between ethnographers and data scientists. *Big Data & Society*, 1(2). <https://doi.org/10.1177/2053951714544337>
- Ford, H. (2016). *What does it mean to be a "participant observer" in a place like Wikipedia?* Pozyskano z <https://medium.com/ethnography-matters/what-does-it-mean-to-be-a-participant-observer-in-a-place-like-wikipedia-89d6727276ba>
- Gagné, J.-F., Karmanov, F. i Hudson, S. (2018). *Global AI Talent Pool Report*. Pozyskano z <http://www.jfgagne.ai/talent/>
- Gitelman, L. i Jackson, V. (2013). Introduction. W: L. Gitelman (red.), *"Raw Data" is an oxymoron*. Cambridge, London: The MIT Press. <https://doi.org/10.1080/1369118X.2014.920042>
- Glaser, B. G. i Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine.
- Golik-Górecka, G. (2017). Types of Big Data and their Use in Strategies of Marketing. *Marketing i Zarządzanie*, 49(3), 43–54. <https://doi.org/10.18276/miz.2017.49-04>
- Golik-Górecka, G. (2018). Innowacyjne dashboardy a zakres Big Data marketingu analitycznego. *Marketing i Rynek*, (8), 129–138.
- Google Akademia Analytics. (2019). *Google Analytics dla początkujących. 1. Wprowadzenie. 1.4. Konfigurowanie widoków za pomocą filtrów*. Pozyskano z <https://analytics.google.com/analytics/academy/course/6/unit/1/lesson/4?authuser=0>
- Google. (2019). *Informacje*. Pozyskano z <https://www.google.com/about/>
- Grush, L. (2015). *Google engineer apologizes after Photos app tags two black people as gorillas*. Pozyskano z <https://www.theverge.com/2015/7/1/8880363/google-apologizes-photos-app-tags-two-black-people-gorillas>
- Grygierek, R. (2017). *Koniec ery nieobliczalności. Technologia, marketerze!* Pozyskano z <https://nowymarketing.pl/a/13937,koniec-ery-nieobliczalnosci-technologie-marketerze>
- Gutierrez, S. (2014). *Data Scientists at Work: Sexy Scientists Wrangling Data And Begetting New Industries*. Berkeley: Appres.
- Halford, S. i Savage, M. (2017). Speaking Sociologically with Big Data: Symphonic Social Science and the Future for Big Data Research. *Sociology*, 51(6), 1132–1148. <https://doi.org/10.1177/0038038517698639>
- Halper, F. (2018). *Five Data Management and Analytics Best Practices for Becoming Data-Driven Consider a Cohesive Platform That Supports*. Pozyskano z <https://tdwi.org/research/2018/06/adv-all-checklist-five-data-management-and-analytics-best-practices-for-becoming-data-driven.aspx>
- Halas, E. (1987). *Spoleczny kontekst znaczeń w teorii symbolicznego interakcjonizmu*. Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego.
- Harari, Y. N. (2017). *Homo Deus: A Brief History of Tomorrow*. London: Vintage.
- Harris, J. (2012). Data Driven. W: R. Smolan i J. Erwitte (red.), *The Human Face of Big Data* (s. 200–203). Sausalito: Against All Odds Production.
- Hippold, S. (2018). *Build a Data-Driven Organization — Smarter With Gartner*. Pozyskano z <https://www.gartner.com/smarterwithgartner/build-a-data-driven-organization/>
- Hosni, H. i Vulpiani, A. (2018). Data science and the art of modelling. *Lettera Matematica*, 6(2), 121–129. <https://doi.org/10.1007/s40329-018-0225-5>
- Hsu, W. F. (2014). Digital Ethnography Toward Augmented Empiricism: A New Methodological Framework. *Journal of Digital Humanities*, 3(1). Pozyskano z <http://journalofdigitalhumanities.org/3-1/digital-ethnography-toward-augmented-empiricism-by-wendy-hsu/>
- Iwasiński, Ł. (2016). Społeczne zagrożenia danetyzacji rzeczywistości. W: B. Sosińska-Kalata (red.), *Nauka o informacji w okresie zmian. Informatologia i humanistyka cyfrowa* (s. 135–146). Warszawa: Wydawnictwo SBP.
- Iwasiński, Ł. (2017). Przyczynek do rozważań nad suwerennością konsumenta w epoce danetyzacji i big data. *Kultura — Historia — Globalizacja*, 21, 119–133.
- Jeffery, M. (2010). *Data-driven marketing: the 15 metrics everyone in marketing should know*. Hoboken: John Wiley i Sons.
- Jemielniak, D. (2018). Socjologia 2.0: o potrzebie łączenia big data z etnografią cyfrową, wyzwaniach jakościowej socjologii cyfrowej i systematyzacji pojęć. *Studia Socjologiczne*, 2(229), 7–29. <https://doi.org/10.24425/122461>
- Jeske, M. (2018). *Data-driven marketing. Poznaj i zrozum swojego klienta dzięki umiejętnemu wykorzystaniu danych*. Pozyskano z <https://www.bluerank.pl/data-driven-marketing-poznaj-i-zrozum-swojego-klienta-dzieki-umiejtnemu-wykorzystaniu-danych/>

- Jifa, G. i Lingling, Z. (2014). Data, DIKW, Big Data and Data Science. *Procedia Computer Science*, 31, 814–821. <https://doi.org/10.1016/J.PROCS.2014.05.332>
- Julia, A., Larson, J., Mattu, S. i Kirchner, L. (2016). *Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks*. Pozyskano z <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Kaggle. (2017). *The State of Data Science i Machine Learning 2017*. Pozyskano z <https://www.kaggle.com/surveys/2017>
- Karasek, D. (2018). *Przewidywanie przyszłości i sztuczna inteligencja dla każdego*. Pozyskano z <https://nowymarketing.pl/a/18613,przewidywanie-przyszlosci-i-sztuczna-inteligencja-dla-kazdego>
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1–12. <https://doi.org/10.1177/2053951714528481>
- Kitchin, R. i Carrigan, M. (2014). *Big data should complement small data, not replace them*. Pozyskano z <http://blogs.lse.ac.uk/impactofsocialsciences/2014/06/27/series-philosophy-of-data-science-rob-kitchin/>
- Knox, H. i Nafus, D. (red.). (2018). *Ethnography for a Data-saturated World*. Manchester: Manchester University Press.
- Konecki, K. T. (2000). *Studia z metodologii badań jakościowych. Teoria ugruntowana*. Warszawa: Wydawnictwo Naukowe PWN.
- Kozielski, R. (2016). Przyszłość marketingu — wymiar koncepcyjny, organizacyjny i implementacyjny. *Handel Wewnętrzny*, 4(363), 101–109.
- Krzemiński, I. (1986). *Symboliczny interakcjonizm i socjologia*. Warszawa: PWN.
- Krzysztofek, K. (2011). W stronę maszyn społecznych. Jaka będzie socjologia, której nie znamy? *Studia Socjologiczne*, 2(201), 123–145.
- Krzysztofek, K. (2012). Big Data Society. Technologie samozapisu i samopokazu: ku humanistyce cyfrowej. *Transformacje*, 1–4(72–75), 223–257.
- Kumar, V., Chattaraman, V., Neghina, C., Skiera, B., Aksoy, L., Buoye, A. i Henseler, J. (2013). Data-driven services marketing in a connected world. *Journal of Service Management*, 24(3), 330–352. <https://doi.org/10.1108/09564231311327021>
- Larose, D. T. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining*. Hoboken: John Wiley & Sons.
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D. i Van Alstyne, M. (2009). Social science. Computational social science. *Science*, 323(February), 721–723. <https://doi.org/10.1126/science.1167742>
- Łapińska, M. (2019). MarTech — Marketing Technology a rozwój platform SaaS do automatyzacji marketingu. *Marketing i Rynek*, 3(3). <https://doi.org/10.33226/1231-7853.2019.3.2>
- Mah, P. (2019). *3 steps to successful data-driven marketing*. Pozyskano z <https://www.enterpriseinnovation.net/article/3-steps-successful-data-driven-marketing-2011594387>
- Manhart, K. (1996). Artificial Intelligence Modelling: Data Driven and Theory Driven Approaches. W: K. Troitzsch, U. Müller, G. Nigel i J. E. Doran (red.), *Social Science Micro Simulation* (s. 416–431). Berlin: Springer.
- Manovich, L. (2012). *Język nowych mediów*. Warszawa: Oficyna Wydawnicza Łódźgraf.
- Marcus, G. E. (1995). Ethnography in / of the World System?: The Emergence of Multi-Sited Ethnography. *Annual Review of Anthropology*, 24, 95–117.
- Mazurek, G. (2018). Transformacja cyfrowa — implikacje dla marketingu. W: B. Gregor i D. Kaczorowska-Spychalska (red.), *Marketing w erze technologii cyfrowych. Nowoczesne koncepcje i wyzwania* (s. 33–57). Warszawa: Wydawnictwo Naukowe PWN.
- Merity, S. (2017). *Bias is not just in our datasets, it's in our conferences and community*. Pozyskano z https://smerity.com/articles/2017/bias_not_just_in_datasets.html
- Metcalf, J. i Crawford, K. (2016). Where are human subjects in Big Data research? The emerging ethics divide. *Big Data & Society*, 3(1). <https://doi.org/10.1177/2053951716650211>
- Micheaux, A. i Bosio, B. (2018). Customer Journey Mapping as a New Way to Teach Data-Driven Marketing as a Service. *Journal of Marketing Education*, 41(2). <https://doi.org/10.1177/0273475318812551>
- Minelli, M., Dhiraj, A. i Chambers, M. (2013). *Big Data, Big Analytics. Emerging business intelligence and analytic trends for today's businesses*. New Jersey: John Wiley i Sons.
- Mojżesz-Zimocznyk, J. (2017). *E-mail. Ewolucja. Data Driven Mailing*. Pozyskano z <https://nowymarketing.pl/a/15811,e-mail-ewolucja-data-driven-mailing>
- Moneta, M. (2018). Jan Zajac (Sotrender): *O matematyce w marketingu*. Pozyskano z <https://nowymarketing.pl/a/17989,jan-zajac-sotrender-o-matematyce-w-marketingu>
- Mulvenna, M., Norwood, M. i Büchner, A. (1998). Data-Driven Marketing. *Electronic Markets*, 8(3), 32–35. <https://doi.org/10.1080/10196789800000038>
- Nadler, A. i McGuigan, L. (2018). An impulse to exploit: The behavioral turn in data-driven marketing. *Critical Studies in Media Communication*, 35(2), 151–165. <https://doi.org/10.1080/15295036.2017.1387279>
- Nicolaus, H., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiesman, B. i Sethupathy, G. (2016). *The age of analytics: Competing in a data-driven world*. McKinsey Global Institute. <https://doi.org/10.1111/bjjet.12230>
- O'Neil, C. (2017). *Broń matematycznej zagłady. Jak algorytmy zwiększają nierówności i zagrażają demokracji*. Warszawa: Wydawnictwo Naukowe PWN.
- Ogilvy. (2018). *The Dress For Respect*. Pozyskano z <https://www.ogilvy.com/work/the-dress-for-respect/>
- Owczarek, K. (2018). *Why Python is growing so quickly. Future trends*. Pozyskano z <https://www.netguru.co/blog/why-python-is-growing-so-quickly-future-trends>
- Parvinen, P., Kaptein, M. i Pöyry, E. (2012). *Data-driven vs. Theory-driven Approaches to e-Selling. 2nd International Conference on Management and Artificial Intelligence*. Pozyskano z <http://ipedr.com/vol35/021-ICMAI2012-E10033.pdf>
- Peltarion. (2018). *The Essential AI Handbook for Leaders*. Pozyskano z <https://peltarion.com/article/ai-handbook>
- Pentland, A. (2009). Reality Mining of Mobile Communications: Toward a New Deal on Data. W: S. Dousta i I. Mia (red.), *The Global Information Technology Report, World Economic Forum 2008–2009. Mobility in a Networked World*. (s. 75–80). INSEAD. Pozyskano z http://hd.media.mit.edu/wef_globalit.pdf
- Pentland, A. i Heibeck, T. (2008). *Honest Signals. How They Shape Our World*. Cambridge: The MIT Press.
- Piatetsky, G. (2017). *Python overtakes R, becomes the leader in Data Science, Machine Learning platforms*. Pozyskano z <https://www.kdnuggets.com/2017/08/python-overtakes-r-leader-analytics-data-science.html>
- Pink, S., Ruckenstein, M., Willim, R. i Duque, M. (2018). Broken data: Conceptualising data in an emerging world. *Big Data & Society*, 5(1). <https://doi.org/10.1177/2053951717753228>
- Provost, F. i Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, 1(1), 51–59. <https://doi.org/10.1089/big.2013.1508>
- PTBRiO. (2017). *Data driven decisions. Przewodnik po źródłach wiedzy w marketingu*. Polskie Towarzystwo Badaczy Rynku i Opinii. Pozyskano z http://datadrivendecisions.pl/ddd_guide.pdf
- PTBRiO. (2019). *Data Driven Decisions 2.0*. Polskie Towarzystwo Badaczy Rynku i Opinii. Pozyskano z <http://www.datadrivendecisions.pl/>
- Richardson, R., Schultz, J. i Crawford, K. (2019). Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. *New York University Law Review Online*, (February), 192–233. Pozyskano z https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333423

- Savage, M. i Burrows, R. (2007). The Coming Crisis of Empirical Sociology. *Sociology*, 41(5), 885–899. <https://doi.org/10.1177/0038038507080443>
- Schmidt, E., Rosenberg, J. i Eagle, A. (2014). *Jak działa Google*. Kraków: Insignis Media.
- Shah, D. V., Cappella Ramesh, J. N. i Neuman, W. R. (2015). Big Data, Digital Media, and Computational Social Science: Possibilities and Perils. *Annals of the American Academy of Political and Social Science*, 659(1). <https://doi.org/10.1177/0002716215572084>
- Silver, N. (2014). *Sygnal i szum*. Gliwice: Helion.
- Strauss, A. L. (1978). A Social World Perspective. *Studies in Symbolic Interaction*, (1), 119–128.
- Strong, A. (2018). *Why I'm Learning Python in 2018*. Pozyskano z: <https://news.codecademy.com/why-learn-python/>
- Surma, J. (2017). *Cyfryzacja życia w erze big data. Człowiek, biznes, państwo*. Warszawa: Wydawnictwo Naukowe PWN.
- Symeonidis, I., Tsormpatzoudi, P. i Preneel, B. (2015). Collateral damage of Facebook Apps: An enhanced privacy scoring model. *IACR Cryptology ePrint Archive*, 1–18. Pozyskano z <http://dblp.uni-trier.de/db/journals/iacr/iacr2015.html#SymeonidisBTP15>
- Szymielewicz, K. i Iwańska, K. (2019). *Śledzenie i profilowanie w sieci. Jak z klienta stajesz się towarem*. Pozyskano z https://panoptykon.org/sites/default/files/publikacje/panoptykon_raport_o_sledzeniu_final.pdf
- Teddlie, C. i Tashakkori, A. (2009). *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Los Angeles: Sage.
- Törnberg, P. i Törnberg, A. (2018). The limits of computation: A philosophical critique of contemporary Big Data research. *Big Data & Society*, 5(2), 1–12. <https://doi.org/10.1177/2053951718811843>
- USACM. (2017). *Statement on algorithmic transparency and accountability, USACM press releases §*. Association for Computing Machinery US Public Policy Council. Pozyskano z https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf
- Vedder, A. (1999). KDD: The challenge to individualism. *Ethics and Information Technology*, (1), 275–281. <https://doi.org/10.1023/A:1010016102284>
- Wang, S. i Qu, H. (2008). Journal of Travel i Tourism Impacts of Accession to WTO on China's Travel Industry. *Journal of Travel i Tourism Marketing*, 12(1), 63–80. <https://doi.org/10.1300/J073v12n01>
- Wang, T. (2013). *Why Big Data Needs Thick Data*. Pozyskano z <https://medium.com/ethnography-matters/why-big-data-needs-thick-data-b4b3e75e3d7>
- Wang, T. (2016). *The human insights missing from big data*. Pozyskano z https://www.ted.com/talks/tricia_wang_the_human_insights_missing_from_big_data#t-9755
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., West, S. M., Richardson, R., Schulz, J. i Schwartz, O. (2018). *AI Now Report 2018*. New York. Pozyskano z https://ainowinstitute.org/AI_Now_2018_Report.pdf
- Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, 59(10), 1–23. <https://doi.org/10.18637/jss.v059.i10>
- Wickham, H. i Grolemund, G. (2017). *R for Data Science*. Pozyskano z <http://r4ds.had.co.nz/>
- Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89. <https://doi.org/10.1057/jit.2015.5>
- Żulicki, R. (2017). Potencjał Big Data w badaniach społecznych. *Studia Socjologiczne*, 3(226), 176–207. Pozyskano z http://www.studiasocjologiczne.pl/pliki/Studia_Socjologiczne_2017_nr3_str175_207.pdf

PWE poleca

Jacek Kamiński

MAKROMARKETING

Nauka o marketingu
wobec problemów społecznych
i środowiskowych



Polskie Wydawnictwo Ekonomiczne

Makromarketing jest obszarem nauki o marketingu, który bardziej niż marketingiem jako funkcją zarządzania i narzędziem menedżera jest zainteresowany społeczną i środowiskową rolą marketingu. Traktuje on marketing jako proces społeczny, główny akcent zainteresowania kładąc na oddziaływanie marketingu na społeczeństwo oraz społeczeństwa na marketing.

Książka jest jednym z nielicznych i pierwszym na polskim rynku wydawniczym opracowaniem z zakresu teorii marketingu oraz historii myśli marketingowej poświęconym makromarketingowi. Jej głównym celem jest przedstawienie istoty, historii, przedmiotu i zakresu makromarketingu jako części nauki o marketingu oraz ukazanie roli, jaką spełnia on w jej rozwoju, a także szans, jakie w związku z rozwojem problematyki makromarketingowej stają przed nauką o marketingu.

www.pwe.com.pl