

**Bartłomiej Łódzki**

<https://orcid.org/0000-0003-3526-6485>

Wydział Nauk Społecznych  
Uniwersytet Wrocławski

## Rola *big data* w kampaniach wyborczych

**Słowa kluczowe:** *big data*, komunikowanie polityczne, kampanie wyborcze, social media, analityka danych

**Key words:** Big Data, political communication, political campaign, social media, data analyst

### Wstęp

Ewolucja technologii cyfrowych wiąże się ze stałym wzrostem generowanych i przetwarzanych danych. Skalę zmian ilustrują liczby. W 1992 roku na świecie powstawało 100 GB danych dziennie, w 1997 roku 100 GB generowano w ciągu godziny, a w 2002 roku potrzebowano na to sekundę. W 2018 roku, według prognoz, przyrost danych miał wynieść 50 000 GB na sekundę (Samsung 2017). Procesy te niosą ze sobą konieczność tworzenia nowej infrastruktury technicznej nie tylko przez podmioty komercyjne, ale też państwowe. Celem wprowadzanych zmian jest usprawnienie przepływu informacji i tworzenie przestrzeni do przetwarzania danych i wyciągania z nich niezbędnej wiedzy. Mechanizmy te określane są jako *big data*. Sam termin pochodzi z sektora komercyjnego (technicznego). Współcześnie *big data* wykorzystuje się na wielu płaszczyznach. Od kilkunastu lat zjawisko to intensywnie wkracza w obszar komunikowania politycznego, w szczególności komunikowania marketingowego (Pawelczyk i Jakubowski 2017: 34–44).

Termin „big data” bywa rozumiany różnie zależnie od branży, u jego podstaw kryją się jednak podobne mechanizmy. Przede wszystkim *big data* to proces polegający na pobieraniu, przetwarzaniu, analizowaniu i wizualizacji danych. To także porównywanie danych i wyciąganie z nich wniosków. Pojęcie to odnosi się zwykle do dużego zestawu danych, który jest prawie niemożliwy do zarządzania przy użyciu tradycyjnych narzędzi (Halevi 2012: 3). Timo Elliott (2013) zwraca uwagę na rozwój pojęcia w czasie. Pierwotnie wiązało się ono z trzema „v” (ang. *volume, velocity, variety* – rozumiane jako objętość, szybkość i różnorodność). Aktualnie z *big data* identyfikowane są kolejne „v” (w tym *validity, veracity, value, visibility* – ważność, wiarygodność, jakość i widoczność).

Objętość jest jedną z najbardziej rozpoznawalnych cech. Odnosi się ona do dużych zbiorów danych. Jak zauważa George Firican (2017), 90% wszystkich dzisiejszych danych powstało w ciągu ostatnich kilku lat. Szybkość odnosi się

do prędkości, z jaką dane są generowane i przetwarzane. Różnorodność wiąże się zaś z rodzajami danych, które są poddawane analizom. Dane strukturalne zdominowane zostają przez dane półstrukturalne, a przede wszystkim niestrukturyzowane (zróznicowane i rozproszone). Mogą to być informacje w postaci dźwięków, obrazów, materiałów wideo, pochodzące z mediów społecznościowych, a także wytwarzane przez urzędy. Kolejną cechą jest zmienność danych. Możliwe, że informacje będą zmieniać źródła swojego pochodzenia, ulegać anomalii czy rozproszeniu. Prawdziwość/wiarygodność (*veracity*) zaczyna odgrywać coraz istotniejszą rolę. W natłoku ilości danych zwiększa się trudność weryfikowania ich źródeł. Umiejętność szybkiej oceny, analizy i wyciągnięcia wniosków jest wyzwaniem w dobie *big data*. Z wiarygodnością wiąże się bezpośrednio kolejna cecha, jaką jest ważność (*validity*). Podejmowanie wiarygodnych decyzji bez brania pod uwagę ważnych (istotnych) danych jest niemożliwe lub może prowadzić do negatywnych następstw. Widoczność (*visibility*) to dostarczanie i przedstawianie informacji w sposób, który ma sens (jest zrozumiały dla odbiorców). Analitycy danych poświęcają coraz więcej czasu na „czyszczenie” i filtrowanie danych celem uzyskania przejrzystych wyników. Zdaniem Georga Firicana (2017) prawdopodobnie najważniejszą ze wszystkich cech jest wartość (*value*). Inne funkcje *big data* nie mają znaczenia, jeśli nie niosą ze sobą wartości biznesowej pochodzącej z danych. Firican rozumie przez to lepsze zrozumienie klientów, optymalizację procesów i poprawę wydajności maszyn lub instytucji. Mechanizmy te coraz częściej są rozumiane i wykorzystywane także przez podmioty polityczne w trakcie kampanii wyborczych.

## **Big data w kampaniach wyborczych**

Prowadzenie współczesnych kampanii wyborczych na podstawie danych płynących w głównej mierze z badań sondażowych staje się rozwiązaniem przestarzałym i nieefektywnym. Klasyczne formy gromadzenia danych są czasochłonne, kapitałochłonne i wymagające dużych mocy przerobowych. Stratedzy polityczni potrzebują coraz więcej i bardziej rozbudowanych informacji o wyborcach. Szybkość gromadzenia tych informacji zaczyna odgrywać coraz większą rolę. Procesy te wymuszają wprowadzenie do kampanii pogłębionej analityki danych. Koszty jej implementacji do niedawna były bardzo wysokie, jednak obecnie stopniowo się to zmienia. Finansowanie zakupu, przechowywanie i zarządzanie danymi nie stanowi już tak istotnej bariery. Jedną z tych, które pozostają, wiąże się z brakiem w szeregach partii politycznych odpowiednio wykwalifikowanych ekspertów. Słabość ta wymusza zatrudnienie specjalistów posiadających kompetencje i umiejętności prognozowania i modelowania danych. Ich wsparcie jest niezbędne ze względu na fakt, że podmioty polityczne coraz częściej decydują się na budowanie, utrzymywanie i regularne rozszerzanie własnych baz danych o wyborcach (Nickerson i Rogers 2014: 51–74). Analityka danych pozwala na tworzenie modeli przewidujących zachowania obywateli,

monitorowanie potencjalnych zmian ich zachowań, a tym samym efektywne przygotowywanie strategii lub wprowadzanie na bieżąco korekt w zaplanowanych działaniach. Technologie *big data* pozwalają na ukierunkowanie kampanii na konkretne osoby, a następnie na agregację prognoz. Działania prowadzone przez podmioty polityczne są najczęściej traktowane jako „tajna broń” kampanii. Rodzaj i zakres gromadzonych danych są ściśle chronione. Powoduje to, że unika się prowadzenia publicznych debat na ten temat, co skutkuje powstawaniem wielu spekulacji i obaw o prywatność wyborców. Zwracano na to uwagę między innymi w artykule zamieszczonym w gazecie „New York Times”: „Strategowie powiązani z kampaniami Obamy i Romneya mówią, że mają dostęp do informacji o życiu osobistym wyborców na skalę, której nigdy wcześniej sobie nie wyobrażano. I używają tych danych, by próbować wpływać na nawyki głosujących – w efekcie szkolą wyborców, by poszli do urn, za pomocą subtelnych wskazówek, nagród i gróźb w sposób zbliżony do działań firm marketingowych wydających karty kredytowe” (Duhigg 2012).

Gromadzone dane wykorzystywane są do konstruowania modeli predykcyjnych, dzięki którym kampanie stają się jeszcze bardziej efektywne. Modele te generują trzy kategorie wyników prognozowania. Należą do nich wskaźniki: zachowań, wsparcia i relacji. W przypadku wskaźnika zachowań wykorzystuje się informacje o wcześniejszym zachowaniu (aktywności), łącząc je z danymi demograficznymi. Służą one do obliczenia prawdopodobieństwa zaangażowania się wyborców w określone formy działalności politycznej: prognozują frekwencję, wynik wyborczy, a także zaangażowanie w wolontariat czy przekazanie darowizn. Dzięki wskaźnikowi wsparcia przewiduje się preferencje polityczne, natomiast wskaźnik relacji pozwala na zbadanie zainteresowania wyborców określonymi tematami kampanii.

## Źródła danych

Kampania prezydencka Baracka Obamy była jedną z pierwszych, w trakcie których analitykę danych użyto na tak szeroką skalę. W opinii ekspertów „Prawdziwym zwycięzcą wyborów w 2012 została analityka” (Shen 2013). Według danych serwisu Datafloq przez półtora roku przed dniem wyborów w listopadzie 2012 roku sztab Obamy zebrał i wykorzystał ponad 1,5 mld dolarów. Przy kampanii pracowało ponad tysiąc opłacanych pracowników i 2,2 miliona ochotników. Wykonano też ponad 100 analiz danych, w trakcie których przeprowadzono ponad 66 tysięcy symulacji komputerowych. Celem działań prowadzonych przez Jima Messina było „zmierzenie wszystkiego”. Dążył, by znać każdy szczegół o tym, co się wydarzyło w trakcie kampanii. Pozyskane informacje wykorzystywał do symulacji i poprawy skuteczności prowadzonych aktywności. Zgromadzono łącznie 10 TB danych, co wówczas było liczącym się woluminem. Dane te pochodziły z wielu różnych kanałów, powodując znaczne rozdrobnienie informacji. Przetwarzano dane terenowe (preferencje wyborcze) i finansowe (pochodzenie pozyskiwanych

funduszy). Za pomocą mediów społecznościowych monitorowano proces rekrutacji wolontariuszy. Analizie poddawano przekaz medialny i kupowano odpowiedni czas antenowy (Datafloq 2013). Dane pochodzące z mediów społecznościowych i Internetu odegrały główną rolę. Proces ich przetwarzania rozpoczynał się w chwili otrzymania dostępu do profili zwolenników – zarówno na Facebooku, jak i w innych portalach społecznościowych. Na ich podstawie tworzono listy kontaktów. Równolegle kupowano i przetwarzano dane demograficzne od firm, które analizują takie szczegóły, jak historia zakupów, skłonność do hazardu, zainteresowanie programami szybkiego wzbogacania się, problemy finansowe czy obecność na portalach randkowych. Oprogramowanie w postaci plików *cookies* instalujące się na komputerach wyborców (zwolenników i wolontariuszy) pozwalało sprawdzać, jakie strony przeglądają. Analizowano przykładowo, czy są to treści religijne, czy erotyczne. Określano w ten sposób kwestie związane z moralnością internautów. Wyborcy przeglądający witryny religijne po powrocie na stronę Mittromney.com lub Barackobama.com mogli być powitani przez wiadomości przyjazne religii (Duhigg 2012).

W trakcie amerykańskiej kampanii prezydenckiej w 2016 roku analityka danych weszła na kolejny poziom. Według ekspertów rynkowych firma Cambridge Analytica, pracująca na rzecz kampanii Donalda Trumpa, zgromadziła ponad pięć tysięcy informacji na temat każdego Amerykanina (Datafloq 2018). Proces wyglądał niewinnie, rozpoczęto bowiem od przeprowadzenia na Facebooku ankiety, która została wypełniona przez setki tysięcy Amerykanów. Na podstawie zgromadzonych danych specjaliści z Cambridge Analytica byli w stanie stworzyć model przewidujący osobowość dorosłych obywateli Stanów Zjednoczonych. To z kolei umożliwiało dostarczanie użytkownikom portalu spersonalizowanych treści dotyczących kampanii wyborczej. Cambridge Analytica, za pomocą oprogramowania i mechanizmów uczenia maszynowego, zoptymalizowała proces wyświetlania reklam dopasowanych do preferencji politycznych milionów potencjalnych wyborców. Analiza danych pozwalała zrozumieć, jakie treści zostaną najlepiej przyjęte przez odbiorców (Hunter 2017). Hanes Grassegger i Michael Krogerus (2017) twierdzą, że Cambridge Analytica gromadziła informacje także ze źródeł zewnętrznych. Dane osobowe kupowano między innymi z rejestrów gruntów, portali zajmujących się sprzedażą części samochodowych czy monitoringiem zakupów. Gromadzono nawet informacje o członkostwie w poszczególnych klubach czy stowarzyszeniach. Oczyszczone i przefiltrowane cyfrowe ślady przekształcano w praktyczne informacje. Na ich podstawie przygotowywano treści wyborcze. Personalizacja polegała między innymi na takich szczegółach, jak zmiany w nagłówkach, kolorach czy podpisach pod tekstami wyświetlanymi w Sieci.

Od lipca 2016 roku sztab Trumpa dysponował specjalną aplikacją mobilną, która agregowała wszystkie treści i tworzyła profil psychologiczny użytkowników. Pozwalała ona na bieżące tworzenie przekazów na temat kampanii i integrowanie wokół nich zwolenników kandydata. W wyniku analiz udało się podzielić populację USA na 32 typy osobowości. Przekaz kampanijny wzmocniano był telefonami lub domowymi wizytami. Kontaktowano się nie tylko

z osobami korzystającymi z aplikacji i będącymi zwolennikami Trumpa, ale też z użytkownikami, którzy odpowiednio zareagowali na wybrane treści. Tematy rozmów dostosowywano do osobowości wyborców i zgromadzonych o nich treści. Po każdym kontakcie lub spotkaniu wprowadzano do aplikacji kolejne szczegóły. Dane te poddawane były dalszej analizie celem usprawnienia prowadzonej kampanii.

Omawiana mobilna aplikacja została stworzona przez Thomasa Petersa i jego firmę uCampaign. Deweloper zaprojektował aplikację w taki sposób, że za jej pomocą możliwe było przebicie się z przekazem wyborczym przez dziesiątki lub setki innych informacji, jakie otrzymują użytkownicy Facebooka każdego dnia. Newsy przygotowane na rzecz Partii Republikańskiej nie tylko były wysoce spersonalizowane, ale też wzmacniane wypowiedziami lub autorytetem bliskiej osoby. Komunikaty tego typu wysyłano użytkownikom w postaci treści polubionych wcześniej przez ich znajomych lub członków rodziny. Przekaz wzmacniano wysyłanym osobno SMS-em w imieniu bliskiej osoby. Jak zauważa twórca aplikacji, wiadomość dostarczana w imieniu kogoś z rodziny niosła ze sobą dodatkowy ładunek emocjonalny (Tani 2016).

Sztab Hillary Clinton również dysponował własną aplikacją podczas kampanii w 2016 roku. Jej autorami byli twórcy DreamWorks Animation. Choć miała służyć Clinton do podejmowania działań, koncentrowała się głównie na integracji środowiska wolontariuszy. Za pomocą gier społecznościowych mogli oni konkurować ze sobą i zdobywać zarówno wirtualne, jak i rzeczywiste nagrody. Aplikacja gromadziła dane o użytkownikach i zachęcała do podzielenia się listą kontaktów, ale nie agregowała danych jak w przypadku narzędzia wykorzystywanego przez sztab Trumpa (Fried 2016).

W trakcie kampanii dotyczącej wyjścia Wielkiej Brytanii ze struktur Unii Europejskiej firma uCampaign ponownie odegrała istotną rolę. Thomas Peterson przyznał, że doświadczenie zdobyte w czasie rozwijania projektu w Stanach Zjednoczonych przyniosło wymierne korzyści. Udoskonalono technologię, co pozwoliło dotrzeć do znacznie szerszego grona odbiorców. W tym przypadku nową aplikację pobrano blisko 150 tysięcy razy, a jej użytkownicy wygenerowali ruch na poziomie jednej czwartej miliona akcji. Kampania dotycząca Brexitu zakończyła się sukcesem. Można przypuszczać, że treści tworzone za pomocą wspomnianego narzędzia i sposób, w jaki jego użytkownicy przekazywali je dalej, mogły w znaczący sposób przyczynić się do sukcesu eurosceptyków (Tani 2016). Jako dowód wskazuje się fakt, że organizacja BeLeave przeznaczyła znaczne środki na usługi cyfrowe podmiotowi powiązanemu strukturalnie z Cambridge Analytica (wsparcie projektu Leave.EU i kampanii Nigela Farage'a i Arrona Banksa). Zdaniem Shahmira Sanniego, byłego wolontariusza, darowizny w wysokości 627 tysięcy funtów miały być przekazane na cel nielegalnie. Analityką danych w trakcie kampanii zajmowała się też kanadyjska firma AggregateIQ (AIQ), która wspierała kampanię Michaela Gove'a, Borisa Johnsona i Dominicka Cummingsa – wszyscy z projektu Vote Leave (Cadwaladr i in. 2018). Ich kampanie zakończyły się sukcesem, a część wymienionych polityków zajęła prominentne stanowiska w nowym rządzie.

Pojawiają się głosy, że analityka danych wykorzystana została także w Polsce podczas kampanii prezydenckiej Andrzeja Dudy w 2015 roku. Wzmiankował na ten temat brytyjski dziennik „The Guardian”, powołując się na wypowiedzi przedstawicieli Facebooka i Szkockiej Partii Narodowej. Ich zdaniem Facebook był integralną częścią zwycięstwa. Na zaangażowanie Cambridge Analytica może wskazywać podobny język i używane sformułowania (Hern 2018). Prezydent Duda oraz osoby odpowiedzialne za jego kampanię odrzucają te zarzuty i wskazują, że dysponowano zbyt niskim budżetem, by móc opłacić profesjonalną firmę zajmującą się prognozowaniem (Wirtualne Media 2018).

## Boty społeczne

Boty społeczne stają się coraz ważniejszym graczem w kampaniach wyborczych (Rogers 2013; Lee i Xu 2018; Keller i Klinger 2019). Są to programy komputerowe, które naśladują ludzi i potencjalnie manipulują ludźmi oraz ich zachowaniami w sieciach społecznościowych, wykorzystując potencjał *big data* (Wagner i in. 2016). Są coraz tańszymi narzędziami, dzięki którym określone treści, tematy lub aktorzy stają się bardziej popularni niż są w rzeczywistości. Nagłaśniają i katalizują zjawiska online, wywołując często efekt oburzenia oraz sztuczny ruch i dyskusję wokół określonych tematów. Ich działania są coraz trudniejsze do rozpoznania zarówno przez ludzi, jak i stworzone do tego celu algorytmy. Pracują jako bezosobowi agenci na rzecz wybranych podmiotów (instytucji komercyjnych czy aktorów politycznych). Wchodzą w interakcje z użytkownikami Sieci, mogąc wpływać na kształtowanie relacji społecznych. Ich zachowania są podobne do ludzkich. Przez niektórych autorów nazywane są „zautomatyzowanymi aktorami społecznymi” (Abokhodair i in. 2015: 2).

Boty społecznościowe przestały już być marginalnym zjawiskiem (Keller i Klinger 2019: 172). Jak wykazały badania Oxford Internet Institute, podczas kampanii wyborczych, które odbyły się w 2017 roku, wskaźnik „wysoce zautomatyzowanych” tweetów wzrósł z 5,7% do 7,4% (dane od lutego do września 2017 roku). W przypadku poszczególnych kampanii wynosił on od 5,2% do 16,5% (Neudert i in. 2017: 6). Badania realizowane poza Stanami Zjednoczonymi wykazały, że partie prawicowe i radykalne partie opozycyjne używały botów społecznych częściej niż inne ugrupowania (Schäfer i in. 2017). Wyniki badań przytaczanych przez Tobiasa Kellera i Ulrike Klinger dowodzą, że boty były obecne i miały wpływ na amerykańską kampanię prezydencką w 2016 roku. W przypadku profilu Donalda Trumpa na Twitterze stanowiły jedną czwartą wszystkich obserwujących go. Automatyczne programy ingerowały w debatę o Brexicie, a internetowa petycja o drugie referendum w sprawie Brexitu w czerwcu 2016 roku została „podpisana” przez 77 tysięcy botów. Odkryto też sieć blisko 14 tysięcy automatycznych profili na Twitterze wspierających kampanię Leave UE. Prowadziły one też kampanię dezinformacyjną #MacronLeaks (Keller i Klinger 2019: 175–176).

## Rola firm technologicznych

Rozwój botów i ich skuteczność powiązane są głównie z możliwościami, jakie media społecznościowe oferują deweloperom tworzącym tego typu aplikacje. Koncerny, takie jak Facebook, Twitter, Google i Microsoft, pozwalają twórcom programów na szybką dystrybucję ich produktów i precyzyjne dotarcie do szerokiego grona odbiorców. Za pomocą wewnętrznych narzędzi przekazują informacje o użytkownikach, ich cechach i aktywności w Sieci, umożliwiając tym samym analizę dużej ilości danych (*big data*). Daniel Kreiss i Shannon C. McGregor (2018) zwrócili uwagę na znaczącą rolę firm technologicznych w kampaniach wyborczych. Są to prawdopodobnie pierwsi badacze, którym udało się przeprowadzić nie tylko obserwacje terenowe, ale też wywiady z pracownikami tych instytucji i szefami kampanii prezydenckich w USA odpowiedzialnymi za działania w mediach cyfrowych. Autorzy dowiedli, że dyspozytorom firm technologicznych bardzo zależało na obecności ich firm podczas kampanii, a ich rola i oddziaływanie były większe, niż dotąd sądzono. Powoływali oni struktury organizacyjne z wyspecjalizowanym personelem, który współpracował z przedstawicielami aktorów politycznych. Koncerny umożliwiały strategom politycznym dotarcie do wyborców na podstawie wyspecjalizowanych kryteriów, takich jak demografia, rodzaje zachowań czy zainteresowania. Na podstawie tych danych przygotowano i kształtowano strategie, a także skuteczne przekazy kampanijne.

Przedstawiciele wszystkich czterech firm potwierdzili, że ich zaangażowanie w politykę wynikało z chęci uzyskania znaczących przychodów z oferowanych usług i produktów. Zależało im również na jak największej widoczności (nagłośnieniu). Implementacja ich poszczególnych usług rosła z kampanii na kampanię. Korzystali z nich zarówno demokraci, jak i republikanie. Firmy technologiczne oferowały narzędzia pozwalające z jednej strony wygrywać wybory, a z drugiej budować równoległe długofalowe relacje z kandydatami, wybranymi urzędnikami czy dziennikarzami. Każdy z podmiotów oferował wyspecjalizowane rozwiązanie. Microsoft podczas kampanii wyborczych zapewniał najlepszą infrastrukturę do zarządzania danymi i operacjami cyfrowymi. Facebook, Twitter i Google zorientowane były natomiast na dostarczanie rozwiązań niezbędnych do prowadzenia akcji reklamowych. Stworzyły też przestrzeń do budowania wizerunku kandydatów i stały się najważniejszymi kanałami komunikowania się z wyborcami (Kreiss i McGregor 2018).

Nowe technologie zaoferowane przez opisywane podmioty znacząco wpłynęły nie tylko na dystrybucję komunikatów kampanijnych, ale także fałszywych treści (*fake news*). Dezinformacja w trakcie kampanii prezydenckiej w 2016 roku w Stanach Zjednoczonych osiągnęła rozmiary nigdy wcześniej niespotykane. W ostatnim etapie kampanii fałszywe wiadomości osiągały w mediach społecznościowych większe zasięgi niż materiały pochodzące od najpoważniejszych podmiotów medialnych (Silverman 2016). Międzynarodowa opinia publiczna oskarżała o ten stan właśnie firmy technologiczne. Początkowo ich przedstawiciele stali na stanowisku, że nie są podmiotami medialnymi i nie mają wpływu na to, jakimi treściami dzielą się ich użytkownicy. Naciski międzynarodowej

opinii publicznej wymusiły jednak zmianę postawy. Firmy podjęły działania korygujące nadszarpnięty wizerunek. Rozpoczęto prace mające na celu wypracowanie mechanizmów weryfikowania i oznaczania fałszywych treści. Powołano programy stypendialne i podjęto współpracę z uczelniami. Działania te zostały wsparte przez wydawców i organizacje non profit. W podjętych krokach upatrywano zmianę jakościową (Łódzki 2017). Pomimo tych decyzji niespełna rok po kampanii dało się zaobserwować kolejne negatywne zjawiska. Na stronach weryfikujących fakty, PolitiFact i Snopes, pojawiły się reklamy przedsiębiorstwa informatycznego Google zawierające nieprawdę. Ich nadawcy, przykładowo, dystrybuowali treści o tym, że Melania Trump opuściła Biały Dom. Facebook i Messenger także zostały wykorzystane do promocji reklam politycznych poprzez niezaweryfikowane fałszywe konta, przynosząc spore zyski firmom technologicznym (Gdak 2017). Zdaniem Rafała Gdaka „[...] ze wspomnianą plagą nie tylko trudno walczyć, ale podatni na chorobę są również walczący z nią. Zjawisko staje się powoli trwałym elementem medialnego krajobrazu”. W 2018 roku dyspozytorzy firmy Google bezterminowo zawiesili system oznaczania fałszywych treści w wynikach wyszukiwań. Decyzja ta, zdaniem przedstawicieli firmy, podjęta została w wyniku oskarżeń o błędne wskazania. Nieprawidłowości nagłaśniali głównie przedstawiciele konserwatywnych serwisów informacyjnych (Funke 2018). W decyzji tej można dopatrywać się też przyczyn ekonomicznych.

## Podsumowanie

Przedstawione w artykule rozważania na temat roli *big data* w kampaniach wyborczych należy traktować jako zarys problematyki. Autor zasygnalizował obszary, które warto dalej eksplorować.

Aktorzy polityczni, przy wsparciu ekspertów zajmujących się analizą danych i monitoringiem Internetu, otrzymują szeroki wachlarz możliwości silniejszego niż dotąd oddziaływania na nastroje społeczne i kształtowanie poglądów. Gromadzą i przetwarzają coraz więcej szczegółowych danych o wyborcach, dostosowując strategię działań niemal w czasie rzeczywistym. Jak zauważył George Monbiot (2017), niepokój badaczy powinna budzić możliwość manipulowania lub zarządzania wyborcami bez ich świadomości. Zdaniem cytowanego autora „Informacje online już nadają się do manipulacji i politycznych nadużyć, a wiek dużych danych jeszcze się nie rozpoczął. W połączeniu z postępowaniem w dziedzinie lingwistyki kognitywnej i neuronauki dane te mogą stać się potężnym narzędziem służącym do zmiany decyzji wyborczych, które podejmujemy” (Monbiot 2017).

Wyodrębnianie potrzebnych informacji z dużych woluminów danych jest jednym z głównych wyzwań stojących nie tylko przed analitykami pracującymi dla aktorów politycznych, ale i badaczami. Piszą o tym Abdul Manan i Muqem Ahmed (2018: 366–369). Dokonując przeglądu i uporządkowania literatury na temat narzędzi, modeli i algorytmów wykorzystywanych do obserwacji

kampanii, dyskusji, przewidywań i analizy wyborów, przywołani autorzy dowodzą, że dzisiejsze badania nie przynoszą jeszcze jednoznacznych wyników i będą musiały być wielokrotnie weryfikowane. Ich zdaniem skomplikowane procedury statystyczne należy uzupełniać o analizę tekstów i obrazów.

Analiza mechanizmów *big data* w kampaniach wyborczych wymaga od badaczy coraz nowszych kompetencji. Powinni umieć podążać za zmieniającym się rynkiem usług cyfrowych i rozumieć siatkę specjalistycznych pojęć i mechanizmów związanych z analityką danych. Wiedza o procesach produkcji, dystrybucji, personalizacji i kontroli informacji umożliwia dokładniejszą analizę zachodzących zjawisk. Pozwala przyjrzeć się i krytycznie ocenić działalność aktorów politycznych i skuteczność tworzonych przez nich komunikatów. A Internet, jak podkreślają Andrew Chadwick i Jennifer Stromer-Galley (2016), daje aktorom politycznym coraz więcej możliwości dywersyfikacji kanałów dotarcia do wyborców.

Wyzwaniem stają się też badania doniesień medialnych w makroskali. Jednym z przykładów ilustrujących sposób ich realizacji jest projekt brytyjskich badaczy (Sudhahar i in. 2015). Autorzy, wykorzystując metody *big data*, zbadałi zawartość ponad 130 tysięcy informacji o wyborach w USA w 2012 roku. Analizowali treści dystrybuowane zarówno w środowisku online, jak i offline, wyodrębniając kluczowych aktorów politycznych i powiązane z nimi kwestie. W badaniu wykorzystano automatyczne językoznawstwo korpusowe połączone z analizą sieciową. Analiza współczesnego komunikowania politycznego wymaga też rozumienia specyfiki mediów społecznościowych oraz narzędzi ich badania (Procter i in. 2013), a także zachowań odbiorców, którzy coraz częściej rozposzechniają treści polityczne na masową skalę, bez wcześniejszej pogłębionej lektury (Bright 2016: 297).

Przedstawione w artykule przykłady zastosowania *big data* w kampaniach wyborczych miały głównie negatywny wydźwięk. Nie oznacza to, że analityka danych niesie ze sobą tylko pejoratywne konsekwencje. Technologie cyfrowe mogą też być potężną, pozytywną siłą. Dostarczając szybkich informacji, dają możliwość lepszego zarządzania. Rządy i partie polityczne mogłyby wykorzystywać wiedzę płynącą z danych, tak aby polityka stawała się bardziej przejrzysta i profesjonalna. Zastosowane do pozytywnych zmian społecznych, mogłyby wesprzeć wzrost i innowacje (Patel 2018). Aby taki warunek mógł być spełniony, technologia musi zostać upowszechniona, a jej wykorzystanie musi być transparentne. Nabywanie nowych kompetencji z obszaru technologii informacyjnych i komunikacyjnych (ang. *information and communications technology* – ICT) przez dziennikarzy czy zwykłych obywateli z pewnością przyspieszy te procesy.

## Bibliografia

- Abokhodair, Norah, Yoo, Daisy, i McDonald, David. 2015. Dissecting a Social Botnet. W: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing – CSCW '15*. New York: ACM Press, s. 839–851.
- Bright, Jonathan. 2016. The Social News Gap: How News Reading and News Sharing Diverge. *Journal of Communication*, 3 (66), s. 343–365.
- Chadwick, Andrew, i Stromer-Galley, Jennifer. 2016. Digital Media, Power, and Democracy in Parties and Election Campaigns: Party Decline or Party Renewal? *The International Journal of Press/Politics*, 3 (21), s. 283–293.
- Halevi, Gali. 2012. The Evolution of Big Data as a Research and Scientific Topic: Overview of the Literature. *Research Trends*, 30, s. 1–40.
- Keller, Tobias R., i Klinger, Ulrike. 2019. Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications. *Political Communication*, 1 (36), s. 171–189.
- Kreiss, Daniel, i McGregor, Shannon C. 2018. Technology Firms Shape Political Communication: The Work of Microsoft, Facebook, Twitter, and Google with Campaigns During the 2016 U.S. Presidential Cycle. *Political Communication*, 2 (25), s. 155–177.
- Lee, Jayeon, i Xu, Weiai. 2018. The More Attacks, the More Retweets: Trump's and Clinton's Agenda Setting on Twitter. *Public Relations Review*, 2 (44), s. 201–213.
- Łódzki, Bartłomiej. 2017. „Fake news” – dezinformacja w mediach internetowych i formy jej zwalczania w przestrzeni międzynarodowej. *Polityka i Społeczeństwo*, 4 (15), s. 19–30.
- Manan, Abdul, i Ahmed, Muqem. 2018. Election Prediction Using Big Data Analytics: A Survey. *International Journal of Engineering and Technology*, 5 (7), s. 366–369.
- Nickerson, David W., i Rogers, Tood. 2014. Political Campaigns and Big Data. *Journal of Economic Perspectives*, 2 (28), s. 51–74.
- Pawelczyk, Piotr, i Jakubowski, Jakub. 2017. Political Marketing in the Times of Big Data. *Przegląd Politologiczny Kwartalnik – Political Science Review*, 3, s. 34–44.
- Procter, Rob, Vis, Farida, i Voss, Alex. 2013. Reading the Riots on Twitter: Methodological Innovation for the Analysis of Big Data. *International Journal of Social Research Methodology*, 3 (16), s. 197–214.
- Rogers, Richard. 2013. *Digital Methods*. London: The MIT Press.
- Schäfer, Fabian, Evert, Stefan, i Heinrich, Philipp. 2017. Japan's 2014 General Election: Political Bots, Right-Wing Internet Activism, and Prime Minister Shinzō Abe's Hidden Nationalist Agenda. *Big Data*, 4 (5), s. 294–309.
- Sudhahar, Saatviga, Veltri, Giuseppe A., i Cristianini, Nello. 2015. Automated Analysis of the US Presidential Elections Using Big Data and Network Analysis. *Big Data and Society*, 1 (2), s. 1–28.

## Netografia

- Cadwalladr, Carole, Graham-Harrison, Emma, i Townsend, Mark. 2018. *Revealed: Brexit Insider Claims Vote Leave Team May Have Breached Spending Limits*. [Online]. The Guardian.com. Dostęp: <https://www.theguardian.com/politics/2018/mar/24/brexit-whistleblower-cambridge-analytica-beleave-vote-leave-shahmir-sanni> [10.11.2018].
- Datafloq. 2013. *Obama Changed the Political Campaign with Big Data*. [Online]. DataFloq.com. Dostęp: <https://datafloq.com/read/big-data-obama-campaign/516> [2.11.2018].
- Datafloq. 2018. *How Big Data Plays Important Role in Politics*. [Online]. DataFloq.com. Dostęp: <https://datafloq.com/read/how-big-data-plays-a-big-role-in-politics/4529> [15.10.2018].
- Duhigg, Charles. 13.10.2012. *Campaigns Mine Personal Lives to Get Out Vote*. [Online]. The New York Times. Dostęp: <https://www.nytimes.com/2012/10/14/us/politics/campaigns-mine-personal-lives-to-get-out-vote.html> [16.11.2018].
- Elliott, Timo. 2013. *7 Definitions of Big Data You Should Know About*. [Online]. Digital Business and Business Analytics. Timo Elliott's Blog. Dostęp: <https://timoelliott.com/blog/2013/07/7-definitions-of-big-data-you-should-know-about.html> [14.12.2017].
- Firican, George. 2017. *The 10 Vs of Big Data. Transforming Data with Intelligence*. [Online]. Tdwi.org. Dostęp: <https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx> [15.11.2018].

- Fried, Ina. 2016. *Hillary Clinton is Launching a Game-Style Mobile App for Campaign Volunteers*. [Online]. Recode. Dostęp: <https://www.recode.net/2016/7/24/12266026/hillary-clinton-campaign-mobile-app> [8.11.2018].
- Funke, Daniel. 2018. *Google Suspends Fact-Checking Feature over Quality Concerns*. [Online]. Poynter.org. Dostęp: <https://www.poynter.org/news/google-suspends-fact-checking-feature-over-quality-concerns> [14.11.2018].
- Gdak, Rafał. 2017. *Google wyświetlał fake newsy w witrynach walczących z nimi*. [Online]. Spidersweb.pl. Dostęp: <https://www.spidersweb.pl/2017/10/fake-news-google.html> [5.12.2017].
- Grassegger, Hannes, i Krogerus, Mikael. 2017. *The Data That Turned the World Upside Down*. [Online]. Motherboard. Dostęp: [https://motherboard.vice.com/en\\_us/article/mg9vvn/how-our-likes-helped-trump-win](https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win) [3.12.2018].
- Hern, Alex. 18.03.2018. *Breach Leaves Facebook Users Wondering: How Safe Is My Data*. [Online]. The Guardian.com. Dostęp: <https://www.theguardian.com/news/2018/mar/18/breach-leaves-facebook-users-left-wondering-how-safe-is-my-data> [10.11.2018].
- Hunter, Fergus. 23.01.2017. *Cambridge Analytica, the 'Psychographic' Data Firm Behind Donald Trump, Eyes Australian Move*. [Online]. The Sydney Morning Herald. Dostęp: <https://www.smh.com.au/politics/federal/cambridge-analytica-the-psychographic-data-firm-behind-donald-trump-eyes-australian-move-20161212-gt926e.html> [1.12.2018].
- Monbiot, George. 2017. *Big Data's Power Is Terrifying. That Could Be Good News for Democracy*. [Online]. The Guardian.com. Dostęp: [www.theguardian.com/commentisfree/2017/mar/06/big-data-cambridge-analytica-democracy](http://www.theguardian.com/commentisfree/2017/mar/06/big-data-cambridge-analytica-democracy) [15.12.2018].
- Neudert, Lisa-Maria, Kollanyi, Bence, i Howard, Philip. 2017. *Junk News and Bots During the German Parliamentary Election: What are German Voters Sharing over Twitter?* [Online]. The Computational Propaganda Research Project. Dostęp: [http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/09/ComProp\\_GermanElections\\_Sep2017v5.pdf](http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/09/ComProp_GermanElections_Sep2017v5.pdf) [15.12.2018].
- Patel, Pranay. 2018. *Role of Big Data in US Presidential Election*. [Online]. Researchgate.net. Dostęp: [https://www.researchgate.net/publication/325333423\\_ROLE\\_OF\\_BIG\\_DATA\\_IN\\_US\\_PRESIDENTIAL\\_ELECTION](https://www.researchgate.net/publication/325333423_ROLE_OF_BIG_DATA_IN_US_PRESIDENTIAL_ELECTION) [15.12.2018].
- Samsung. 2017. *Jak wiele danych produkujemy każdego dnia?* [Online]. Forbes.pl. Dostęp: <https://www.forbes.pl/technologie/jak-wiele-danych-produkujemy-kazdego-dnia/4mn4w69> [12.11.2018].
- Shen, George. 2013. *Big Data, Analytics and Elections*. [Online]. Analytics-magazine.org. Dostęp: <http://analytics-magazine.org/big-data-analytics-and-elections/> [17.11.2018].
- Silverman, Craig. 2016. *This Analysis Shows How Viral Fake Election News Stories Outperformed Real News on Facebook*. [Online]. BuzzFeedNews. Dostęp: [https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm\\_term=.elpwEEP2oo#.rrVvkk6oKK](https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm_term=.elpwEEP2oo#.rrVvkk6oKK) [18.04.2017].
- Tani, Maxwell. 2016. *Trump Campaign Is Using Same App 'Leave' Campaign Used During Brexit*. [Online]. Business Insider. Dostęp: <https://www.businessinsider.com/donald-trumps-phone-app-brexit-2016-11?IR=T> [2.11.2018].
- Wagner, Claudia, i in. 2016. *When Social Bots Attack: Modeling Susceptibility of Users in Online Social Networks*. [Online]. Ceur-ws.org. Dostęp: <http://ceur-ws.org/Vol-838> [2.12.2018].
- Wirtualne Media. 2018. *Andrzej Duda: nie było współpracy z Cambridge Analytica w mojej kampanii wyborczej, to fake news*. [Online]. Wirtualne Media.pl. Dostęp: <https://www.wirtualnemedia.pl/artykul/andrzej-duda-nie-bylo-wspolpracy-z-cambridge-analytica-w-mojej-kampanii-wyborczej-to-fake-news> [10.11.2018].

## Streszczenie

Technologie informacyjne i komunikacyjne (ICT) odgrywają coraz ważniejszą rolę we współczesnych społeczeństwach. Służą do przetwarzania, gromadzenia i przesyłania stale rosnącej ilości danych i informacji. Celem autora artykułu jest zaprezentowanie możliwości, jakie oferują mechanizmy *big data* i ich praktyczne wykorzystanie w kampaniach wyborczych. Wskazane zostały wyzwania, szanse i zagrożenia płynące

ze stosowania i przetwarzania „dużych danych”. Podjęta została też próba odpowiedzi na pytanie, czy analityka danych może oddziaływać na przebieg kampanii i wyniki wyborów.

## **The Role of Big Data in Political Campaigns**

### **S u m m a r y**

Information and communication technologies (ICTs) play an increasingly important role in modern societies. They are used to gather, process and share an ever-growing amount of data and information. The aim of the author of the article is to present the possibilities offered by Big Data mechanisms and their practical use in political campaigns. The challenges, opportunities, and dangers of using and processing “big data” are presented. An attempt to answer the question whether the data analysis may affect the campaigns and election results is taken.