

Szanse i zagrożenia rozwoju koncepcji *Big Data* na przykładzie sektora publicznego

JEL classification: F61, F63, L86, M31, O30

Słowa kluczowe: gospodarka cyfrowa, *Big Data*, dane osobowe, nowe technologie, Internet, sieci społecznościowe

Keywords: digital economy, Big Data, personal data, new technologies, Internet, social networking

Abstrakt: Współczesny postęp technologiczny implikuje dynamiczny rozwój gospodarki cyfrowej, w tym przede wszystkim przestrzeni internetowej. W konsekwencji wpływa to na pojawianie się nowych usług i narzędzi, wykorzystywanych do analizy ruchu w sieci, przede wszystkim koncepcji *Big Data*. Rozwój tego rynku stanowi szanse rozwoju nie tylko podmiotów gospodarczych, lecz także gospodarek narodowych. Wykorzystanie tego narzędzia przez sektor publiczny niesie z sobą również zagrożenia, co może przyczynić się do ograniczenia zaufania obywateli do państwa. Celem artykułu jest przedstawienie problematyki *Big Data* w obszarze funkcjonowania jednostek rządowych oraz rozważań na temat potencjalnych możliwości związanych z dalszym rozwojem tego narzędzia.

Opportunities and threats for the development of the Big Data concept based on the example of the public sector

Abstract: Modern technological progress entails a dynamic development of digital economy, including in particular the Web. As a result, this affects the evolution of new services and tools used to analyze network traffic, especially the concept of Big Data. The development of this market represents opportunities not only for business area but also the entire national economies. The use of this tool by the public sector also brings risks, which may result in decreasing confidence of citizens in the state. The purpose of the paper is to present the issues connected with Big Data in the functioning of government departments and to deliberate on the potential opportunities for the further development of this tool.

Wstęp

Współczesne koncepcje zarządzania przedsiębiorstwem w zdecydowanej większości opierają się na pozyskiwaniu, przetwarzaniu i wykorzystywaniu danych w celu budowania przewagi konkurencyjnej. Szczególnie umiejętność wykorzystania użytecznych informacji zdaje się tym czynnikiem, który może decydować o skuteczniejszym dotarciu do kluczowych klientów. Dynamiczny rozwój systemów informatycznych, jak również szeroko rozumianej przestrzeni internetowej, pozwala współczesnym przedsiębiorstwom wykorzystywać dużo dokładniejsze i efektywniejsze narzędzia do analizy danych, niż to miało miejsce przed erą gospodarki cyfrowej. Korzystają na tym nie tylko podmioty gospodarcze, lecz także sami klienci, którzy otrzymują produkt precyzyjniej zaspokajający ich konkretne potrzeby. Rynek *Big Data*, będący prężnie rozwijającą się gałęzią gospodarki cyfrowej, zdaje się najlepszym tego przykładem. Co ważne, możliwość wykorzystania pozyskanych za jego pomocą danych nie ogranicza się jedynie do komunikacji marketingowej. Coraz częściej znajduje on zastosowanie także w innych niż handel sferach życia człowieka, co może mieć pozytywny wpływ na dalszy rozwój całego społeczeństwa informacyjnego. Jednak wraz z dynamicznym rozwojem tej gałęzi gospodarki cyfrowej coraz częściej stawiane są pytania o potencjalne zagrożenia, jakie może on z sobą nieść. Kwestia ochrony prywatności konsumentów zdaje się tutaj sprawą szczególnie istotną i delikatną, co można zaobserwować zarówno w liczbie publikacji prasowych, jak i podejmowanych na ten temat dyskusji publicznych. Jednak rozważania nad tym problemem wiązać się także powinny z odpowiedzią na pytanie, do jakiego stopnia sami konsumenci potrafią zadbać o swoją prywatność w sieci internetowej.

Istota *Big Data* we współczesnym świecie

Wiedza na temat zachowań zakupowych konsumentów stanowi jeden z niezbędnych czynników powodzenia danego przedsięwzięcia gospodarczego. Przedsiębiorstwa od wielu lat dążą do gromadzenia wszelkich informacji na temat swoich klientów, aby w możliwie najskuteczniejszy sposób móc dotrzeć do nich ze swoją ofertą bądź też w odpowiedni sposób dostosować ją do ich oczekiwań¹. Informacje te potrafią istotnie wpłynąć na postawę i strategiczne podejście przedsiębiorstw, dlatego też niezbędne w ich działalności jest tworzenie i wykorzystywanie odpowiednich narzędzi pozwalających na pozyskanie takich danych.

Pierwotne narzędzia wykorzystywane do tego celu, między innymi karty lojalnościowe, badania opinii publicznej czy też ankiety bezpośrednie, pozwalały zdobyć podstawowe informacje na temat pewnych wzorców zachowań klientów,

¹ *Consumer privacy – protecting the consumer (from himself?)*, Euromonitor International 2007, www.portal.euromonitor.com/portal/default.aspx?8HpsQJ3P0tg%3d (dostęp: 18.05.2014).

choć wiązały się z istotnymi nakładami czasu oraz pieniędzy. Rozwój i popularyzacja Internetu przyczyniły się z kolei do opracowania dużo skuteczniejszych narzędzi, wymagających niższych zasobów zarówno finansowych, jak i czasowych, między innymi: fora internetowe, skrzynki pocztowe czy też tzw. pliki cookies². Jednak dopiero wraz z rozwojem i popularyzacją współczesnych portali społecznościowych oraz takich narzędzi, jak wyszukiwarki internetowe, nastąpiła istotna zmiana jakościowa w możliwościach pozyskiwania przez firmy wszelkiego rodzaju informacji. Z czasem przyczyniło się to do intensywnego wzrostu gromadzonych przez przedsiębiorstwa danych, samo zaś zjawisko określono mianem „koncepcji *Big Data*”.

Użycie tego pojęcia po raz pierwszy przypisuje się Rogerowi Magoulasowi z O'Reilly Media. W 2005 roku próbował on w ten sposób zdefiniować potężną ilość danych, którym tradycyjne metody zarządzania danymi nie potrafiły już sprostać ani ich przetworzyć z powodu zbytnej złożoności oraz rozmiarów samych informacji³. Stanowiło to konsekwencję intensywnego wzrostu liczby internautów na całym świecie (w 2012 roku liczba ta sięgnęła ponad 2,5 mld osób, co stanowiło około 1/3 światowej populacji⁴) oraz przeobrażeń, jakie dokonały się w świecie wirtualnym, a w wyniku których tradycyjny odbiorca treści stał się ich bezpośrednim nadawcą, co spotęgowało liczbę generowanych danych w przestrzeni internetowej. Prognozy sugerujące dalszy dynamiczny przyrost użytkowników Internetu (mający do 2020 roku sięgnąć niemal 44% całej populacji⁵) pozwalają stwierdzić, że ilość generowanych przez internautów danych będzie w przyszłości rosła w równie intensywnym tempie, co może stanowić wyzwanie dla firm starających się nimi zarządzać i przetwarzać je w celu osiągnięcia przewagi konkurencyjnej.

Współcześnie koncepcja *Big Data* opisywana jest za pomocą trzech najważniejszych, będących ściśle z sobą powiązanych czynników⁶. Pierwszy jest właśnie odpowiedzią na problemy związane z technologią pobierania, przechowywania i analizowania coraz większej ilości danych gromadzonych przez przedsiębiorstwa. Obejmuje on zarówno wiele innowacji technicznych, jak i nowe typy danych, dotychczas nieuwzględnianych, które umożliwiają dokładniejsze analizy

² Niewielkie dokumenty zapisywane na komputerze użytkownika, pozwalające na odczytywanie przez serwer podstawowych informacji, na przykład w jaki sposób użytkownik korzysta z portalu, jakich dokonał ustawień optymalizacyjnych *etc.*

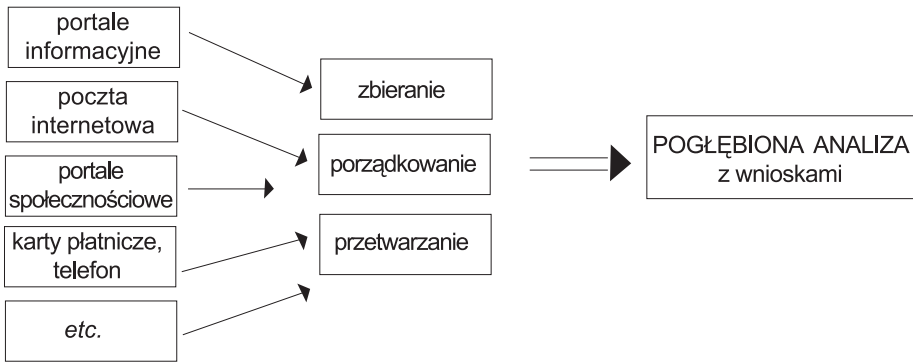
³ E.G. Ularu *et al.*, *Perspectives on Big Data and Big Data analytics*, „Database Systems Journal” III, 2012, nr 4, Bucharest 2012, s. 3.

⁴ Euromonitor International, www.portal.euromonitor.com/portal/default.aspx?itidoS5OFNI%3d (dostęp: 18.05.2014).

⁵ *Why Is the Online Domain Increasingly Displacing More Traditional Mediums?*, „Euromonitor International” 2013, www.portal.euromonitor.com/portal/default.aspx?tiaNpqH10x8%3d (dostęp: 18.05.2014).

⁶ D. Nunan, M.L. Di Domenico, *Market research and the ethics of big data*, „International Journal of Market Research” 2013, nr 55 (4), s. 506.

zachowań konsumentów. Uproszczony schemat zarządzania *Big Data* przedstawiono na rysunku 1⁷.



Rysunek 1. Schemat procesu zarządzania *Big Data*

Źródło: opracowanie własne na podstawie E.G. Ularu *et al.*, *op. cit.*, s. 6.

W porównaniu z tradycyjnymi metodami przetwarzania danych w celu wydobycia wiedzy z określonego zbioru danych *Big Data* wymaga głębszej analizy niż tylko dokonanej za pomocą wygenerowania prostego raportu⁸. To też stanowi o meritum *Big Data* — gromadzenie czy pobieranie danych nie jest faktem tak istotnym, jak wiedza będąca rezultatem umiejętności wyciągnięcia z tych danych odpowiednich wniosków. Do osiągnięcia tego celu niezbędna jest zatem infrastruktura komputerowa, będąca w stanie zgromadzić i przetworzyć taką ilość danych napływających z wielu obszarów aktywności konsumentów. Zgodnie z raportem Deloitte, *A wave of digital change: Trends in digital*, przeciętne współczesne amerykańskie przedsiębiorstwo zatrudniające ponad 1000 pracowników przetwarza rocznie około 200 terabajtów danych dotyczących klientów⁹. Dalszy wzrost liczby internautów wymagać zatem będzie jeszcze wydajniejszych i skuteczniejszych narzędzi do zarządzania taką ilością informacji.

Drugi czynnik określający koncepcję *Big Data* skupia się na wartości handlowej, która może zostać dodana do organizacji dzięki odpowiedniemu wykorzystaniu posiadanych przez nią informacji. Związana jest ona zarówno z rozwojem samej technologii, jak i z większą skłonnością konsumentów do udostępniania danych osobowych w ramach oferowanych przez firmy usług internetowych.

⁷ Należy zauważyć, że istnieją różne metody i algorytmy przetwarzania danych w ramach *Big Data*, niemniej w niniejszej publikacji przedstawiono najprostszą z nich.

⁸ H.F. Qin, Z.H. Li, *Research on the Method of Big Data Analysis*, „Information Technology Journal” 12, 2012, nr 10, s. 1974.

⁹ R. Antczak *et al.*, *A wave of digital change: Trends in digital*, raport Deloitte 2013, s. 6, www.deloitte.com/assets/Dcom-Poland/Local%20Assets/Documents/Raporty,%20badania,%20rankingi/pl_Digital_Trends_2013_Final2.pdf (dostęp: 18.05.2014).

W rezultacie możliwości wykorzystania *Big Data* w komunikacji marketingowej zdają się praktycznie nieograniczone. Wiedza na temat danego konsumenta, uzyskana na podstawie analizy jego zachowania czy też poszukiwanych przez niego informacji w sieci, pozwala przygotować ofertę komercyjną idealnie do niego dopasowaną. Z kolei wiedza na temat sposobu zarządzania domowym budżetem określonego klienta pozwala instytucjom finansowym zaproponować najlepsze rozwiązania dla jego oszczędności. Wykorzystanie *Big Data* w sferze handlu umożliwia także zastosowanie tzw. systemu spersonalizowanych cen, polegającego na dostosowywaniu w danym obiekcie handlowym cen produktów do konkretnego konsumenta, na podstawie historii jego koszyka zakupów¹⁰.

Według raportu The Boston Consulting Group, *The value of our digital identity*, wartość ekonomiczna tego typu danych przetwarzanych przez przedsiębiorstwa z możliwością ich komercyjnego wykorzystania do 2020 roku w samej tylko Europie sięgnie 330 mld euro¹¹. Niemniej jednak, co podkreślają autorzy raportu, korzyści dla konsumentów będą stanowiły wartość znacznie wyższą niż dla przedsiębiorstw, wynoszącą około 670 mld euro. To także jest jednym z najważniejszych argumentów przytaczanych przez zwolenników wykorzystania *Big Data* w działalności marketingowej przedsiębiorstw, starających się podkreślić, że dobrze dopasowana oferta handlowa może przyczynić się między innymi do oszczędności czasu każdej osoby poszukującej odpowiedniej dla siebie oferty¹².

Trzeci czynnik *Big Data* sprowadza się do podkreślenia istoty problemu ochrony prywatności konsumentów, a zatem konieczności ustanowienia odpowiednich regulacji prawnych oraz wytycznych dotyczących etyki komercyjnego wykorzystania pozyskiwanych danych. Czynniki ten staje się coraz ważniejszym elementem koncepcji *Big Data*, co związane jest przede wszystkim ze wzrostem liczby osób świadomych ewentualnych zagrożeń związanych z udostępnianiem swoich danych, dla których sprawa ochrony własnej prywatności ma coraz większe znaczenie. Niektóre kwestie, mające na celu zwiększenie ochrony konsumentów przed nieświadomie przekazywanymi przez nich informacjami, doczekały się już prawnego uregulowania¹³, niemniej intensywny rozwój tego narzędzia

¹⁰ O. Kharif, *Cena szyta na miarę każdego klienta*, „Bloomberg Businessweek Polska” 2013, nr 51–52, s. 19.

¹¹ The Boston Consulting Group, *The value of our digital identity*, 2012, s. 3, www.libertyglobal.com/PDF/public-policy/The-Value-of-Our-Digital-Identity.pdf (dostęp: 18.05.2014).

¹² Przykładowo, wyszukiwarka internetowa wykorzystująca algorytm wyszukiwania behawioralnego przewiduje hasło, jakie dany użytkownik chce znaleźć, opierając się na historii jego wcześniejszych zapytań oraz lokalizując jego obecne miejsce przebywania.

¹³ Przykładowo, zgodnie w wytycznymi Parlamentu Europejskiego Polska znowelizowała ustawę o prawie telekomunikacyjnym, uwzględniając w art. 173 obowiązek uzyskania przez serwis internetowy zgody od użytkownika na wykorzystanie jakichkolwiek narzędzi do gromadzenia danych. Zob. A. Krasuski, *Dane osobowe w obrocie tradycyjnym i elektronicznym. Praktyczne problemy*, Warszawa 2012, s. 86–95.

powoduje konieczność stałego dostosowywania przepisów uwzględniających nowe możliwości wykorzystania danych przeciwko konsumentom. Z powodu znacznego wzrostu zainteresowania tym zagadnieniem warto poszerzyć analizę tego czynnika o kwestię odpowiedniego dbania o swoją prywatność przez użytkowników sieci internetowej.

Problematyka ochrony prywatności w dobie *Big Data*

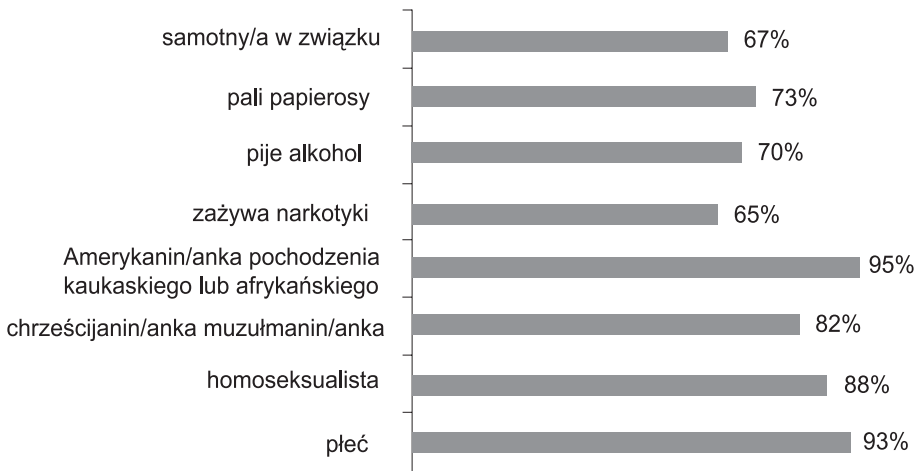
Wraz z rosnącą popularyzacją sieci społecznościowych oraz rozpropagowaniem informacji na temat gromadzenia przez przedsiębiorstwa wszelkiego rodzaju danych o swoich użytkownikach pojawił się dylemat dotyczący odpowiedniej ochrony danych osobowych. O ile poziom świadomości użytkowników na temat ich praw do zachowania prywatności stopniowo wzrasta, o tyle zasadniczy problem zdaje się pozostawać nierozwiązany. W literaturze można znaleźć kilka głównych wyjaśnień tej sprzeczności. Najważniejsze z nich związane jest z praktyką, do której użytkownicy Internetu zostali przyzwyczajeni na przestrzeni wielu lat. Większość podstawowych usług online, jak na przykład internetowa skrzynka pocztowa czy też portale informacyjne, oferowane były (i najczęściej ciągle są) bezpłatnie. Brak konieczności płacenia za świadczone usługi ma jednak swoje następstwa. Pierwszym z nich jest przymusowy kontakt z komunikatami reklamowymi (widocznymi na stronie portalu czy też w postaci tzw. mailingu w skrzynce pocztowej), drugi zaś zasadniczo pozbawia użytkownika prawa do własności w ramach korzystania z obszaru danej usługi. Zgodnie z takim podejściem należy mieć na uwadze, że każda informacja czy też plik z chwilą wprowadzenia do sieci przestaje należeć do jego właściciela¹⁴. A zatem każdy opublikowany na portalu komentarz czy też wygenerowana wiadomość e-mail w momencie utworzenia może stać się informacją pozyskaną, a następnie przetworzoną w ramach narzędzi *Big Data*. Inną kwestią pozostaje, że pewne dane w sposób świadomy i dobrowolny bywają przekazywane przedsiębiorstwom w celu otrzymania określonych korzyści materialnych (według badań 57% internautów udostępniłoby swój adres e-mail w zamian za otrzymanie zniżki na zakupy czy też powiadomienie o wyprzedży¹⁵).

O ile problem ten można wytłumaczyć ponoszeniem pewnego rodzaju kosztów za korzystanie z darmowych usług internetowych, o tyle niebezpieczne w opinii badaczy przedmiotu staje się zarządzanie przez internautów swoją prywatnością na portalach społecznościowych. Pomijając fakt częstego niezapoznawania się z regulaminem korzystania z danego portalu (którego najważniejsze założenia są powszechnie znane), można zaobserwować w tego rodzaju sieciach skłonność do wręcz świadomego pozbywania się prywatności.

¹⁴ E. Mistewicz, *Czas monetyzacji*, „Nowe Media” 3, Gdynia 2013, s. 123.

¹⁵ Cisco Systems Inc., *Gen Y: New Dawn for work, play, identity. Report*, www.cisco.com/en/US/solutions/ns341/ns525/ns537/ns705/ns1120/2012-CCWTR-Chapter1-Global-Results.pdf.

W rezultacie dane publikowane w sieciach społecznościowych mogą stać się istotnym źródłem informacji o określonym użytkowniku. Potwierdzeniem tej tezy mogą być wyniki badań przeprowadzonych na grupie ponad 60 tysięcy użytkowników najpopularniejszego współcześnie portalu społecznościowego Facebook¹⁶. Specjalnie stworzony algorytm komputerowy, analizując tylko publicznie dostępne, polubione na portalu strony (tzw. kliknięcia „Lubię to”), z dużą dokładnością potrafił stworzyć portret psychologiczny danej osoby. Najważniejsze cechy przedstawiono poniżej.



Rysunek 2. Dokładność predykcji sklasyfikowanych cech (wybrane wartości)

Źródło: M. Kosinski *et al.*, *op. cit.*

Jak można zauważyć, model komputerowy niemal bezbłędnie potrafił określić płeć czy przynależność rasową, jak również dane powszechnie uznawane za wrażliwe, takie jak orientacja seksualna czy też wyznawana religia. Uzyskane wyniki są tym istotniejsze, że opracowany algorytm nie opierał się na informacjach relatywnie prostych (jak na przykład polubienie strony *Barack Obama* czy *Being Gay*), lecz na mniej oczywistych zachowaniach i preferencjach (na przykład przeczytanie jakiejś książki, preferowanie konkretnej kawiarni *etc.*). Dlatego też należy podkreślić, że nawet brak otwartej manifestacji swoich poglądów w sieci społecznościowej nie gwarantuje pełnego zachowania prywatności.

Należy w tym miejscu nadmienić, że nie tylko portale społecznościowe mogą stanowić zagrożenie utraty tożsamości. W opinii niektórych badaczy większy problem

¹⁶ M. Kosinski *et al.*, *Private traits and attributes are predictable from digital records of human behavior*, „PNAS” vol. 110, 2013, nr 15, www.pnas.org/content/early/2013/03/06/1218772110.full.pdf+html?sid=1b67721e-32d0-4efa-a486-67c018015f53.

z ochroną danych osobowych związany jest z usługami firmy Google, przede wszystkim jej wyszukiwarką internetową. Spółka ta, stosująca strategię udostępniania internautom wszelkiego rodzaju bezpłatnych usług, staje się powoli konglomeratem agregującym ogromną ilość danych¹⁷. W związku z tym wszelkie informacje, jakie pojawiają się w przestrzeni internetowej na temat danego użytkownika, choćby były zamieszczone bez jego wiedzy i zgody, mogą być wygenerowane w wynikach wyszukiwania tej wyszukiwarki. O ile na część publikowanych informacji można nie mieć wpływu, o tyle należy pamiętać o ostrożności w przypadku wszelkich innych danych pozostawianych w sieci internetowej. Eksperyment przeprowadzony przez autora niniejszej publikacji dowiódł, że wystarczy niespełna pięć minut, aby korzystając głównie z narzędzi Google, stworzyć profil psychologiczny osoby pierwotnie anonimowej. Posługując się tylko danymi zamieszczonymi na ogólnodostępnych stronach internetowych, można uzyskać między innymi informacje dotyczące: adresu e-mail, sytuacji zawodowej, preferencji co do napojów alkoholowych, przybliżonej sytuacji materialnej (subiektywnie, na podstawie zdjęć z wakacji) czy też na temat najbliższych planów (rozpoczęcie kursu języka obcego)¹⁸.

Możliwość uzyskania tych informacji uzależniona jest przede wszystkim od danej osoby. Im większa aktywność danego użytkownika w sieci oraz skłonność do zamieszczania informacji *stricte* prywatnych (głównie adresu e-mail oraz numeru telefonu, na przykład w ogłoszeniach internetowych), tym większe prawdopodobieństwo otrzymania relatywnie dużej ilości danych. Taka sytuacja może rodzić przede wszystkim obawy o możliwość wykorzystania tego typu informacji przez osoby niepowołane. Komunikaty marketingowe zdają się w tym przypadku najmniejszym ryzykiem. Stąd też, korzystając z możliwości współczesnych technologii, należy w opinii autora stale kontrolować zakres i rodzaj udostępnianych przez siebie informacji.

Wykorzystanie *Big Data* w sektorze publicznym

Możliwości wykorzystania *Big Data* są znacznie szersze niż tylko *stricte* komercyjne. Coraz częściej podejmowanym tematem jest wykorzystanie tego narzędzia przez sektor publiczny. Także w tym wypadku należy wykazywać potencjalne szanse i zagrożenia z tym związane. Sprawa związana z udostępnieniem poufnych informacji przez amerykańskiego obywatela Edwarda Snowdena z 2013 roku przyczyniła się do stworzenia negatywnego wizerunku instytucji publicznej jako podmiotu zbierającego i przetwarzającego wszelkiego rodzaju informacje o swoich obywatelach¹⁹. Kontrowersyjność tej sprawy nie powinna jednakże przysłańać

¹⁷ J.P. Oury, *Kto się boi Google?*, „Nowe Media” 4, red. E. Mistewicz, Gdynia 2013, s. 137–139.

¹⁸ Informacje zgromadzone przy wykorzystaniu Google.pl, Facebook.pl, Allegro.pl, Gumtree.pl.

¹⁹ M. Jones, P. Silberzahn, *Snowden and the Challenge of Intelligence: The Practical Case against the NSA's Big Data*, www.forbes.com/sites/silberzahnjones/2013/07/11/snowden-and-the-challenge-of-intelligence-the-practical-case-against-nsa-big-data/ (dostęp: 18.05.2014).

pozytywnych aspektów możliwości wykorzystania przez sektor publiczny koncepcji *Big Data*.

Należy jednak nadmienić, że inne cele przyświecać będą wykorzystywaniu *Big Data* w działalności przedsiębiorstw komercyjnych, inne zaś w przypadku podmiotów publicznych. W biznesie głównym założeniem jest osiągnięcie przez dany podmiot gospodarczy jak największych zysków, przez dostarczanie produktów i usług w możliwie największym stopniu zaspokajających potrzeby klientów, a przez to tworzenie i utrzymywanie przewagi konkurencyjnej. Nadrzędnym celem jednostki publicznej, przede wszystkim rządu, jest zaś dążenie do osiągnięcia zrównoważonego rozwoju, zabezpieczenia podstawowych praw obywateli czy też promowania powszechnego dobrobytu i wzrostu gospodarczego.

Przykładem możliwości wykorzystania koncepcji *Big Data* przez instytucje rządowe jest analiza zapytań podejmowanych przez internautów w wyszukiwarkach internetowych. Dzięki takiemu narzędziu obywatele są w stanie zasygnalizować instytucjom rządowym miejsce wybuchu ognisk chorób zakaźnych, na przykład przez generowanie zapytań o sposób leczenia podobnych objawów choroby²⁰. Podmioty medyczne otrzymują szansę lepszego dostosowania medykamentów do walki z daną przypadłością, na bieżąco analizując postępy w leczeniu chorych pacjentów. Jednostki zarządzające rynkiem pracy otrzymują narzędzie do zrekrutowania idealnego kandydata do danej oferty pracy, z pominięciem drogich i pracochłonnych testów psychologicznych. Dostęp do wielu danych generowanych przez obywateli każdego dnia pozwala z kolei tworzyć dokładniejsze modele makroekonomiczne, w lepszym stopniu przewidujące rozwój gospodarki, co przyczynić się może do napływu kapitału zagranicznego. Analiza różnego rodzaju danych pozwala odpowiednim jednostkom publicznym skuteczniej pełnić swoje funkcje czy też wyprzedzać pewnego rodzaju zjawiska, na przykład o charakterze kryminogennym. Ma to szczególnie istotne znaczenie w budowaniu przez obywateli zaufania do organów państwa stojących za ich bezpieczeństwem.

Należy jednak zauważyć, że wykorzystanie narzędzi *Big Data* przez sektor publiczny może stanowić problem związany z zagadnieniami integracji danych z wielu źródeł (np. państw, instytucji, agencji i departamentów), występujących w różnych formatach, nie wspominając o kosztach ich pozyskania. O ile podmioty gospodarcze agregują pozyskiwane informacje w ramach swojego własnego systemu informatycznego, o tyle gromadzenie danych przez sektor publiczny związane już jest z operowaniem różnymi systemami, niekoniecznie z sobą kompatybilnymi. Brak ustandaryzowanego systemu sprawia, że udostępnianie danych między różnymi departamentami i agencjami rządowymi staje się niemożliwe lub trudne do osiągnięcia, ograniczając tym samym potencjalne możliwości płynące z wykorzystania narzędzi *Big Data*. Skuteczność prowadzonych analiz zależeć także będzie

²⁰ M. Stańczyk, *Dokąd zmierza komputeryzacja*, „Wprost” 2010, nr 13, s. 46.

od umiejętności określonych jednostek publicznych, jeśli chodzi o przekształcenie posiadanych informacji w odpowiednią wiedzę i skuteczne jej wykorzystanie.

Co równie istotne, udostępnianie danych i informacji między krajami staje się szczególnym wyzwaniem, czego przykładem mogą być próby zapobiegania współczesnym atakom terrorystycznym. Rządy krajowe powinny być więc przygotowane i chętne do udostępniania danych w budowaniu systemów przeciwdziałania przestępczości. To jednak może rodzić obawy o nadmierną ingerencję państwa w prywatne życie obywateli. O ile kwestia nabywania i wykorzystywania przez podmioty handlowe wiedzy o swoich klientach jest sprawą co do swej celowości raczej oczywistą, o tyle działania sektora publicznego w tym samym zakresie stanowić mogą źródło społecznych niepokojów. Granica między troską o bezpieczeństwo obywateli a poszanowaniem prawa do ich prywatności jest bardzo cienka, co może stworzyć pole do politycznych nadużyć. Należy bowiem zauważyć, że skala i zakres posiadanych przez jednostki rządowe informacji o obywatelach jest nieporównywalna z tymi agregowanymi przez sektor komercyjny. O ile istnieje pewna kontrola nad udostępnianiem przez obywateli informacji podmiotom gospodarczym, o tyle w przypadku sektora publicznego dane gromadzone są poza możliwością ich ingerencji. Spornym zagadnieniem pozostaje także kwestia udostępniania organom rządowym informacji posiadanych przez podmioty gospodarcze o swoich klientach, co dotyczy głównie sektora bankowego oraz telekomunikacyjnego. Także w tym wypadku mogą istnieć obawy przed nadmierną ingerencją instytucji państwowych w prywatne życie obywateli.

Mając to na uwadze, należy zauważyć, że mimo pozytywnych aspektów możliwości wykorzystania narzędzi *Big Data* przez sektor publiczny istnieją realne zagrożenia dotyczące nadużywania władzy przeciwko obywatelom. Niezbędne wydaje się ustanowienie odpowiednich regulacji prawnych, ściśle określających zakres i możliwości gromadzenia i przetwarzania informacji przez jednostki rządowe. Transparentność ich funkcjonowania przyczynić się bowiem może do wykorzystania potencjalnych możliwości *Big Data* przy jednoczesnym zapewnieniu ochrony prywatnego życia obywateli.

Wnioski

Rynek *Big Data* stanowi istotną gałąź w rozwoju współczesnej gospodarki cyfrowej. Dynamika przyrostu osób korzystających z sieci internetowej oraz spektrum możliwości wykorzystania narzędzi online sprawiają, że nieustannie zmienia się obraz przestrzeni internetowej, implikując między innymi coraz większą ilość danych wytwarzanych, a w związku z tym agregowanych i przetwarzanych, przez przedsiębiorstwa każdego dnia. Także jednostki sektora publicznego dostrzegają możliwość wykorzystania *Big Data*, co w konsekwencji rodzić może szansę na zwiększenie wzrostu gospodarczego. Problematyka koncepcji *Big*

Data w funkcjonowaniu organów rządowych jest znacznie szersza niż w przypadku podmiotów komercyjnych. Wynika to zarówno z ilości i złożoności przetwarzanych danych, jak i konieczności operowania nimi w ujednoczonym systemie informatycznym. Zagrożeniem dalszego rozwoju tego narzędzia mogą stać się obawy obywateli przed nadmierną ingerencją władzy w ich życie prywatne, stąd niezbędne zdaje się stworzenie odpowiednich regulacji prawnych określających możliwości pozyskiwania i przetwarzania danych przez konkretne jednostki sektora publicznego.

Bibliografia

- Antczak R. *et al.*, *A wave of digital change: Trends in digital*, Deloitte 2013, www.deloitte.com.
- Cisco Systems Inc., *Gen Y: New Dawn for work, play, identity. Report*, www.cisco.com.
- Consumer privacy — protecting the consumer (from himself?)*, Euromonitor International 2007, www.portal.euromonitor.com.
- Euromonitor International, www.portal.euromonitor.com.
- Jones M., Silberzahn P., *Snowden and the challenge of intelligence: The practical case against the NSA's Big Data*, www.forbes.com.
- Kharif O., *Cena szyta na miarę każdego klienta*, „Bloomberg Businessweek Polska” 2013, nr 51–52.
- Kosinski M. *et al.*, *Private traits and attributes are predictable from digital records of human behavior*, „PNAS” 110, 2013, nr 15, www.pnas.org.
- Krasuski A., *Dane osobowe w obrocie tradycyjnym i elektronicznym. Praktyczne problemy*, Warszawa 2012.
- Mistewicz E., *Czas monetyzacji*, „Nowe Media” 3, Gdynia 2013.
- Nunan D., Di Domenico M.L., *Market research and the ethics of big data*, „International Journal of Market Research” 2013, nr 55 (4).
- Oury J.P., *Kto się boi Google?*, „Nowe Media” 4, red. E. Mistewicz, Gdynia 2013.
- Qin H.F., Li Z.H., *Research on the method of Big Data analysis*, „Information Technology Journal” 12, 2012, nr 10.
- Stańczyk M., *Dokąd zmierza komputeryzacja*, „Wprost” 2010, nr 13.
- The Boston Consulting Group, *The value of our digital identity*, 2012.
- Ularu E.G. *et al.*, *Perspectives on Big Data and Big Data analytics*, „Database Systems Journal” III, 2012, nr 4, red. I. Lungu, Bucharest 2012.
- Why Is the Online Domain Increasingly Displacing More Traditional Mediums?*, Euromonitor International 2013, www.portal.euromonitor.com.

Opportunities and threats for the development of the Big Data concept based on the example of the public sector

Summary

Big Data market is an important branch in the development of the modern digital economy. The growing rate of people using the Internet and various online tools makes the image of the Web change constantly, entailing among other things, the increasing amount of data generated, and therefore aggregated and processed by the companies each day. Public sector entities recognize the possibility of using Big Data as well, which in turn can contribute to an increase in economic growth. The

problem of the concept of Big Data in the functioning of the government is much wider than in the case of commercial entities. This is due both to the amount and complexity of the data processed, and the need to use them in a unified system. A threat to further development of this tool may be citizens' fears of excessive interference of the authorities in their private life, hence it seems necessary to create appropriate legal regulations defining the possibilities of obtaining and processing data by certain public sector entities.