

Olga Rodak

Akademia Leona Koźmińskiego w Warszawie

## **TWITTER JAKO PRZEDMIOT BADAŃ SOCJOLOGICZNYCH I ŹRÓDŁO DANYCH SPOŁECZNYCH: PERSPEKTYWA KONSTRUKTYWISTYCZNA\***

Serwis społecznościowy Twitter zyskał szczególne zainteresowanie naukowców społecznych z co najmniej dwóch powodów. Po pierwsze, stanowi on fenomen społeczny, wpływając na praktyki komunikacyjne oraz sposoby organizowania się ludzi. Po drugie, Twitter to repozytorium ustrukturyzowanych danych, które mogą zostać wykorzystane w celu badania zjawisk związanych lub niezwiązanych bezpośrednio ze społecznym oddziaływaniem tego medium. Celem niniejszego artykułu jest zebranie wiedzy, która, w opinii autorki, jest niezbędna do sprawnego i owocnego wykorzystania Twittera do prowadzenia badań socjologicznych; zarówno tych, w których Twitter jest przedmiotem analizy, jak i tych, w których Twitter dostarcza jedynie danych do analizy. Jako że technologia rekonfiguruje praktyki społeczne, będąc jednocześnie przez praktyki społeczne kształtowaną, socjologia Twittera wymaga przede wszystkim zrozumienia charakteru tej technologii oraz wpływu, jaki ona wywiera na praktyki badawcze, będące praktykami społecznymi. W artykule pokazany zostaje przekształcający wpływ Twittera w tych dwóch aspektach. Perspektywa konstruktywistyczna pozwala na zrozumienie zarówno samego medium, jak i charakteru oraz ograniczeń wytwarzanych za jego pośrednictwem danych.

Główne pojęcia: Twitter; socjologia Internetu; socjologia cyfrowa; badania społeczne w Internecie, serwisy społecznościowe; dane z mediów społecznościowych; Big Data, *small data*.

---

Katedra Zarządzania w Społeczeństwie Sieciowym, e-mail: orodak@kozminski.edu.pl

\*Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2012/05/E/HS4/01498.

Serdecznie dziękuję dr. Andrew Asherowi z Indiana University za wprowadzenie mnie w techniczne i metodologiczne tajniki badań z wykorzystaniem danych z Twittera, kolegom i koleżankom z Katedry Zarządzania w Społeczeństwie Sieciowym oraz uczestnikom i uczestniczkom seminarium *Organizing Uncertain Futures* w Akademii Leona Koźmińskiego za konsultowanie tekstu na różnych etapach jego powstawania, a także Marcinowi Olendrowi i Michałowi Wieluńskiemu z Działu Obsługi Klienta i Wdrażania Innowacyjnych Rozwiązań ALK za techniczne umożliwienie realizacji badań z użyciem danych z Twittera. Szczególnie dziękuję anonimowym recenzentom za trafne uwagi do pierwszej wersji niniejszego artykułu.

Serwis społecznościowy Twitter, pomimo nieporównywalnej w stosunku do Facebooka bazy użytkowników, zyskał szczególne zainteresowanie naukowców społecznych z co najmniej dwóch powodów. Po pierwsze, serwis ten stanowi fenomen społeczny, trwale i znacząco wpływając na praktyki komunikacyjne oraz sposoby organizowania się ludzi. Po drugie, Twitter to repozytorium ustrukturyzowanych danych – stu czterdziestoznakowych wiadomości generowanych przez użytkowników wraz z ich metadanymi – które mogą zostać wykorzystane w celu badania zjawisk związanych lub niezwiązanych bezpośrednio ze społecznym oddziaływaniem Twittera. Cechy te czynią z Twittera istotny obiekt zainteresowania socjologii i nauk pokrewnych.

Celem niniejszego artykułu jest próba obniżenia progu dostępu do badań z wykorzystaniem danych z Twittera dla polskich socjologów i badaczy społecznych poprzez dzielenie się własnym doświadczeniem w korzystaniu z danych z Twittera. W tym sensie artykuł ten jest przejawem polityki metody, to znaczy aktywnego działania na rzecz włączenia badań typu Big Data do stałego repertuaru metodologii socjologicznej oraz zabierania przez socjologię głosu w debacie na temat dostępu i wykorzystania komputerowo zapośredniczonych danych (Savage i Burrows 2007). Nie ograniczam się w nim jednak do omówienia technicznych i finansowych aspektów dostępu do danych, ponieważ zarówno badania społecznych zjawisk związanych z Twitterem, jak i korzystanie z jego danych w celach badawczych wymagają dogłębnego zrozumienia charakteru tego medium. Przyjęłam perspektywę technologicznego konstruktywizmu, by w kolejnych częściach tego artykułu pokazać, jak społecznie konstruowany jest kształt techniczny Twittera, ale również w jaki sposób technologia ta przekształca praktyki badawcze, które są przecież praktykami społecznymi. Socjologdy i inni badacze społeczni zyskują w ten sposób zrozumienie zarówno samego medium, jak i charakteru oraz ograniczeń wytwarzanych za jego pośrednictwem danych.

Serwis, założony w 2006 roku w San Francisco przez Jacka Dorsey'a, w chwili pisania tego artykułu ma około 313 milionów aktywnych użytkowników w miesiącu (Twitter 2017a), a w 2013 roku przekroczył próg miliarda zarejestrowanych użytkowników (Długosz 2014). Dla porównania, w przypadku Facebooka jest to odpowiednio 1,86 miliarda aktywnych użytkowników w miesiącu oraz ponad dwa miliardy zarejestrowanych użytkowników (Facebook 2017). Cechą charakterystyczną serwisu są ograniczone do stu czterdziestu znaków publiczne wiadomości, tzw. *tweety* – z tego powodu jest on określany mianem serwisu mikroblogowego. Wiadomości jeszcze do niedawna publikowane były w odwrotnej chronologicznej kolejności, a aktualizacja strumienia wymagała odświeżenia strony<sup>1</sup>. Medium to ma efemeryczny charakter – tweety znikają z serwisu po

<sup>1</sup> W 2016 roku zwiększono rolę algorytmu podsuwającego użytkownikom potencjalnie interesujące dla nich treści (Jenkins 2017)(L. D. Jenkins 2017).

kilku dniach. Wybrani użytkownicy mogą być obserwowani, jednak nie powoduje to automatycznej wzajemności, jak w przypadku Facebooka. Użytkownicy Twittera wchodzą w interakcje za pomocą specyficznych funkcjonalności technicznych serwisu, takich jak hashtag (#, ang. *hash*), adresowanie wiadomości (@, ang. *at*) czy podawanie dalej wiadomości (ang. *retweet*) (Rogers 2014).

Choć prawdopodobnie przesadą jest stwierdzenie, że Twitter zmienił świat (Chaykovsky 2017b), to z pewnością medium to istotnie wpłynęło na sposób komunikacji i organizacji społecznej. Stało się tak pomimo jego efemerycznego charakteru, interfejsu niesprzyjającego budowie relacji społecznych (Honeycutt i Herring 2009) oraz ograniczonej liczby stałych użytkowników. Dzięki publicznemu charakterowi wiadomości i możliwości obserwowania wybranych kont, Twitter stał się przestrzenią błyskawicznego rozprzestrzeniania się informacji i opinii, dostarczającą pożywki tradycyjnym mediom i kulturze popularnej. Z kolei hashtag umożliwił podpinanie publicznych wiadomości pod szersze kanały konwersacyjne. W ten sposób Twitter stał się punktem węzłowym interakcji i samoorganizacji ruchów społecznych (Segeberger i Bennett 2011; Lindgren i Lundström 2011), wspólnot profesjonalistów (Lewis i Rush 2013; Stephansen i Couldry 2014) czy oddolnych inicjatyw mających na celu zarządzanie kryzysowe (Aljohani i in. 2011; Marjak 2014). Między innymi z tych powodów Twitter traktowany jest jako fenomen społeczny, a zatem potencjalny obiekt zainteresowania socjologii. Badania dotyczące społecznego wpływu Twittera opierają się na danych zastanych (tweets) i danych wywołanych (obserwacje, wywiady, ankiety), ograniczając stanowisko badawcze do serwisu, jak i w razie potrzeby wychodząc poza jego granice.

Dzięki publicznej dostępności większości tweetów, ich przejrzystej strukturze oraz przyrostowi w tempie 500 milionów wiadomości miesięcznie (Internet Live Stats 2017), Twitter jest też wykorzystywany jako rezerwuuar pokaźnych wolumenów danych. Socjolog ma tu jednak do czynienia ze szczególnym rodzajem danych, a mianowicie Big Data. Są to intencjonalnie bądź nieintencjonalnie pozostawione w sieci ślady cyfrowo zapośredniczonych praktyk, w tym przypadku treść tweetów wraz z metadanymi, takimi jak geolokalizacja czy informacja o liczbie polubień. Big Data ma dawać „obraz interakcji, chwila po chwili, przez dłuższy czas, dostarczający informacje zarówno o strukturze, jak i zawartości relacji” (Lazer i in. 2009). Dzięki bezprecedensowym rozmiarom i gęstości dane te umożliwiają wykrywanie wzorców niewidocznych gołym okiem oraz niemożliwych do uchwycenia za pomocą konwencjonalnych metod socjologicznych, w których dane zbiera i interpretuje badacz. Lev Manovich porównywał Big Data do zdjęcia cyfrowego, które nie traci dokładności przy powiększeniu, ale za każdym razem przynosi nowe informacje (Manovich 2011; Krzysztofek 2012a).

Socjologia na razie podejrzliwie traktuje trend Big Data, głównie ze względu na trudności w uzyskaniu wiedzy na temat reprezentatywności danych.

Możliwość uzyskania reprezentatywnej albo przynajmniej przejrzystej próby kont lub tweetów jest ograniczona (Liang i Fu 2015). Możliwe jest pobranie pewnej liczby tweetów wyodrębnionych na podstawie określonych filtrów oraz jednego procenta losowo wybranych tweetów. Ponadto, większość treści w serwisie wytwarza niewielka liczba użytkowników, a jego profil demograficzny nie odzwierciedla zróżnicowania społeczeństwa (Lindgren i Lundström 2011). W związku z tymi ograniczeniami, badania typu Big Data wymuszają przesunięcie nacisku z reprezentatywności na interpretatywność; z pytania „jak potrzebne dane pozyskać?” na „co z dostępnymi danymi zrobić?”. Otwiera to pole do popisu dla socjologów i innych badaczy społecznych, dysponujących umiejętnością kontekstualnego rozumienia danych oraz zastosowania teorii społecznej do ich interpretacji. Ażeby więc można było właściwie rozeznąć się w ograniczeniach dostępnych danych, potrzebne jest dokładne poznanie technologii, która umożliwi ich generowanie, oraz modelu biznesowego jej dostarczycieli, którzy kontrolują do nich dostęp (Brooker i in. 2016).

Wyjątkowa atrakcyjność Twittera jako źródła danych z mediów społecznościowych – pomimo jego niszowości – wynika z jednej strony z faktu, że dane są zestandaryzowane i przejrzyste, z drugiej zaś z istnienia zintegrowanej społeczności badaczy, którzy dostarczają skodyfikowanej wiedzy oraz narzędzi do pobierania i analizy tych danych. Dane te mogą zostać wykorzystane w celu badania problemów związanych bądź niezwiązanych bezpośrednio z wpływem serwisu na rzeczywistość społeczną. Przykładem pierwszego rodzaju badania jest studium sposobów budowania relacji organizacji z jej interesariuszami na Twitterze (Brindusa Albu i Etter 2016), a drugiego – analiza treści tweetów i relacji pomiędzy nimi w celu zbadania sposobów konceptualizowania własnej pracy przez użytkowników Twittera (Sergi i Bonneau 2016). Ze względu na dynamiczny przyrost wiadomości oraz chaotyczność ich ukazywania się w serwisie bywa, że rekomenduje się automatyczne pobieranie wiadomości oraz stosowanie w odniesieniu do nich analiz ilościowych. To podejście umożliwia choćby zaobserwowanie dynamiki interakcji albo zidentyfikowanie najbardziej aktywnych i wpływowych użytkowników (Stephansen i Couldry 2014; Marwick 2014). Z tego powodu niniejszy artykuł w dużej mierze redukuje temat do wytwarzanych w serwisie danych, choć oczywiście do badania interakcji w serwisie oraz jego społecznego wpływu używa się też bardziej konwencjonalnych podejść, w tym badawczych metod jakościowych (zob. Simonetto 2016; Loureiro-Koechlin i Butcher 2013).

Twitter stwarza więc wiele możliwości dla socjologów i innych badaczy społecznych, ale postulat automatycznego pozyskiwania i analizy danych sprawia, że na drodze do ich realizacji piętrzą się liczne przeszkody. Podstawową barierę stanowi konieczność pogłębienia umiejętności technicznych bądź nawiązania współpracy z odpowiednimi specjalistami, a niekiedy również zdobycia środków finansowych w celu uzyskania dostępu do danych czy narzędzi

analizy. Równie istotna jest jednak bariera mentalna: konieczność zrozumienia jakiego rodzaju wiedza może zostać wytworzona dzięki wykorzystaniu danych z Twittera w kontekście tradycyjnych sposobów wytwarzania wiedzy w socjologii, takich jak badania ankietowe, badania terenowe czy analiza dokumentów. Jaka będzie „teoria *dobrej* roboty naukowej” w socjologii Big Data (Sułek 2002: 17).

Polscy badacze korzystają oczywiście z danych z Twittera, uzyskując ciekawe rezultaty. Prowadzą badania dotyczące wpływu serwisu na różne aspekty życia społecznego: relacje pomiędzy organizacjami i ich interesariuszami (Łapczyński 2016; Tereskiewicz 2017), aktywność polityków (Kubczak 2014) oraz dziennikarzy (Zagórska 2014), sposób komunikacji na temat literatury (Każmierczak 2012), wirusowe rozprzestrzenianie się opinii oraz ramowanie dyskursu politycznego (Annusewicz 2016). Wykorzystują również dane w celu badania tradycyjnych dla danej dziedziny problemów, np. społecznego odbioru przestrzeni (Rzeszewski 2015) oraz nastrojów na temat marki (Michalak 2016; Wolny 2016). Wśród badaczy dominują językoznawcy, literaturoznawcy, medioznawcy, politolodzy i geografowie. Twitterowe dane wykorzystują również przedstawiciele nauk ścisłych, przede wszystkim w celu badania rozpowszechniania się informacji i opinii (Dzida i Markowski 2013)<sup>2</sup>. Na ich tle wyróżnia się brak publikacji w czasopismach socjologicznych czy antropologicznych. Ponadto nie napisano do tej pory tekstu zbierającego i porządkującego wiedzę, której zdobycie umożliwiłoby płynne wejście w świat badań Twittera badaczom „nie-technicznym”.

Stosunkowo umiarkowane zainteresowanie Twitterem wśród polskich badaczy wynika prawdopodobnie z faktu, że serwis ten nie przyjął się szeroko w naszym społeczeństwie. W Polsce Twittera odwiedza ponad 3,5 miliona użytkowników, ale zarejestrowanych jest tylko kilkaset tysięcy (Skrzypek 2016; zob. Wachnicki 2013). Informacje o aktywności niektórych użytkowników na Twitterze pojawiają się w mediach tradycyjnych głównie ze względu na obecność dziennikarzy w serwisie (Machnik 2014). Istnieje jednak kilka powodów, dla których serwis ten ma istotne znaczenie dla socjologii. Po pierwsze, choć Twitter jest w Polsce medium niszowym, to ma oddane grupy użytkowników, które można by uznać za „wspólnoty praktyki” (Lave i Wenger 1991; zob. Wachnicki 2013). Stwarza to możliwość badania strategii kreowania własnej marki albo norm kulturowych panujących w różnych kręgach profesjonalnych. Ponadto wokół serwisu narosło sporo stereotypów. Uważa się na przykład, że typowy użytkownik Twittera to wykształcony mężczyzna z dużego miasta między 25. a 34. rokiem życia: dziennikarz, polityk, przedstawiciel wolnego zawodu albo

<sup>2</sup> Zob. też: projekt RENOIR kierowany przez prof. dr. hab. inż. Janusza Hołysta z Wydziału Fizyki Politechniki Warszawskiej (Politechnika Warszawska 2015).

student (Skrzypek 2016). Tymczasem w 2015 roku ponad 1,3 mln spośród użytkowników było mieszkańcami wsi (Wirtualne Media 2016). Po drugie, liczba użytkowników Twittera w Polsce sukcesywnie wzrasta, zwiększa się również ilość publikowanych przez nich treści i zaangażowanie w nie (Skrzypek 2016; Poprawa 2017). Po trzecie, pomimo mniejszej popularności niż Facebook, Twitter jest znacznie poręczniejszym i lepiej opisanym przez badaczy społecznych zbiorem dużych danych. Po czwarte, badacz Twittera nie musi się ograniczać do polskojęzycznych danych. Wreszcie po piąte, dzięki teoretycznej ramie tego artykułu wnioski mogą być ekstrapolowane na inne platformy internetowe.

### **Konstruktywizm technologiczny a socjologiczne rozumienie Twittera**

Socjologia i pokrewne nauki mają długą tradycję badania relacji pomiędzy wynalazkami technicznymi i społeczeństwem w duchu konstruktywizmu, podważającego determinizm, czyli założenie, że technologia wywiera nieunikniony i z góry zaprogramowany wpływ na sposób organizacji społeczeństwa. Studia nad nauką i technologią opisują, w jaki sposób technologia jest współkonstruowana przez jej użytkowników (Bińczyk 2010). Jednymi z głównych teorii tego nurtu są teoria społecznego konstruowania technologii (SCOT – *Social Construction of Technology*) oraz postkonstruktywistyczna teoria aktora-sieci (ANT – *Actor-Network Theory*). Według założeń tej pierwszej, sukces danej technologii oraz jej ostateczny kształt zależą od społecznych warunków jej adaptacji, to znaczy tego, kto i kiedy danej technologii używa. SCOT obejmuje nie tylko teoretyczne założenia, ale również metodologiczne wskazówki badania społecznych czynników adaptacji danych technologii (Bijker, Hughes i Pinch 1987). ANT dąży do przekroczenia podziału na czynniki społeczne, naturalne i technologiczne, opisując w jaki sposób wspólnie wytwarzają „to, co społeczne” (Latour 2010).

Spółeczeństwo kształtuje technologię, ale istotne znaczenie ma również ruch w przeciwną stronę, to znaczy sposób, w jaki to technologia współkonstruuje społeczeństwo poprzez kształtowanie praktyk oraz tworzenie nowego rodzaju zapośredniczonych technicznie zbiorowości. Na przykład nowe media nie dlatego są nowe, że opierają się na innowacyjnych technologiach, ale dlatego, że urządzenia te modyfikują „formy aktywności i praktyki komunikacyjne”, które z kolei wpływają na „porządek społeczny i społeczne formy organizacyjne” (Pokorna-Ignatowicz i in. 2013). Już w latach pięćdziesiątych XX wieku socjologodzy przemysłu, badający skutki wprowadzenia linii produkcyjnej w brytyjskich kopalniach węgla, ukuli termin „system socjotechniczny” (Trist i Bamforth 1951). Termin ów określa szczególnie układ organizacyjny, będący efektem wprowadzenia nowej technologii w określony kontekst społeczny. Na gruncie polskich



nauk społecznych podobne wnioski wysnuł w 1956 roku Stefan Czarnowski, gdy opisywał, jak zastąpienie drewnianego radła pługiem zmieniło organizację pracy, a pośrednio strukturę społeczną wsi. Użycie radła wymagało wysiłku wielu osób nadzorowanych przez gospodarza, natomiast pług mógł być stosunkowo łatwo używany przez każdego członka gospodarstwa, co zrównało kobiety, parobków i młodzież względem gospodarza, oraz wytworzyło dużą ilość wolnego czasu (Czarnowski 1956).

W późniejszych teoriach starano się tę relację opisać nieco dokładniej. Na przykład Wanda Orlikowski prowadziła badania nad sprawczą rolą nowych technologii informacyjno-komunikacyjnych, oraz wpływem kontekstu, w jakim są one stosowane, na ostateczny efekt społeczny, jaki wywierają. Zainspirowana pracami Anthony'ego Giddensa na temat relacji pomiędzy strukturą społeczną i sprawczością jednostek, Orlikowski zaproponowała teorię technologii jako zasobu, który w określonych warunkach jednocześnie ogranicza użytkowników, jak i umożliwia im działanie. W świetle tej teorii organizacja społeczna jest zatem rezultatem zarówno tego, w jaki sposób została zaprojektowana technologia, jak i tego, jak jest ona używana w określonym kontekście, np. w miejscu pracy (Orlikowski 1992, 2007; Orlikowski i Robey 1991).

W tym tekście nie opowiedziałam się konkretnie za żadną z teorii konstruktywistycznych; powyższy przegląd ma zresztą charakter poglądowy i nie wyczerpuje tematu. Celem jest raczej osadzenie opowieści o wybranej technologii w konstruktywistycznej ramie, w czym naśladuję zresztą innych badaczy Twittera (Bruns i Burgess 2016; Halavais 2014; Rogers 2014; Weller i in. 2014). Odrzucenie determinizmu implikuje, że samo zrozumienie sposobu, w jaki zaprojektowana jest technologia, a także stojącego za nią modelu biznesowego, nie wyczerpuje wiedzy ani o tej technologii, ani o społecznych efektach jej użytkowania. W architekturę technologii wbudowana jest pewna sugestia na temat sposobu jej używania (McGrenere i Ho 2000), natomiast to, co dzieje się z technologią i z jej użytkownikami później, pozostaje otwartym tematem badawczym dla socjologów, antropologów czy badaczy organizacji. Dotychczasowe badania Twittera, jak również pojawiające się stale doniesienia prasowe, wydobywają społeczne konstruowanie tego medium oraz jego znaczny wpływ na praktyki społeczne, w tym praktyki badawcze. W kolejnych częściach artykułu omówiony zostanie przypadek Twittera w tych dwóch aspektach, z podkreśleniem tego, jak konstruktywistyczne ujęcie pozwala lepiej zrozumieć Twittera jako obiekt badań oraz źródło danych.

## Serwis społecznościowy Twitter jako rezultat wzajemnego oddziaływania interesariuszy platformy

Przypadek Twittera egzemplifikuje „społeczne współkonstruowanie technologii nowych mediów w często konfliktowym wzajemnym oddziaływaniu użytkowników platformy, jej dostawców i innych interesariuszy” (Weller i in. 2014: xxxi). By zrozumieć dynamikę tego oddziaływania, należy przede wszystkim uzmysłowić sobie, że Twitter Inc. jest firmą nastawioną na zysk, która funkcjonuje w specyficznym kontekście ekonomicznym. Źródłem konfliktu jest napięcie pomiędzy tym, że Twitter pełni funkcję dobra publicznego, bazuje na darmowej pracy użytkowników, a równocześnie dąży do wytworzenia zysku i zbudowania stabilnego modelu biznesowego. W tej części artykułu opisuję kontekst ekonomiczny funkcjonowania Twittera, a także ewolucję tej platformy jako wyniku oddziaływania jej interesariuszy. Ze względu na cele tego artykułu, szczególną uwagę poświęciłam dwóm kategoriom interesariuszy: użytkownikom końcowym platformy oraz programistom korzystającym z danych podanych przez API (*Application Programming Interface*), czyli interfejsu, który daje dostęp do ustrukturyzowanych, choć surowych danych wytwarzanych na platformie.

Generację serwisów internetowych znaną jako Web 2.0. oparto na technologiach umożliwiających współtworzenie treści przez użytkowników, co podważyło tradycyjny dla mediów model biznesowy, w którym jasno oddziela się producentów i konsumentów dóbr (van Dijck 2016, 2011). W branży internetowej wykształcił się określony model rozwojowy firm dostarczających te serwisy, który polega na przyjmowaniu formy *start-upu* w początkowej fazie istnienia oraz rozwijaniu projektu dzięki pozyskanemu kapitałowi *venture*. Odrzucając z góry określony model biznesowy, twórcy obserwują sposób adaptacji technologii przez użytkowników zakładając, że to umożliwi późniejsze określenie strumieni przychodów. W tym okresie firma buduje relację z użytkownikami adaptującymi technologię. Użytkownicy ci często woluntarystycznie rozwijają technologię i nadają jej społeczne znaczenie poprzez sposób, w jaki z niej korzystają. Celem tej fazy jest również zgromadzenie dużej, stabilnej sieci użytkowników, która umożliwi pokonanie konkurencji i będzie podstawowym źródłem wartości. W tych modelach bowiem to zwykle aktywność użytkowników ma centralne znaczenia dla wytwarzania zysku (Fuchs i Sevignani 2013). W pewnym momencie jednak zwiększa się presja na zwrot z inwestycji, a co za tym idzie na ustabilizowanie modelu biznesowego. Wybrane formy generowania zysku nie powinny zrażać użytkowników, chociaż kiedy osiągnięty zostaje efekt sieci, użytkownicy stają się zależni od danej platformy i mogą pozostać jej klientami ze względu na brak alternatywy. Firma może ogłosić ofertę



publiczną na giełdzie, by pozyskać kolejne środki, jednak wówczas pojawia się presja na szybki zwrot z inwestycji ze strony kolejnych akcjonariuszy, wzmacniana przez kwartalne publikacje wyników.

Twitter Inc. wpisuje się w ten model rozwojowy, a wręcz jest jego podręcznikowym przykładem. W pierwszych latach istnienia firma budowała swoją pozycję jako niezależna platforma mikroblogowa, jednocześnie obserwując użytkowników i czerpiąc zysk z przynoszonych przez nich innowacji. Mniej więcej od 2009 roku uwidocznił się lepiej sprecyzowany kierunek rozwoju firmy. Stało się jasne, że platforma ma być raczej globalnym medium transmisji i cyrkulacji informacji, niż intymną przestrzenią współpracy społeczności internetowych (van Dijck 2011). Model biznesowy Twittera zaczął się stabilizować, a obecnie oparty jest na spieniężaniu danych z treści generowanych przez użytkowników oraz ze śladów ich ruchu na platformie (Puschmann i Burgess 2014; van Dijck 2011). W 2016 roku Twitter wykazał dwa źródła przychodu (Twitter 2017b). Prawie 90% pochodzi z reklam. Twitter oferuje w tym obszarze trzy podstawowe rodzaje produktów promowanych: tweety, konta lub tematy. Reklamodawca może wykupić pakiet, dzięki któremu algorytm umieści odpowiednie treści bezpośrednio w strumieniu wiadomości potencjalnie zainteresowanych użytkowników, a Twitter otrzyma prowizję od każdego kliknięcia (Gadkari 2013). Drugie, znacznie skromniejsze źródło dochodu, stanowi licencjonowanie danych (nieco ponad 10%), a konkretnie sprzedawanie surowych danych za pośrednictwem firmy Gnip, którą zresztą Twitter zakupił w roku 2014 (Kepes 2014). Podstawowe znaczenie dla serwisu ma zatem przyrost liczby użytkowników oraz ich zaangażowanie w tworzenie treści (Puschmann i Burgess 2014).

Badacze Twittera często opisują ewolucję tej platformy jako efekt napięcia na linii użytkownik końcowy – dostawca platformy. Twitter powstał jako usługa autonomiczna wobec innych serwisów i otwarty kanał do komunikacji tekstowej na różnych urządzeniach (van Dijck 2011). Było to proste narzędzie do ekspresowej komunikacji pomiędzy wielkomięskimi grupami przyjaciół; rodzaj internetowych sms-ów czy też „błyskawiczny komunikator publiczny”. Bezpośrednią inspirację dla platformy miały stanowić systemy komunikacyjne dla kurierów rowerowych, pogotowie ratunkowe, GPS oraz CB radio (Rogers 2014). Twórcy serwisu przyznają, że sami do końca nie wiedzieli, czym jest Twitter i w jakim kierunku należy go rozwijać (Williams 2013).

Znaczny wpływ na rozwój platformy mieli jej pierwsi użytkownicy, którzy ciężyli ku interakcji i społecznościowości pomimo niesprzyjającej architektury serwisu (Rogers 2014). W tamtym czasie platforma pozwalała głównie na transmitowanie wiadomości do wybranego grona obserwatorów (van Dijck 2011). Twittera zaludniali użytkownicy, którzy adaptowali rozwiązania zaczerpnięte z kultury wczesnych środowisk internetowych, takich jak powstały w 1988 roku

*Internet Relay Chat* (IRC): znaki @, # oraz RT (*retweet*). Firma wbudowywała te tekstowe rozwiązania w architekturę platformy, jeśli pozostawały one w zgodzie z jej kształtującą się wówczas wizją rozwoju serwisu (Rogers 2014; Puschmann i Burgess 2014). Wbrew pozorom, nie powodowało to zawsze entuzjazmu użytkowników, ponieważ oznaczało uniformizację praktyk społecznych na platformie (Puschmann i Burgess 2014). W tamtym czasie pojawiły się również programy służące do skracania linków, co umożliwiło umieszczanie ich w stuczterdziestoznakowych tweetach i społecznościowe komentowanie informacji (Rogers 2014).

Zamiast wspierać te tendencje do współpracy między użytkownikami (Honeycutt i Herring 2009), w 2009 roku dostarczyciele technologii zdecydowali się jednak zmienić projekt platformy poprzez umożliwienie jednostronnego obserwowania kont użytkowników (przycisk „obserwuj”), tym samym wzmacniając wizję Twittera jako medium rozprzestrzeniania informacji (Burgess 2014). Spowodowało to zmianę w profilu demograficznym użytkowników: profesjonalistów po trzydziestce zastąpiły młodsze osoby szukające rozrywki i informacji. Widocznym przejawem zmiany strategii było pytanie umieszczone w polu publikacji: już nie „co słyhać?”, ale „co się dzieje?”, stało się głównym tematem tweetów. Serwis został połączony z innymi narzędziami w środowisku internetowym oraz pokonał konkurencyjne platformy mikroblogowe (van Dijck 2011). W ten sposób z autonomicznej usługi komunikacyjnej Twitter stał się serwisem społecznościowym, czyli medium, którego istotą jest możliwość podtrzymywania, rozbudowywania i wizualizacji własnej sieci kontaktów przez użytkownika za pośrednictwem jego publicznego profilu (boyd i Ellison 2008).

Decyzja o wprowadzeniu możliwości jednostronnego obserwowania kont oraz adaptacja hasztagu przez użytkowników sprawiły, że Twitter awansował do rangi dobra publicznego – przestrzeni błyskawicznego rozprzestrzeniania się informacji i opinii, wyprzedzającej inne media i cieszącej się bazą wiernych użytkowników (Jenkins 2017; zob. Machnik 2014; Szews 2014). Dzięki hasztagowi możliwe stało się „podpinanie” tweetów pod szersze kanały konwersacyjne, wykraczające poza poziom interakcji z siecią znajomych (Bruns i Moe 2014). Za pośrednictwem Twittera, użytkownicy transmitowali wydarzenia, a także wpływali na przebieg konferencji (zob. *South by Southwest* w Austin w 2007 roku), akcji ratowniczych (zob. trzęsienie ziemi w Syczuanie w 2008 roku, atak terrorystyczny w Bombaju w 2008 roku) czy rewolucji społecznych (zob. Euromajdan 2013). Twitter zyskał na znaczeniu zarówno dla zwykłych użytkowników-obywateli, jak i dla służb porządkowych, agencji rządowych i badaczy. Dużą rolę w kształtowaniu się nowego typu użytkownika Twittera, opartego na zarządzaniu prosumpcją dużych wolumenów danych, odegrały urządzenia mobilne (Rogers 2014), z których w 2013 roku logowało się do serwisu już 75% twitrowiczów (Gadkari 2013).

Rozrost bazy użytkowników Twittera, wypracowanie narzędzi i metodologii badawczych, a także ujednoczenie kształtu technicznego platformy, uczyniły z Twittera poręczną bazę danych z wbudowanymi kategoriami analitycznymi (Rogers 2013, 2014; Halavais 2014). Należą do nich między innymi: *retweet* (RT), czyli cytat, „podanie dalej” tweeta przez kolejnego użytkownika (wyróżnienie szczególnie istotnych, prominentnych treści); hasztag (#), którego podstawową funkcją jest tematyczne kategoryzowanie tweetów (folksonomia); relacja wzajemnego lub jednostronnego obserwowania, która umożliwia analizę sieci społecznych; adresowanie wiadomości (*@reply*), które pozwala na wyróżnienie tweetów związanych z tą samą konwersacją (aspekt społecznościowy); czy geolokalizacja (aspekt geograficzny).

Wydaje się, że kolejny etap rozwoju Twittera rozpoczął się w roku 2013 (Luckerson 2013), gdy firma weszła na giełdę i znalazła się pod presją regularnej prezentacji wyników finansowych. Od tamtego czasu widać nerwowe próby nadawania platformie usystematyzowanego kierunku rozwoju, spieniężania takich dostępnych zasobów jak dane oraz wprowadzana zmian w celu przyciągnięcia nowych użytkowników, głównie poprzez adaptację bieżących trendów w technologiach społecznościowych: personalizację czy transmisję wideo na żywo bezpośrednio z platformy (Chaykovsky 2017b; Jenkins 2017). Do tej pory działania te nie przyniosły oczekiwanych skutków. Udało się nieco zwiększyć zaangażowanie użytkowników w wytwarzanie treści, a Twitter notuje stały niewielki wzrost liczby użytkowników, lecz wciąż przynosi straty (Chaykovsky 2017b)<sup>3</sup>. Niektórzy komentatorzy winą za ten stan rzeczy obarczają niedostateczne zaangażowanie kierownictwa, które ma uciekać z firmy niczym z tonącego statku albo dzielić uwagę pomiędzy Twittera a bardziej rozwojowe projekty (Chaykovsky 2017a, 2017b). Inni sugerują, że Twitter jest na tyle wymagającym i specyficznym medium, że osiągnął już maksymalne nasycenie użytkowników, zaspokajając potrzeby pewnej niszy (Goldman 2016). Kolejne próby upraszczania platformy nie sprawiły, że platforma stała się konkurencyjna wobec innych popularnych serwisów społecznościowych, a niektórych użytkowników rozżołościło do tego stopnia, że zaczęli promować ideę spółdzielczej własności Twittera. Argumentują oni, że ze względu na społeczną wartość Twittera, należy wyłączyć go spod presji Wall Street (Foley 2017; Schneider 2016; zob. też: Mikołajewska-Zajac i Rodak 2016). W związku z tą sytuacją od lat spekuluje się o sprzedaży Twittera, wskazując na takie koncerny, jak Google, Disney czy Salesforce<sup>4</sup>. Ze względu na fakt, że Twitter działa jak informacyjny sejsmograf, najbardziej prawdopodobny wydaje się zakup przez koncern medialny (Jenkins

<sup>3</sup> Kwartalne wyniki firmy można śledzić na jej stronie (Twitter 2017d).

<sup>4</sup> W lutym 2017 roku Google zdecydował się na zakup Fabric – platformy deweloperskiej do tworzenia aplikacji (Weinberger 2017).

2017). Paradoksalnie więc, „społeczna wartość” serwisu pozwala mu trwać mimo strat i braku perspektywy zysku (Goldman 2016; Jenkins 2017).

Drugą grupą, która miała znaczący wpływ na rozwój platformy, są programiści. W 2007 roku Twitter zdecydował się udostępnić swoje API (Bruns i Burgess 2016). Ekonomia API (Savitz 2012) polega na tym, że programiści mogą bez ograniczeń stwarzać aplikacje na bazie danych, podczas gdy serwis zyskuje innowacyjne dodatki. Twitter może następnie zawrzeć umowę z twórcami wartościowych rozwiązań bądź wykupić wybrane projekty. Twitter to zatem nie tyle pojedyncza platforma, ile raczej ekosystem narzędzi, wśród których można znaleźć aplikacje do zarządzania kontami albo programy do analizy danych. Wszystkie te innowacje zwiększają wartość serwisu dla użytkowników. Wśród korzystających z API znaleźli się też badacze niezależni, których analizy również pomogły firmie, przyczyniając się do lepszego zrozumienia użytkowników oraz rozpowszechnienia pozytywnej opinii na temat społecznej wagi serwisu (Bruns i Burgess 2016; Rogers 2014). Być może zresztą to właśnie nadreprezentacja naukowców wśród użytkowników Twitter przyczyniła się do nieproporcjonalnego zainteresowania tym serwisem w stosunku do innych platform społecznościowych.

Delikatna równowaga pomiędzy Twitterem a środowiskiem twórców korzystających z API została zachwiana, gdy firma postanowiła spieniężyć dostęp do danych z API, prawdopodobnie z powodu presji akcjonariatu na rentowność oraz potężnych kosztów utrzymania infrastruktury do obsługi danych. Jeszcze w 2010 roku Ryan Sarver pisał na blogu firmowym Twittera o ekosystemie, który liczył wtedy ponad 50 tysięcy aplikacji „wzbogacających doświadczenie serwisu”, że zapewnienie otwartego dostępu do danych wszystkim zainteresowanym podmiotom stanowi priorytet firmy (Sarver 2010). Mimo to zaczęto stopniowo ograniczać otwarty dostęp do danych z API, by skłonić najbardziej zasobnych klientów do zakupu danych poprzez licencjonowanych pośredników, w owym czasie przede wszystkim firmy Gnip i DataSift (Bruns i Burgess 2016). Najpierw Twitter ograniczył dostęp do archiwalnych danych przez platformę. Tweety znikaly z serwisu w na podstawie nieprzejrzystych reguł. W rezultacie tylko pobieranie tweetów tworzonych na bieżąco w serwisie gwarantuje tworzenie kompletnych korpusów danych. W 2012 roku Twitter Inc. niespodziewanie ograniczył darmowy dostęp do danych przez API. Wprowadzono limit liczby zapytań możliwych do ustawienia w tym samym czasie, a także liczby dostępnych tweetów przepływających w określonym momencie przez serwis. Ponadto wprowadzono warunek uwierzytelnienia tożsamości (Lane 2012; Sippey 2012).

W ten sposób mniej zasobne podmioty, jak również niezależni badacze, zostali w praktyce pozbawieni wolnego dostępu do większych zbiorów tweetów. Jednocześnie mniejsi deweloperzy zniechęcili się do opierania swoich narzędzi

na danych z Twittera, znalazłszy się na łasce firmy, która w każdej chwili potrafi zmienić warunki dostępu do danych (Lunden 2015). W odpowiedzi niektórzy przedstawiciele branży wystosowali petycję protestacyjną, a nawet otworzyli niezależny serwis mikroblogowy oparty na subskrypcjach i gwarantujący otwartość API, AppNet, który jednak nie zdołał zyskać większej popularności (Indvig 2012). Wreszcie w 2014 roku Twitter przejął Gnipa, a w 2015 roku zakończył współpracę z pozostałymi licencjonowanymi sprzedawcami nieograniczonych strumieni danych, planując sprzedawanie surowych danych bezpośrednio do zainteresowanych firm (Hofer-Shall 2015). W 2015 roku dochód z licencjonowania danych wzrósł o 34% w porównaniu z rokiem ubiegłym (Twitter 2016), w 2016 roku zaś o 14%, przy czym jego ogólny udział w dochodach firmy spadł i waha się w granicach 11% (Twitter 2017b). Posunięcie to nie poprawiło więc spektakularnie sytuacji firmy, za to naraziło ją na oskarżenia o wykorzystanie darmowej pracy programistów (Bucher 2013; Halstead 2015; Willmott 2015) i o niszczenie ducha otwartej współpracy (Lane 2015).

### **Wpływ Twittera na praktyki badawcze**

Dotychczasowe rozważania dotyczyły społecznego konstruowania technologii, tymczasem kolejne części tego artykułu poświęcone będą odwrotnemu kierunkowi analizy, a mianowicie wpływowi technologii na praktyki badawcze, rozumiane jako forma praktyk społecznych. Jeszcze w 2007 roku Dominik Batorski i Krzysztof Olechnicki pisali o „względnej łatwości badania Internetu (w sensie kosztów i możliwości dostępu do tego pola badawczego)”. Narzekali wówczas, że polska socjologia zbyt wolno dostosowuje się do zmian społecznych i nie wkracza wraz z badanymi do Internetu (Batorski i Olechnicki 2007: 6). Dziesięć lat później, pomimo wszechobecności Internetu oraz rosnących kompetencji badaczy, przepaść jest być może jeszcze większa. Dzieje się tak, ponieważ komputery umożliwiły powstanie nowego typu danych – Big Data.

Te nowe dane, będące szczegółowym zapisem technicznym naszej zapośredniczonej cyfrowo aktywności, stanowią wyzwanie dla socjologii na kilku płaszczynach. Po pierwsze jest to wyzwanie epistemologiczne, które polega na konieczności zrozumienia charakteru tych danych oraz ich relacji do tradycyjnych sposobów wytwarzania wiedzy w socjologii (Lazer i in. 2009; Ruppert 2013; Kitchin 2014). Z kolei wyzwanie logistyczne wynika z faktu, że w związku z prywatnym charakterem platform internetowych dane te stanowią zwykle czyjąś własność, a ponadto ich pobranie, gromadzenie i obróbka wymagają umiejętności technicznych oraz dostępu do odpowiedniego sprzętu i oprogramowania. Pojawia się również kwestia zapewnienia prywatności osób, które te dane

pozostawiają, z czym wiążą się wyzwania natury prawnej i etycznej. Wreszcie socjolog mierzy się z wyzwaniem natury politycznej. Spekuluje się, jak dalece pojawienie się tych nowych danych zmusza badaczy do przemyślenia roli ich dyscypliny w sytuacji, gdy przedstawiciele nauk ścisłych tworzą na ich bazie modele aspirujące do odkrywania „praw” rządzących światem społecznym bez odniesienia do teorii socjologicznej (Krzysztofek 2012b; zob. Barabási 2002), a dostęp do danych przeważnie wymaga dużych nakładów finansowych. Sławna stała się wypowiedź Jimmy’ego Lina z Twittera, zniechęcającego akademików do podejmowania badań, które badacze Twittera, ze względu na dostęp do licznych zasobów, mogą zrobić lepiej (zob. Boyd i Crawford 2012; Borra i Rieder 2013). Mimo to naukowcy społeczni angażują się w politykę metody, zarówno poprzez promowanie roli socjologicznej interpretacji w badaniach z użyciem Big Data (Savage i Burrows 2007), jak i poprzez walkę o równy dostęp do danych (Bruns i Burgess 2016).

Dane z serwisów społecznościowych typu Twitter to właśnie dane typu Big Data. W przypadku Twittera są to ślady aktywności użytkowników w postaci tweetów oraz metadanych powiązanych z tweetami. Twitter posiada centralny serwer, który daje firmie kontrolę nad każdym tweetem. Zakładając konto w serwisie, użytkownicy wyrażają zgodę na dalsze udostępnianie przez firmę opublikowanych przez siebie tweetów. Warunki korzystania z usługi oraz przeznaczone dla twórców oprogramowania dokument *Rules of the Road* (2012) określają charakter kontroli Twittera nad danymi. Firma może odmówić dostępu do API oraz do określonych treści, a także pod dowolnym pretekstem wycofać licencję na redystrybucję danych. Z perspektywy niezależnego badacza warto zapamiętać, że publiczne udostępnianie zbiorów danych może stanowić naruszenie warunków stawianych przez Twittera (Beurskens 2014).

Niejednoznaczna jest kwestia prywatności użytkowników Twittera. Z jednej strony przyjęło się zakładać, że użytkownicy wiedzą, iż generowane przez nich treści są publicznie dostępne, o ile nie ograniczą oni widoczności tweetów do własnych obserwatorów. Z drugiej zaś, wczesna retoryka twórców Twittera podkreślała efemeryczność tego medium. Dyskusja wokół propozycji gromadzenia tweetów przez Archiwum Biblioteki Kongresu pokazała, że wielu użytkowników pozostawało nieświadomymi możliwości gromadzenia i dalszego udostępniania wiadomości, pomimo ich publicznej natury (Zimmer i Proferes 2014). Kilka lat później, w dobie krystalizowania się modeli biznesowych platform oraz powszechnej wiedzy o wykorzystaniu danych z mediów społecznościowych, użytkownicy są z pewnością bardziej świadomi tego faktu. Badacze powinni jednak dołożyć wszelkich starań, by wyprzedzić politykę platformy oraz regulacje i chronić prywatność użytkowników w sytuacji dzielenia się zbiorami danych oraz publikowania wyników badań (Williams, Burnap i Sloan 2017).



Dostęp do danych możliwy jest z dwóch źródeł: poprzez interfejs serwisu oraz przez API. W interfejs Twittera wbudowana jest wyszukiwarka, za pomocą której można dotrzeć do tweetów spełniających określone kryteria, to znaczy słowa kluczowe, autorstwo, lokalizację i przedział czasowy. Twitter nie udostępnia jednak w ten sposób wszystkich opublikowanych wiadomości, a jedynie ich część, nie informując przy tym o charakterze ubytków w zbiorze. Im większy przedział czasowy, tym ilość dostępnych tweetów mniejsza, jako że z czasem coraz większa ich liczba „odpada z serwisu”. Ponadto, nie ma raczej możliwości pobrania tych danych inaczej niż poprzez „zeskrobywanie”, czyli kopiowanie tweetów prosto z interfejsu.

Pozyskanie danych w zestandaryzowanym formacie umożliwiającym ich komputerową analizę wymaga skorzystania z API. W literaturze z zakresu nauk społecznych zwykle opisywane są jego trzy rodzaje: *Streaming API*, *REST API* i *Search API*, który jest częścią *REST API*. *Streaming API* pozwala na gromadzenie tweetów w momencie, gdy publikowane są w serwisie. Zwykli użytkownicy mogą mieć dostęp jedynie do jednego procenta wszystkich treści generowanych w danej chwili. W przypadku wielu projektów badawczych, zwłaszcza tych o charakterze eksploracyjnym, jest to ilość w pełni wystarczająca. Stuprocentowy dostęp do danych otrzymują jedynie partnerzy biznesowi Twittera, w tym autoryzowani sprzedawcy danych. Poprzez *Streaming API* możliwe jest losowe pobranie jednego procenta wszystkich treści publikowanych w danym momencie w serwisie albo pobranie treści na podstawie specyficznych zapytań o słowa kluczowe i hasztagi, konkretnych użytkowników (do pięciu tysięcy kont) oraz lokalizację (Gaffney i Puschmann 2014).

Drugi rodzaj API – *REST (REpresentational State Transfer)* – pozwala na dostęp do innego rodzaju danych niż treść publikacji. Mowa na przykład o danych na temat obserwowania się przez użytkowników albo o tzw. trendach, czyli najpopularniejszych tematach w serwisie, mierzonych słowami kluczowymi. By kontrolować ruch, Twitter ogranicza jednak ilość zapytań, które obsługuje w ciągu jednej godziny. W końcu *Search API*, umożliwia gromadzenie archiwalnych tweetów, jednak tylko do siedmiu dni wstecz. Inaczej niż w przypadku *Streaming API*, w *Search API* pobieranie danych odbywa się co określoną jednostkę czasu, wstecznie. W związku z faktem, że *Search API* jest zaprojektowana w celu wyszukiwania najbardziej trafnych tweetów, pomija niektóre rezultaty wyszukiwania. Niemożliwe jest zebranie pełnego korpusu danych przez *REST API*, a nawet zorientowanie się w brakach skompletowanego zbioru (Gaffney i Puschmann 2014). Szczegółowy opis bieżących warunków dostępu do API znajduje się na stronie internetowej Twittera (Twitter 2017c).

W jaki sposób można pobrać dane z API inaczej niż bezpośrednio, przez zbudowanie odpowiednich narzędzi? W 2014 roku Eric Borra i Bernhard Rieger zestawili potencjalne możliwości (Borra i Rieder 2014). Wybierając jedną

z nich, należy mieć na uwadze, że każda z nich niesie ze sobą pewne szanse, ale również ograniczenia.

Po pierwsze, dane można zakupić przez autoryzowanego pośrednika, firmę Gnip. Plusem tego rozwiązania jest dostęp do wszystkich nieusuniętych tweetów spełniających kryteria wyszukiwania. Z drugiej strony, koszt takiego rozwiązania dla wielu badaczy jest zaporowy. Ponadto Gnip dostarcza surowe dane w formacie *json*, który należy przekonwertować do formatu umożliwiającego zaimportowanie ich do programów analizy danych.

Drugim i najprzystępniejszym rozwiązaniem dla „nietechnicznych” badaczy społecznych jest wykorzystanie jednej z internetowych platform do pobierania i analizy danych. Niektóre z nich stworzone zostały dla badaczy akademickich (np. DiscoverText, Truthy), inne przeznaczone są do analiz marketingowych (Topsy, Twitonomy, Hootsuite). Zdecydowaną zaletą korzystania z tych narzędzi jest łatwość obsługi oraz koszt – z niektórych można skorzystać za darmo. Platformy te mają jednak liczne ograniczenia, szczególnie istotne z perspektywy badań naukowych. Umożliwiają bowiem tylko określone operacje analityczne, ale nie dopuszczają modyfikacji do własnych celów, a niekiedy też eksportu danych do innych programów. Korzystając z takiego narzędzia, należy zwrócić uwagę, na bazie którego API jest zbudowane (Puschmann i Burgess 2014). Dla przykładu, łatwa w obsłudze aplikacja internetowa TAGS korzysta z REST API, co znaczy, że tworzy niepełne zbiory danych (TAGS 2017).

Trzecia możliwość to skorzystanie z otwartoźródłowych programów do pobierania danych: Twapper Keeper, 140kit czy współtworzonego przez Borę i Riedera DMI-TCAT. Zwłaszcza to pierwsze narzędzie zyskało swego czasu dużą popularność wśród badaczy społecznych (Yang i Kavanaugh 2011). W związku ze swoją polityką ograniczania dostępu do danych, w roku 2011 Twitter zażądał zamknięcia 140kit i TwapperKeepera, argumentując, że naruszają one regulamin, ponieważ pozwalają na dalsze dystrybuowanie tweetów (Bruns i Burgess 2016). Bazy tweetów tworzone przez użytkowników były bowiem udostępniane w sieci, by umożliwić ich wykorzystanie przez kolejnych badaczy, a także replikowanie badań. Kod tych programów jest jednak wciąż udostępniony w sieci (dla Twapper Keepera pod nazwą Your Twapper Keeper).

Z kolei DMI-TCAT (Digital Methods Initiative – Twitter Capture and Analysis Toolset) to projekt amsterdamskiej grupy badawczej Digital Methods Initiative, stworzony już po zamknięciu Twapper Keepera i 140kit. Członkowie grupy to przeważnie medioznawcy, świadomi dorobku badaczy społecznych w zakresie pracy na twitterowych danych, jak również wpływu narzędzi technicznych na charakter wytwarzanej wiedzy. Stworzony przez nich program opiera się na założeniu, że programowanie jest również teoretyzowaniem (Heiberger i Riebling 2016): program ma gwarantować przejrzystość danych oraz replikowalność badań, maksymalnie wykorzystywać możliwość bezpłatnego dostępu do danych

gwarantowane przez API, jak również umożliwić jak najszerszy wachlarz operacji analitycznych. Za pośrednictwem DMI-TCAT można zatem pobierać jeden procent losowo wybranych tweetów istniejących w serwisie w danych momencie, zbiory tweetów zawierających określone słowa kluczowe oraz do 3200 ostatnich tweetów opublikowanych przez pięć tysięcy różnych użytkowników. Program umożliwia wydzielanie podzbiorów z zebranych korpusów danych oraz porównywanie ich ze zbiorem macierzystym. Jeśli zaś chodzi o operacje analityczne, DMI-TCAT pozwala uzyskać statystyki tweetów i wskaźniki aktywności, eksportować tweety wyselekcjonowane według wybranych kryteriów do dalszej ręcznej analizy i dokonywać analizy sieci społecznych (Borra i Rieder 2014).

Wyżej opisane programy stworzone zostały po to, by umożliwić prowadzenie badań na bazie danych z Twittera, które spełniać będą standardy naukowe. Obecnie korzystanie z nich wymaga jednak posiadania serwera. DMI-TCAT w formie aplikacji internetowej jest dostępny tylko dla studentów i badaczy afiliowanych przy Wydziale Medioznawstwa Uniwersytetu Amsterdamskiego (Digital Methods Initiative 2017).

Kolejne wyzwanie stanowi uzyskanie dostępu do danych starszych niż siedem dni na sposób inny niż za pośrednictwem Twitterowej wyszukiwarki. Wydaje się, że za darmo jest to w zasadzie niemożliwe. Wyjątkiem jest opisana powyżej możliwość dotarcia do tweetów poprzez konta poszczególnych użytkowników. Tweety starsze niż 6 miesięcy są gromadzone przez Bibliotekę Kongresu Stanów Zjednoczonych, jednak, mimo wysiłków, nie znaleziono do tej pory sposobu ich udostępniania (Zimmer 2015). Pozostaje zakup danych przez Gnipa albo pośrednio do platformy analizującej dane. Na przykład firma Sifter umożliwia zakup historycznych danych do programu Discover Text, który służy do maszynowej analizy treści. To rozwiązanie jest o tyle korzystne, że dane są od razu gotowe do analizy i można stopniowo pobierać je w formacie *csv*. Gnip nie przedstawia cennika, tylko za darmo szacuje cenę na podstawie zapytania, ale koszt zakupu danych za jego pośrednictwem jest prawdopodobnie porównywalny z kosztem zakupu przez Siftera, który wynosi 20 dolarów za dzień obserwacji i 30 dolarów za 100 tys. tweetów. W przypadku Siftera, należy jeszcze doliczyć koszt licencji na oprogramowanie.

Przykład Twittera pokazuje, jak technologia pośredniczy w badaniach społecznych, wpływając na praktyki badawcze i charakter wytwarzanej wiedzy, a nawet na tożsamość dyscypliny (Latour 2010). Tak jak pojawienie się pakietu do analizy statystycznej SPSS, umożliwiając badaczom szeroko zakrojone badania empiryczne, przyczyniło się do powojennego „utwardzenia” socjologii, tak platformy społecznościowe i wytwarzane za ich pośrednictwem dane, skłaniając socjologów do wychodzenia poza tradycyjne sposoby pozyskiwania i analizy danych, być może przyczynią się do kolejnej rewolucji w łonie tej dyscypliny. Twitter tworzył swoje API z myślą o programistach; narzędzie to nie jest

więc dostosowane do potrzeb niezależnych naukowców, którzy mimo wszystko wykorzystują stwarzane przez nie możliwości. Jednak ze względu na sposób, w jaki zaprojektowana jest ta technologia, badacze zmuszeni są dostosowywać pytania badawcze do tego, jakiego rodzaju dane da się pobrać i w jaki sposób będzie można je analizować.

Istotne znaczenie ma nie tylko projekt technologii, ale również forma własności. Twitter najpierw udostępnił swoje API, „niechęć” przyczyniając się do rozkwitu badań społecznych opartych na Big Data oraz do wypracowania standardowych praktyk w badaniach nad platformą, po czym zaczął stopniowo ograniczać dostęp do danych, czym uderzył przede wszystkim w badaczy niekorporacyjnych. Sytuacja ta uprzywilejowuje aktorów komercyjnych oraz wiodące ośrodki badawcze, które dysponują możliwościami technologicznymi i programistycznymi, a jednocześnie zdolne są przyciągać środki finansowe na realizację badań (Bruns i Burgess 2016). Wytwarza się cyfrowy podział na tych, którzy mogą sobie pozwolić na badania z użyciem Big Data i tych, którzy mają znacznie mniejsze wyjściowe szanse na udział w tym badaniu (boyd i Crawford 2012; Bruns i Burgess 2016).

Specyficzny kształt technologii oraz liczne przeszkody na drodze do wykorzystania tworzonych dzięki niej danych inspirują do większej refleksyjności nad charakterem danych i stwarzanymi przez nie możliwościami badawczymi, a nawet do formułowania innowacji metodologicznych. Uwaga przesuwa się z ograniczeń na możliwości (Brooker i in. 2016), co wymaga użycia charakterystycznego dla nauk społecznych podejścia interpretatywnego do materiału badawczego, które generalnie postulują krytycy bezrefleksyjnego zachwyty nad możliwościami Big Data. Na przykład danah boyd i Kate Crawford zauważyły, że w badaniach opartych na Big Data dominuje poznawcze uprzywilejowywanie wzorów i zależności zaobserwowanych w wolumenie danych kosztem kontekstualnego zrozumienia danych oraz formułowania pytań badawczych w oparciu o teorię społeczną (boyd i Crawford 2012; zob. też: Benthall 2016). Relacje wyrażone przez zapis technologiczny, na przykład relacje obserwowania, mogą mieć różne znaczenia dla użytkowników, podczas gdy niektóre zapośredniczone interakcje w ogóle nie znajdują odzwierciedlenia w zapisie technologicznym (zob. też: Lindgren i Lundström 2011). Ponadto duży zbiór danych jest niekoniecznie pełen; zjawisko może na przykład wykraczać poza ramy platformy, z której pobrano dane (Burrell 2009; boyd i Crawford 2012).

W związku z powyższym, niektórzy badacze społeczni Twittera zaczęli formułować nowy paradygmat badawczy, który z czasem zaczął być określany jako podejście typu *small data*. Wysuwa ono na nowo na pierwszy plan osobę badacza oraz uzasadnienie decyzji, które podejmuje w odniesieniu do dostępnych danych. W podejściu tym łączy się podstawowe operacje ilościowe, analizę treści wybranych fragmentów z korpusu danych, metody jakościowe oraz wiedzę

kontekstualną spoza danego medium czy nawet spoza świata *online* (Stephansen i Couldry 2014). Badania ilościowe i jakościowe wspierają się nawzajem, by doprowadzić badacza do sformułowania pytań badawczych oraz odpowiedzi na nie (zob. boyd i Crawford 2012; Lindgren i Lundström 2011; Sergi i Bonneau 2016; Stephansen i Couldry 2014).

Podejście *small data* nie jest równoznaczne z wezwaniem do zarzucenia ilościowych analiz danych cyfrowych na rzecz metod jakościowych czy etnografii albo z rezygnacją z walki na rzecz bardziej sprawiedliwych reguł dostępu do danych. Jak przyznaje wielu badaczy Twittera, „szukanie, filtrowanie czy klasyfikowanie to jedyne sposoby, które umożliwiają badaczowi interpretację wielkiej liczby treści; stanowią zatem logiczny pierwszy krok każdej analizy, również jakościowej” (Puschmann i Burgess 2014: 65; zob. też: Marwick 2014). Metody jakościowe, takie jak wywiad, obserwacja czy analiza treści mogą służyć pogłębieniu obserwacji ilościowych (Letierce i in. 2010). Z kolei analizy ilościowe pozwalają zidentyfikować ważne skupiska społeczne i ich dynamikę, na przykład analiza aktywności trwających przez jakiś czas pomaga określić ważne momenty dla społeczności użytkującej hashtag (Thelwall 2014). Niekiedy badacz od początku jest świadomy znaczenia Twittera dla pewnej grupy i zaczyna badać jej aktywność w tym serwisie, jednak także tego typu badanie będzie znacznie wzbogacone dzięki operacjom ilościowym na danych wytworzonych przez tych użytkowników (Stephansen i Couldry 2014). Wraz z rosnącą biegłością w posługiwaniu się technologiami informacyjnymi, wzrasta zdolność badaczy do rozumienia funkcjonalności technologii jako swobodnego języka natywnego (Rogers 2013), który pod różnymi kątami rzuca światło na swoich użytkowników, ale nie wyczerpuje rozumienia ich aktywności.

### **Podsumowanie: pisanie o Twitterze jako uprawianie polityki metody**

Niniejszy tekst pisałam z myślą o polskich socjologach i badaczach społecznych, którzy chcieliby skorzystać z danych wytwarzanych w serwisie Twitter w celu badania związanych z nim zjawisk społecznych albo w innych projektach, w których dane te mogłyby okazać się przydatne. Pierwotnie artykuł ten miał zatem cel praktyczny – obniżenie progu dostępu do badań z użyciem danych z Twittera dzięki dostarczeniu informacji przydatnych do zrozumienia ich specyfiki, jednak bez pretensji do wyczerpania tematu.

Na ten pierwotny cel nałożyły się jednak dwa kolejne. Po pierwsze, jako że badawcze wykorzystanie technologii wymaga jej uprzedniego zrozumienia, posłużyłam się perspektywą konstruktywistyczną w socjologii techniki, aby uporządkować zebrane przeze mnie informacje oraz uzasadnić, dlaczego właśnie

ich znajomość może znacząco pomóc socjologom odnaleźć się w badaniach z użyciem danych z Twittera. Chodzi o to, by potraktować platformę jako efekt napięcia pomiędzy różnymi interesariuszami, jak w soczewce ukazujący szersze konflikty wbudowane w polityczną ekonomię Internetu. Jednocześnie należy dostrzec ruch w przeciwnym kierunku, to znaczy wpływ wywierany przez platformy na praktyki społeczne, w tym na praktyki badawcze. Oprócz celu praktycznego artykuł spełnia więc też funkcję rozwijania teorii społecznej: jego konkluzje można ekstrapolować na przykłady innych platform internetowych.

Po drugie, dostrzeżenie społecznej roli platform oraz ich potencjału jako źródła danych na temat zachowań społecznych skłania do krytycznego odniesienia się do polityki ich dostarczycieli oraz regulatorów Internetu. Przykład Twittera pokazuje, jak zapośredniczenie ogromnej części interakcji społecznych przez platformy własnościowe może prowadzić do wytworzenia się różnego rodzaju wykluczeń, również w dostępie do danych, a co za tym idzie – do ograniczenia dostępu do wiedzy na temat społeczeństwa. Trzecim celem artykułu jest przedstawienie krytyki kierowanej wobec obecnej polityki platform internetowych oraz uwrażliwienie badaczy na polityczne i etyczne aspekty pracy z danymi z serwisów społecznościowych.

Badania tego typu są zawsze „dokonaniem wspólnym” (Marres 2012) wielu aktorów: przedstawicieli różnych dyscyplin naukowych, dostarczycieli technologii i ich użytkowników, osób wspierających projekt od strony technicznej, merytorycznej oraz finansowej. Ten artykuł z pewnością nie powstałby, gdyby nie pomoc wielu specjalistów: napisałam go bowiem jako badaczka „nietechniczna”, przystępując do badania Twittera z minimalną wiedzą na jego temat oraz bez wykształcenia informatycznego. Nie powstałby też zresztą, gdyby nie „praca” użytkowników Twittera. Ze względu na swą radykalną transdyscyplinarność (Ruppert 2013), socjologia cyfrowa<sup>5</sup> wymaga ustawicznej nauki, ale, co za tym idzie, niesie ze sobą ryzyko błędów, nieścisłości czy przeoczeń, które jednak trzeba „wpisać w koszt” tego typu przedsięwzięć. Być może więc pozytywnym rezultatem wyzwania Big Data dla socjologii okaże się przesunięcie tej dyscypliny w stronę pracy zespołowej oraz rozwijania umiejętności translacji pomiędzy różnymi twórcami i interesariuszami wiedzy socjologicznej.

---

<sup>5</sup> Termin *computational social science* (Lazer i in. 2009), oznaczający nauki społeczne angażujące się w pracę z cyfrowymi danymi, tłumaczę za wpływową notką blogową Radosława Bomby jako „socjologia cyfrowa” (Bomba 2011), mimo że w polskiej literaturze określa się tym mianem również socjologię Internetu jako taką (zob. np. Szpunar 2016).



## Literatura

- Aljohani, Naif R., Saad A. Alahmari i Ali M. Aseere. 2011. *An Organized Collaborative Work Using Twitter in Flood Disaster*. „ACM Web Science Conference Proceedings”, 2011: 1–2; [http://www.websci11.org/fileadmin/websci/Posters/172\\_paper.pdf](http://www.websci11.org/fileadmin/websci/Posters/172_paper.pdf), dostęp: 6.09.2016.
- Annusewicz, Olgierd. 2016. *Ramowanie „dobrej zmiany”*. *Ramy językowe kampanii wyborczej Prawa i Sprawiedliwości na Twitterze w 2015 roku*. „e-Politikon” 17: 75–96.
- Barabási, Albert-László. 2002. *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Penguin Group.
- Batorski, Dominik i Krzysztof Olechnicki. 2007. *Wprowadzenie do socjologii Internetu*. „Studia Socjologiczne” 3: 5–14.
- Benthall, Sebastian. 2016. *Philosophy of Computational Social Science*. „Cosmos and History: The Journal of Natural and Social Philosophy” 12 (2): 13–30.
- Beurskens, Michael. 2014. *Legal Questions of Twitter Research*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 123–136.
- Bijker, Wiebe E., Thomas P. Hughes i Trevor J. Pinch (red.). 1987. *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge, MA: MIT Press.
- Bińczyk, Ewa. 2010. *(Post)konstruktywizm na temat technonauki*. „Zagadnienia Naukoznawstwa” 64(1): 231–251.
- Bomba, Radosław. 2011. *Socjologia cyfrowa. Nowy paradygmat nauk społecznych w gospodarce informacyjnej*. Bomba.Blog; <http://rbomba.pl/archives/1140>, dostęp: 16.05.2017.
- Borra, Erik i Bernhard Rieder. 2014. *Programmed method: developing a toolset for capturing and analyzing tweets*. „Aslib Journal of Information Management” 66(3): 262–278.
- boyd, danah i Kate Crawford. 2012. *Critical Questions for Big Data*. „Communication & Society” 15 (5): 662–679.
- boyd, danah i Nicole B. Ellison. 2008. *Social Network Sites: Definition, History, and Scholarship*. „Journal of Computer-Mediated Communication” 13: 210–230.
- Brindusa Albu, Oana i Michael Etter. 2016. *Hypertextuality and Social Media: A Study of the Constitutive and Paradoxical Implications of Organizational Twitter Use*. „Management Communication Quarterly” 30(1): 5–31.
- Brooker, Phillip, Julie Barnett, Timothy Cribbin i Sanjay Sharma. 2016. *Have We Even Solved the First ‘Big Data Challenge?’ Practical Issues Concerning Data Collection and Visual Representation for Social Media Analytics*. W: H. Snee, C. Hine, Y. Morey, S. Roberts i H. Watson (red.). *Digital Methods for Social Science. An Interdisciplinary Guide to Research Innovation*. New York: Palgrave Macmillan, s. 34–50.
- Bruns, Axel, i Jean Burgess. 2016. *Methodological Innovation in Precarious Spaces: The Case of Twitter*. W: H. Snee, C. Hine, Y. Morey, S. Roberts i H. Watson (red.).

- Digital Methods for Social Science. An Interdisciplinary Guide to Research Innovation*. New York: Palgrave Macmillan, s. 17–33.
- Bruns, Axel, i Hallward Moe. 2014. *Structural Levels of Communication on Twitter*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 15–28.
- Bucher, Taina. 2013. *Objects of Intense Feeling: The Case of the Twitter API*. „Computational Culture” 3: 1–17; <http://computationalculture.net/article/objects-of-intense-feeling-the-case-of-the-twitter-api>, dostęp: 6.09.2016.
- Burgess, Jean. 2014. *From ‘Broadcast Yourself’ to ‘Follow Your Interests’: Making Over Social Media’*. „International Journal of Cultural Studies” 18(3): 281–285.
- Burrell, Jenna. 2009. *The Field Site as a Network: A Strategy for Locating Ethnographic Research*. „Field Methods” 21 (2): 181–199.
- Chaykovsky, Kathleen. 2017a. *Jack Dorsey: Twitter Lacked ‘Focus And Discipline’ For Growth*. „Forbes” Feb. 16, <https://www.forbes.com>, dostęp: 16.05.2017.
- Chaykovsky, Kathleen. 2017b. *Twitter Shares Plummet As Revenue Growth Nearly Slows To A Halt*. „Forbes” Feb. 9, <https://www.forbes.com>, dostęp: 16.05.2017.
- Czarnowski, Stefan. 1956. *Kultura*. Warszawa: KiW.
- Dijck, José van. 2016. *The Platform Society*. #AoIR2016: Opening Keynote; <https://www.youtube.com/watch?v=ypiiSQTnqo>.
- Dijck, José van. 2011. *Tracing Twitter: The Rise of a Microblogging Platform*. „International Journal of Media & Cultural Politics” 7(3): 333–348.
- Długosz, Dawid. 2014. *Twitter w liczbach. Jak przedstawiają się statystyki popularnej społecznościówki?*. „KomputerŚwiat”; <http://www.komputerswiat.pl/artykuly/redakcyjne/2014/12/twitter-w-liczbach.aspx>, dostęp: 10.07.2016.
- Dzida, Damian i Adam Markowski. 2013. *Rozpowszechnianie opinii w sieciach złożonych. Model wirusowy i jego weryfikacja na podstawie danych z Twittera*. „Symulacja w Badaniach i Rozwoju” 4(1): 5–21.
- Foley, Stephen. 2017. *Twitter Shareholders Voice Benefits of a Co-Operative*. „Financial Times” Feb. 3, <https://www.ft.com>.
- Fuchs, Christian i Sebastian Sevignani. 2013. *What Is Digital Labour? What Is Digital Work? What’s Their Difference? And Why Do These Questions Matter for Understanding Social Media?*. „TripleC” 11(2): 237–293.
- Gadkari, Pia. 2013. *How Does Twitter Make Money?*. „BBC”; <http://www.bbc.com/news/business-24397472>, dostęp: 31.08.2016.
- Gaffney, David i Cornelius Puschmann. 2014. *Data Collection on Twitter*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 55–67.
- Goldman, David. 2016. *10 Years Later, Twitter Still Isn’t Close to Making Money*. „CNN Money”; <http://money.cnn.com/2016/03/21/technology/twitter-10th-anniversary/>, dostęp: 31.08.2016.
- Halavais, Alexander. 2014. *Structure of Twitter: Social and Technical*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 29–42.
- Halstead, Nick. 2015. *Twitter Ends Its Partnership with DataSift – Firehose Access Expires on August 13, 2015*. DataSift Blog; <http://blog.datasift.com/2015/04/11/twitter->

- ends-its-partnership-with-datasift-firehose-access-expires-on-august-13-2015/, dostęp: 23.08.2016.
- Heiberger, R. H. i J. R. Riebling. 2016. *Installing Computational Social Science: Facing the Challenges of New Information and Communication Technologies in Social Science*. „Methodological Innovations” 9: 1–11.
- Hofer-Shall, Zach. 2015. *Working Directly With the Twitter Data Ecosystem*. Gnip Blog; <https://blog.gnip.com/twitter-data-ecosystem/>, dostęp: 23.08.2016.
- Honeycutt, Courtenay i Susan C. Herring. 2009. *Beyond Microblogging: Conversation and Collaboration via Twitter*. W: *Proceedings of the 42nd Hawaii International Conference on System Sciences*, s. 1–10.
- Indvig, Lauren. 2012. *Online Petition Asks Twitter to Keep Ecosystem Open*. „Mashable” Ang. 7, <http://mashable.com>, dostęp: 23.08.2016.
- Jemieliński, Dariusz. 2013. *Netnografia, czyli etnografia wirtualna – nowa forma badań etnograficznych*. „Prakseologia” 154: 97–116.
- Jenkins, Lisa D. 2017. *12 Social Media Marketing Predictions for 2017 From the Pros*. „Social Media Examiner” Feb. 15, <http://www.socialmediaexaminer.com>.
- Kaźmierczak, Marek. 2012. *Użytkownik, nadawca i odbiorca w Web 2.0. Uwagi o różnych sposobach odnoszenia się do literatury w serwisie Twitter*. „Teksty Drugie” 6: 270–286.
- Kepes, Ben. 2014. *Twitter Buys Gnip - It's All About The Data*. „Forbes” Apr. 15, <http://www.forbes.com>, dostęp: 31.08.2016.
- Kitchin, Rob. 2014. *Big Data, New Epistemologies and Paradigm Shifts*. „Big Data & Society” 1: 1–12.
- Krzysztofek, Kazimierz. 2012a. *Big Data Society. Technologie samoopiszu i samopokazu: ku humanistycy cyfrowej*. „Kultura i Historia” 1–4(72–75): 223–257.
- Krzysztofek, Kazimierz.. 2012b. *Zmiana permanentna? Refleksje o zmianie społecznej w epoce technologii cyfrowych*. „Studia Socjologiczne” 4(207): 7–39.
- Kubczak, Paulina. 2014. *Kandydaci na prezydentów miast Górnośląskiego Związku Metropolitalnego na portalach Facebook i Twitter*. „Vademecum Śląsk” 2: 253–267.
- Lane, Kim. 2012. *Twitter Continues to Restrict Access to Our Tweets*. „API Evangelist”; <http://apievangelist.com/2012/06/29/twitter-continues-to-restrict-access-to-our-tweets/>, dostęp: 23.08.2016.
- Lane, Kim. 2015. *On Twitter, Gnip, DataSift, And Making The Hard Platform Decision*. „API Evangelist”; <http://apievangelist.com/2015/04/16/on-twitter-gnip-datasift-and-making-the-hard-platform-decision/>, dostęp: 23.08.2016.
- Latour, Bruno. 2010. *Splatając na nowo to, co społeczne*. Tłum. A. Derra i K. Abri-szewski. Warszawa: Universitas.
- Lave, Jean i Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.
- Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, László-Barabási, Albert Devon Brewer, Nicholas Christakis, Noshir Contractor, Marshall Van Alstyne James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy i Deb Roy. 2009. *Computational Social Science*. „Science” 323: 721–723.

- Letierce, Julie, Alexandre Passant, John G. Breslin i Stefan Decker. 2010. *Using Twitter During an Academic Conference: The #iswc2009 Use-Case*. W: *Proceedings of the Fourth International {AAAI} Conference on Weblogs and Social Media*, s. 279–282.
- Lewis, Bex i David Rush. 2013. *Experience of Developing Twitter-Based Communities of Practice in Higher Education*. „Research in Learning Technology” 21: 18598 <http://dx.doi.org/10.3402/rlt.v21i0.18598>.
- Liang, Hai i King-Wa Fu. 2015. *Testing Propositions Derived from Twitter Studies: Generalization and Replication in Computational Social Science*. „PLoS ONE” 10(8): 1–14.
- Lindgren, Simon i Ragnar Lundström. 2011. *Pirate Culture and Hacktivist Mobilization: The Cultural and Social Protocols of #WikiLeaks on Twitter*. „New Media & Society” 13(6): 999–1018.
- Luckerson, Victor. 2013. *Twitter Goes Public*. „Time”; <http://business.time.com/2013/11/07/live-updates-twitter-goes-public/>, dostęp: 31.08.2016.
- Lunden, Ingrid. 2015. *Twitter Cuts Off DataSift To Step Up Its Own Big Data Business*. „Tech Crunch”; <https://techcrunch.com/2015/04/11/twitter-cuts-off-datasift-to-step-up-its-own-b2b-big-data-analytics-business/>, dostęp: 23.08.2016.
- Łapczyński, Mariusz. 2016. *Możliwości wykorzystania komentarzy klientów banku w współtworzeniu wartości – zastosowanie eksploracji opinii (opinion mining)*. „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu” 459: 121–129.
- Machnik, Bartłomiej. 2014. *Twitter jako „broń taktycznego rażenia”. Rola serwisu w przekazywaniu informacji na przykładzie „afery podsłuchowej”*. *Kwartalnik Naukowy OAP UW „e-Politikon”* 12: 150–176.
- Manovich, Lev. 2011. *Trending: The Promises and the Challenges of Big Social Data*. W: M.K. Gold (red.). *Debates in the Digital Humanities*. Minneapolis and London: University of Minnesota Press, s. xx.
- Marjak, Henryk. 2014. *Potencjał i rola mediów społecznościowych w zarządzaniu kryzysowym na kolejnych etapach sytuacji kryzysowej*. „Logistyka” 5: 1009–1018.
- Marres, Noortje. 2012. *The Redistribution of Methods: On Intervention in Digital Social Research, Broadly Conceived*. „Sociological Review” 60(S1): 139–165.
- Marwick, Alice E. 2014. *Ethnographic and Qualitative Research on Twitter*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 109–122.
- Mcgreneire, Joanna i Wayne Ho. 2000. *Affordances: Clarifying and Evolving a Concept*. W: *Proceedings of Graphics Interface*. 2000: Montreal, Québec, Canada, 15–17 May 200, 179–186.
- Michalak, Joanna. 2016. *Wykorzystanie danych tekstowych do monitorowania nastrojów i sposobu postrzegania marki na przykładzie portali społecznościowych*. „Acta Universitatis Nicolai Copernici” 1: 139–155.
- Mikołajewska-Zajac, Karolina i Olga Rodak. 2016. *Platformy spółdzielcze jako próba rewizji korporacyjnego modelu gospodarki współdzielonej*. „E-mentor” 4(66): 67–73.

- Moe, Hallvard. 2012. *Who Participates and How? Twitter as an Arena for Public Debate about the Data Retention Directive in Norway*. „International Journal of Communication” 6 (1): 1222–1244.
- Orlikowski, Wanda J. 2007. *Sociomaterial Practices: Exploring Technology at Work*. „Organization Studies” 28: 1435–1448.
- Orlikowski, Wanda J. 1992. *The Duality of Technology: Rethinking the Concept of Technology in Organizations*. „Organization Science” 3(3): 398–427.
- Orlikowski, Wanda J. i Daniel Robey. 1991. *Information Technology and the Structuring of Organizations*. „Information Systems Research” 2(2): 143–169.
- Pokorna-Ignatowicz, Katarzyna, Stanisław Jędrzejewski, i Joanna Bierówka. 2013. *Nowe media a praktyki komunikacyjne*. Kraków: Krakowska Akademia im. Andrzeja Frycza Modrzewskiego.
- Poprawa, Magdalena. 2017. *Twitter w Polsce. Podsumowanie 2016 r.* Sotrender; <https://www.sotrender.com/blog/pl/2017/01/twitter-w-polsce-podsumowanie-2016-r-infografika/>, dostęp: 17.05.2017.
- Puschmann, Cornelius i Jean Burgess. 2014. *The Politics of Twitter Data*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 43–54.
- Rogers, Richard. 2013. *Digital Methods*. Cambridge, London: The MIT Press.
- Rogers, Richard. 2014. *Debanalising Twitter: The Transformation of an Object of Study*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. ix–xxvii.
- Ruppert, E. 2013. *Rethinking Empirical Social Sciences*. „Dialogues in Human Geography” 3(3): 268–273.
- Rzeszewski, Michał. 2015. *Cyberpejzaż miasta w trakcie megawydarzenia: Poznań, Euro 2012 i Twitter*. „Studia Regionalne i Lokalne” 1(59): 123–137.
- Sarver, Ryan. 2010. *Enabling a Rush of Innovation*. Twitter Blog; <https://blog.twitter.com/2010/enabling-a-rush-of-innovation>, dostęp: 23.08.2016.
- Savage, Mike i Roger Burrows. 2007. *The Coming Crisis of Empirical Sociology*. „Sociology” 41(5): 885–899.
- Savitz, Eric. 2012. *Welcome To The API Economy*. „Forbes”; <http://www.forbes.com/sites/ciocentral/2012/08/29/welcome-to-the-api-economy/#62e937546d39>, dostęp: 10.07.2016.
- Schneider, Nathan. 2016. *Here's My Plan to Save Twitter: Let's Buy It*. „The Guardian”; <https://www.theguardian.com/commentisfree/2016/sep/29/save-twitter-buy-platform-shared-ownership>.
- Seegerberg, Alexandra i W. Lance Bennett. 2011. *Social Media and the Organization of Collective Action: Using Twitter to Explore the Ecologies of Two Climate Change Protests*. „The Communication Review” 14: 197–215.
- Sergi, Viviane i Claudine Bonneau. 2016. *Making Mundane Work Visible on Social Media: A CCO Investigation of Working out Loud on Twitter*. „Communication Research and Practice” 2(3): 378–406.
- Simonetto, Deana. 2016. *Expanding Our Methodological Toolbox: The 'place' of Twitter in the Ethnographic Endeavor*. „Qualitative Sociology Review” 12(1): 98–112.



- Sippey, Michael. 2012. *Changes Coming in Version 1.1 of the Twitter API*. Twitter Blog; <https://blog.twitter.com/2012/changes-coming-in-version-11-of-the-twitter-api>, dostęp: 23.08.2016.
- Skrzypek, Adam. 2016. *Potencjal rozwoju Twittera w Polsce i wykorzystania go przez marki*. „Nowy Marketing”; <http://nowymarketing.pl/a/7969,potencjal-rozwoju-twittera-w-polsce-i-wykorzystania-go-przez-marki>, dostęp: 16.05.2017.
- Stephansen, Hilde C. i Nick Couldry. 2014. *Understanding Micro-Processes of Community Building and Mutual Learning on Twitter: A ‘small Data’ Approach*. „Information, Communication & Society” 17(10): 1212–1227.
- Sulek, Antoni. 2002. *Ogród metodologii socjologicznej*. Warszawa: Wydawnictwo Naukowe Scholar.
- Szews, Przemysław. 2014. *Facebook, Twitter i YouTube w mediach tradycyjnych. Jak prasa, radio i telewizja wykorzystują serwisy społecznościowe?*. „Media i Społeczeństwo” 4: 56–73.
- Szpunar, Magdalena. 2016. *Humanistyka cyfrowa a socjologia cyfrowa. Nowy paradygmat badań naukowych*. „Zarządzanie w Kulturze” 17(4): 355–369.
- Tereszkiewicz, Anna. 2017. *Przykro nam to słyszeć :( – reakcje firm telekomunikacyjnych na skargi i zażalenia klientów publikowane w serwisie Twitter*. „Poradnik Językowy” 3: 17–29.
- Thelwall, Mike. 2014. *Sentiment Analysis and Time Series with Twitter*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 83–96.
- Trist, Eric i Ken Bamforth. 1951. *Some Social and Psychological Consequences of the Longwall Method of Coal Getting*. „Human Relations” 4: 3–38.
- van Dijck, José. 2011. *Tracing Twitter: The rise of a microblogging platform*. „International Journal of Media & Cultural Politics” 7(3): 333–348.
- van Dijck, José. 2016. *The Platform Society*. #AoIR2016: Opening Keynote.
- Wachnicki, Michał. 2013. *Dlaczego Polacy Nie Lubią Twittera?*. „Newsweek”; <http://www.newsweek.pl/polska/dlaczego-polacy-nie-korzystaja-z-twittera-newsweek,artykuly,270698,1.html>, dostęp: 10.07.2016.
- Weinberger, Matt. 2017. *Twitter just sold its developer platform to Google*. „Business Insider”; <http://www.businessinsider.com/google-buys-twitter-fabric-2017-1?IR=T>, dostęp: 17.05.2017.
- Weller, Katrin, Axel Bruns, Jean Burgess, Merja Mahrt i Cornelius Puschmann. 2014. *Twitter and Society: An Introduction*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Puschmann (red.). *Twitter and Society*. New York: Peter Lang, s. xxix–xxxviii.
- Wellman, Barry. 1979. *The Community Question*. „American Journal of Sociology” 84: 1201–1231.
- Wellman, Barry. 1999. *The Network Community*. W: B. Wellman (red.). *Networks in the Global Village*. Boulder, CO: Westview, s. 1–48.
- Wellman, Barry. 2001. *Physical Place and Cyber-Place: Changing Portals and the Rise of Networked Individualism*. „International Journal for Urban and Regional Research” 25(2): 227–252.



- Williams, Ev. 2013. *Twitter's Early Years*. „Inc”; <http://www.inc.com/issie-lapowsky/ev-williams-twitter-early-years.html?cid=em01011week40day04b>, dostęp: 31.08.2016.
- Williams, Matthew L., Pete Burnap i Luke Sloan. 2017. *Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking Into Account Users' Views, Online Context and Algorithmic Estimation*. „Sociology”, doi: 10.1177/0038038517708140.
- Willmott, Steve. 2015. *How Not to Be a Platform: Twitter's Firehose Mistake*. „3Scale”; <https://www.3scale.net/2015/04/losing-innovation-twitters-firehose-mistake/>, dostęp: 31.08.2016.
- Wolny, Wiesław. 2016. *Sentiment analysis of Twitter data using emoticons and emoji ideograms*. „Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach” 296: 163–171.
- Yang, Seungwon i Andrea L. Kavanaugh. 2011. *Half-Day Tutorial : Collecting, Analyzing and Visualizing Tweets Using Open Source Tools*. „Digital Government” 11: 374–375.
- Zagórska, Martyna Małgorzata. 2014. *Dziennikarze w mediach społecznościowych. Studium przypadku – profile Krzysztofa Skórzyńskiego (Fakty TVN) na Facebooku i Twitterze*. „Civitas et Lex” 3: 25–37.
- Zimmer, Michael i Nicholas Proferes. 2014. *Privacy on Twitter, Twitter on Privacy*. W: K. Weller, A. Bruns, J. Burgess, M. Mahrt i C. Punschmann (red.). *Twitter and Society*. New York: Peter Lang, s. 169–182.
- Zimmer, Michael. 2015. *The Twitter Archive at the Library of Congress: Challenges for Information Practice and Information Policy*. „First Monday. Peer-Reviewed Journal on the Internet” 20(7), <http://firstmonday.org/article/view/5619/4653>, dostęp: 6.09.2016.

## Źródła internetowe

- Digital Methods Initiative. 2017. Twitter Capture and Analysis Toolset (DMI-TCAT), <https://wiki.digitalmethods.net/Dmi/ToolDmiTcat>, dostęp: 7.05.2017.
- Facebook. 2017. Investor Relations, <https://investor.fb.com/home/default.aspx>, dostęp: 10.04.2017.
- Internet Live Stats. 2017. Twitter Usage Statistics, <http://www.internetlivestats.com/twitter-statistics>, dostęp: 10.07.2016.
- Politechnika Warszawska. 2015. Wspólny projekt Politechniki Warszawskiej i Uniwersytetu Stanforda, <https://www.pw.edu.pl/Badania-i-nauka/Aktualnosci/Wspolny-projekt-Politechniki-Warszawskiej-i-Uniwersytetu-Stanforda>, dostęp: 16.05.2017.
- TAGS. 2017. About, <https://tags.hawksey.info/>, dostęp: 31.04.2017.
- Twitter. 2017a. O firmie, <https://about.twitter.com/pl/company>, dostęp: 10.04.2017.
- Twitter. 2017b. Q4 2016. Letter to Shareholders, [http://files.shareholder.com/downloads/AMDA-2F526X/4445498605x0x927279/AC2897AE-A9F6-43A7-B150-52300045809C/Q416\\_ShareholderLetter.pdf](http://files.shareholder.com/downloads/AMDA-2F526X/4445498605x0x927279/AC2897AE-A9F6-43A7-B150-52300045809C/Q416_ShareholderLetter.pdf), dostęp: 10.05.2017.

- Twitter. 2017c. Twitter Developer Documentation, <https://dev.twitter.com/overview/api>, dostęp: 7.05.2017.
- Twitter. 2016. Twitter Q1 2016 Shareholder Letter, [https://www.sec.gov/Archives/edgar/data/1418091/000156459016016642/twtr-ex991\\_6.htm](https://www.sec.gov/Archives/edgar/data/1418091/000156459016016642/twtr-ex991_6.htm) dostęp: 31.08.2016.
- Wirtualne Media. 2016. Twitter z 4 mln polskich użytkowników. Coraz więcej starszych, dobrze wykształconych i z miast, <http://www.wirtualnemedia.pl/artykul/twitter-z-4-mln-polskich-uzytownikow-coraz-wiecej-starszych-dobrze-wykształconych-i-z-miast>, dostęp: 17.05.2017.

### **Twitter as an Object of Sociological Inquiry and a Source of Social Data: Constructivist Perspective**

#### Summary

Twitter received special attention of social scientists for at least two reasons. Firstly, this social networking site is a social phenomenon, influencing communication and organizational practices. Secondly, Twitter is a repository of structured data, which may be used in order to study phenomena either related or not directly related with social impact of this medium. The purpose of the following paper is to summarize knowledge, which, in the author's opinion, is necessary to smoothly and fruitfully use Twitter for conducting sociological inquiry: either this, in which Twitter is an object of analysis, or the one in which the service only delivers analytical data. As technology reconfigures social practices, being at the same time shaped by those practices, sociology of Twitter demands first of all understanding of the character of this technology as well as the influence of this technology on research practices, which are here understood as social practices. The article shows transformative character of Twitter in these two aspects. Constructivist perspective enables to understand the medium itself, as well as the character and limitations of data produced through it.

Key words: Twitter; digital sociology; computational social science; social network sites; social media data; Big Data; small data.