Jędrzej Wieczorkowski[1]

Big data a prywatność. Naruszenie prywatności w świecie wirtualnym – wyniki badań

1. Wstęp

Nowe możliwości technologiczne w zakresie informatyki i telekomunikacji znacząco wpływają na modele biznesowe stosowane przez przedsiębiorstwa, a także na metody wykorzystywane przez administrację publiczną. W szczególności dotyczy to rozwoju zastosowań przetwarzania danych masowych, określanych zazwyczaj jako *big data*. Jednym z podstawowych zastosowań takich metod jest gromadzenie i przetwarzanie danych osobowych. Rodzi to znaczące konsekwencje społeczne, jeśli chodzi o zmianę spojrzenia na prywatność.

Celem artykułu jest przedstawienie problemu prywatności oraz zbadanie subiektywnego poczucia jej naruszenia w kontekście stosowania metod masowego przetwarzania danych osobowych. W kolejnych podpunktach przedstawiono społeczny aspekt funkcjonowania *big data*, w szczególności zagadnienie prywatności, następnie opisano przeprowadzone przez autora badanie dotyczące postrzegania jej naruszenia związanego z metodami *big data* oraz przedstawiono i zinterpretowano jego wyniki.

2. Prywatność jako aspekt społeczny *big data*

Pojęcie *big data*, choć w praktyce dopiero się kształtuje, jest w ostatnich latach bardzo chętnie stosowane. W literaturze można odnaleźć artykuły, w których dokonano przeglądów jego definicji[2]. Najbardziej typowe próby zdefiniowania

[1] Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych.

[2] D. Boyd, K. Crawford, *Critical questions for big data in Information*, „Communication & Society" 2012, vol. 15, issue 5, s. 662–679; M. Tabakow, J. Korczak, B. Franczyk, *Big data – definicje, wyzwania i technologie informatyczne*, „Business Informatics", z. 1(31), Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2014, s. 138–153.

odnoszą się do zbiorów danych, których rozmiary przekraczają możliwości typowych narzędzi bazodanowych w zakresie gromadzenia, przechowywania, zarządzania i analizowania tych danych[3]. Pojęcie to wiąże się więc z coraz szerszymi możliwościami przetwarzania danych masowych, lecz w praktyce łączy w sobie kilka dość odrębnych aspektów. W celu zrozumienia zjawiska *big data* należy, zdaniem autora, połączyć kilka punktów widzenia i odpowiadające im trzy podstawowe aspekty *big data*:

- nowe możliwości w zakresie dostępnych technologii i metod analizy danych – **aspekt techniczny**;
- nowe zastosowania tych możliwości w różnych obszarach, takich jak biznes, administracja publiczna, nauka – **aspekt biznesowy**;
- konsekwencje powyższych zastosowań (w szczególności o obszarze społecznym) związane z masowym przetwarzaniem danych osobowych – **aspekt społeczny**.

W kontekście niniejszego artykułu najistotniejszy jest aspekt społeczny. Wiąże się ściśle z **problemem prywatności**. Dzisiejsza technologia informatyczna i technologia telekomunikacyjna całkowicie zmieniają dotychczasowe rozumienie pojęcia prywatności. Dotyczy to zarówno kontekstu społecznego, czyli kontaktów ze znajomymi oraz rodziną, jak i kontekstu instytucjonalnego, czyli relacji z różnorodnymi instytucjami oraz organizacjami o charakterze publicznym i biznesowym. Współczesna rzeczywistość w zakresie podejścia do prywatności jest niekiedy określana jako społeczeństwo postprywatne, w którym sfera prywatności jest stopniowo ograniczana. W kontekście instytucjonalnym można zaobserwować rozproszenie systemu kontroli podmiotów mających dostęp do danych związanych z prywatnością. Są to, poza instytucjami państwa, podmioty związane z konsumpcją, rynkiem i ogólnie działalnością gospodarczą[4].

W obszarze kontroli publicznej szczegółowe dane osobowe są gromadzone i przetwarzane przez różne instytucje w celu zapewnienia szeroko rozumianego bezpieczeństwa państwa. Analizuje się dane ogólnodostępne (np. treści zamieszczane w Internecie), dane obrazowe (monitoring TV, zdjęcia lotnicze i satelitarne), a także na podstawie specjalnych praw pozyskuje się całkowicie prywatne dane, np. policja, służby specjalne, prokuratura pozyskują dane z telefonii (lokalizacyjne i bilingi), prywatne treści z Internetu: korespondencję

[3] *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute, 2011.

[4] R. Dopierała, *Prywatność w perspektywie zmiany społecznej*, Zakład Wydawniczy Nomos, Kraków 2013.

e-mailową, zawartość plików przechowywanych w chmurze itp. Można także przypuszczać, że część danych jest uzyskiwana przez służby pozaprawnie.

W obszarze biznesowym typowym zastosowaniem przetwarzania typu *big data* jest opracowywanie zindywidualizowanego przekazu reklamowego. W tym celu śledzi się i analizuje aktywność użytkowników w Internecie, a dzięki temu oglądane przez nich reklamy można oprzeć na danych profilowanych w czasie rzeczywistym. Podstawowy model biznesowy serwisów społecznościowych zakłada udostępnienie społeczności platformy użytkowej w zamian za dostęp do spersonalizowanych strumieni informacji współtworzonych i współdzielonych przez społeczność[5]. Wysoką wartość biznesową mają także uzyskiwane w czasie rzeczywistym osobowe dane lokalizacyjne pochodzące np. z telefonii komórkowej.

Poza reklamą do biznesowych dziedzin zastosowań *big data* można zaliczyć m.in. finanse, medycynę i energetykę. Zastosowania te nie opierają się wyłącznie na danych zagregowanych i w związku z tym, że operują danymi elementarnymi, wiążą się z przetwarzaniem danych dotyczących bezpośrednio poszczególnych osób. Ewidencja i przetwarzanie operacji finansowych dokonywanych kartami płatniczymi, np. na potrzeby wykrywania w czasie rzeczywistym potencjalnych nadużyć, wiążą się z przetwarzaniem informacji o zakupach i lokalizacji poszczególnych osób. Analiza komputerowa szczegółowych danych medycznych opiera się na przetwarzaniu wrażliwych danych dotyczących zdrowia poszczególnych pacjentów. Kontrola sieci energetycznej i poboru prądu w czasie rzeczywistym w celu zarządzania całą siecią skutkuje przetwarzaniem danych dotyczących zachowań poszczególnych osób (np. uruchamiania odbiorników energii). Obecnie ogromna wiedza na temat obywateli znajduje się w rękach komercyjnych organizacji, których celem jest (w przeciwieństwie do instytucji państwa) przynoszenie zysku akcjonariuszom.

Najistotniejszym elementem aspektu społecznego *big data* jest zapewnienie odpowiedniego poziomu prywatności. Powstaje pytanie, jaki jest ten odpowiedni poziom, a także jak rozumieć obecnie samą prywatność. Społeczne rozumienie pojęcia prywatności w ostatnich latach szybko się zmieniało i będzie prawdopodobnie nadal ewoluować. Zapewne nie ma już powrotu do rozumienia prywatności np. sprzed 20 lub 10 lat. Na te zmiany ma ogromny wpływ w szczególności technologia informatyczna. Dane osobowe przetwarzano od dłuższego czasu, ale dopiero ostatnie lata i metody *big data* przyniosły tak szerokie możliwości

[5] K. Polańska, A. Wassilew, *Analizy big data w serwisach społecznościowych*, „Nierówności Społeczne a Wzrost Gospodarczy" 2015, nr 4, cz. 2, s. 117–128.

operowania indywidualnymi danymi osobowymi, a nie tylko danymi zagregowanymi. Aktualnie dostęp do indywidualnych danych osobowych możliwy jest często w czasie rzeczywistym lub do niego zbliżonym, a nie z opóźnieniem wymaganym do wstępnego przetworzenia danych.

Sytuacja ta wymusza zapewnienie niezbędnego minimum ochrony prywatności oraz wolności obywatelskiej przez **system prawny**. Powinien on dotyczyć przetwarzania danych na potrzeby zarówno biznesu, jak i celów ogólnospołecznych, w szczególności bezpieczeństwa publicznego, obronności i zarządzania finansami publicznymi. Zagadnienia prawne są wyjątkowo aktualne. W kwietniu 2016 r. Parlament Europejski przyjął reformę unijnych przepisów o ochronie danych. Ma ona przynieść korzyści europejskiej gospodarce i jednocześnie zapewnić obywatelom większą kontrolę nad swoimi danymi osobowymi. Nowe przepisy mają: zapewnić przestrzeganie podstawowego prawa do ochrony danych, gwarantowanego przez Kartę praw podstawowych UE, przynosić korzyści obywatelom, przedsiębiorstwom i administracji publicznej, a także sprawdzać się w przyszłości i być otwarte na innowacje. Przepisy zaczną obowiązywać w 2018 r.[6] Wiosną i latem 2016 r. w Polsce toczyła się dyskusja związana z wejściem w życie tzw. ustawy antyterrorystycznej, ułatwiającej organom państwa m.in. inwigilację obywateli, blokowanie stron internetowych, a także wymuszającej rejestrowanie kart SIM[7]. Trudno obecnie stwierdzić, jak w praktyce ustawa będzie wykorzystywana. Niemniej, co do zasady, prawo nie powinno nadmiernie ograniczać możliwości wykorzystania metod *big data* w celu sprawnego zarządzania państwem, a także rozwoju całej gospodarki np. poprzez blokowanie możliwości funkcjonowania biznesu. Jednocześnie prawo musi zabezpieczać interesy poszczególnych osób, zapewniając im niezbędny poziom prywatności.

3. Opis i metoda przeprowadzonego badania

Otwarte pozostaje pytanie: jak w kontekście funkcjonowania współczesnych technologii oraz systemu prawnego zagrożenie naruszenia prywatności jest postrzegane przez osoby, których dane się przetwarza? Jaki jest stopień poczucia tego zagrożenia w zależności od metod przetwarzania i typu danych?

---

[6] J. Albrecht, M. Lauristin, V. Jourová, *Nowe przepisy o ochronie danych godne epoki cyfrowej*, „Gazeta Wyborcza", 15.04.2016.

[7] Ustawa z dnia 10 czerwca 2016 r. o działaniach antyterrorystycznych.

W literaturze można odnaleźć badania z ostatnich lat dotyczące problemów prywatności pośrednio związanych z metodami *big data*. Najczęściej jednak są ukierunkowane na ustawienia prywatności w serwisach społecznościowych i ograniczają się do prywatności w Internecie, jak badania przeprowadzone przez J. Surmę[8] oraz Ł. Kołodziejczyka[9].

W celu pełniejszego zrozumienia społecznego aspektu *big data* autor niniejszego artykułu prowadzi badania dotyczące poczucia zagrożenia wynikającego z naruszenia prywatności w sposób szerszy, poruszając też zagadnienia masowego przetwarzania danych osobowych pochodzących również spoza Internetu. Badanie prowadzone od 2014 r. ma charakter cykliczny i dotyczy, z jednej strony, powszechnego rozumienia pojęcia *big data*, z drugiej – poczucia zagrożenia wynikającego ze stosowania tych metod. Jako badaną grupę przyjęto studentów Szkoły Głównej Handlowej w Warszawie. Jest to oczywiście grupa dość homogeniczna, niereprezentatywna dla ogółu społeczeństwa, lecz taki dobór ankietowanych pozwala na zadawanie dość szczegółowych pytań dzięki wystarczającemu zrozumieniu problemu. Studenci, jako ludzie młodzi, są zazwyczaj otwarci na wykorzystywanie nowych technologii i traktują je jako rzecz w pełni naturalną. Ponadto studenci uczelni ekonomiczno-biznesowej powinni w miarę dobrze rozumieć możliwości zastosowań nowoczesnych technologii i jednocześnie podchodzić do nich jak typowi użytkownicy: z jednej strony, jako osoby prywatne, z drugiej – jako przyszli użytkownicy biznesowi.

Ankieta jest anonimowa, prowadzona na zajęciach niezwiązanych z *big data* w tradycyjnej formie papierowej, dzięki czemu uzyskiwany jest prawie 100-procentowy zwrot. Pierwsza część ankiety dotyczy znajomości i rozumienia pojęcia *big data*, druga – poczucia naruszenia prywatności. W niniejszym artykule autor skupia się na drugiej jej części, odnosząc się do ogólnych wyników części pierwszej. W trakcie trzech kolejnych semestrów poddano dotąd badaniu ponad 250 respondentów.

Część pierwsza ankiety składa się z 20 zamkniętych pytań dotyczących spotkania się z pojęciem *big data* i zaliczenia do niego wskazanych zagadnień oraz cech charakterystycznych. Ponieważ pojęcie w praktyce dopiero się kształtuje i jest wieloaspektowe, nie można, zdaniem autora, jednoznacznie ocenić jego definicyjnej znajomości. Część pytań odnosi się przede wszystkim do opinii

[8] J. Surma, *The Privacy Problem in Big Data Applications: An Empirical Study on Facebook*, ASE/IEEE International Conference on Social Computing, 2013, s. 955–958.

[9] Ł. Kołodziejczyk, *Prywatność w Internecie: postawy i zachowania dotyczące ujawniania danych prywatnych w mediach społecznych*, Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 2014.

studentów i nie może być wykorzystana do oceny ich wiedzy. Inne natomiast są związane z powszechnie akceptowanymi elementami charakterystyki pojęcia – te pozwoliły na przypisanie ankietowanym ogólnego poziomu rozumienia pojęcia *big data*, określanego dalej jako wskaźnik wiedzy. Powyższe zagadnienia zostały szerzej omówione w kontekście edukacji informatycznej w odrębnym artykule[10].

Druga część ankiety opiera się na problemach opisanego powyżej społecznego aspektu *big data*. Wykorzystując wstępne badania autora dotyczące zrozumienia pojęcia *big data* w popularnej prasie i czasopismach[11], sformułowano ogólne polecenie „Oceń w skali od 1 do 5 poziom swojego poczucia naruszenia prywatności w kontekście poniższych zjawisk" wraz z 12 szczegółowymi zagadnieniami (pytaniami) dotyczącymi zjawisk związanych z możliwościami masowego przetwarzania danych osobowych. Ankietowani mieli w przypadku każdego wskazać poziom swojego subiektywnego poczucia naruszenia prywatności w skali Likerta o wartościach od 1 do 5 (poziom 1 to brak poczucia naruszenia prywatności, poziom 5 to poważne naruszenie). Dobierając pytania, autor starał się odnieść do różnego typu zagrożeń wynikających z przetwarzania danych masowych, w szczególności związanych z nowymi możliwościami technologii informatycznych oraz tymi, które w ostatnim czasie są coraz częściej wykorzystywane. W tabeli 1 wymieniono szczegółowe zagadnienia z tej części ankiety wraz z podstawowymi wynikami badania.

4. Interpretacja wyników badania

Średnie wyniki odpowiedzi na poszczególne pytania kształtują się w przedziale od 2,3 do 3,8 (przy możliwych odpowiedziach w przedziale od 1 do 5). Odchylenia standardowe dla poszczególnych pytań zawierały się w zakresie od 0,98 (pytanie 3) do 1,26 (pytanie 10). Wynika z tego, że na większość pytań ankietowani odpowiadali w sposób dość wyważony, unikając zazwyczaj odpowiedzi

[10] I. Pawełoszek, J. Wieczorkowski, *Big data as a business opportunity: an Educational Perspective*, „Annals of Computer Science and Information Systems", vol. 5, *Proceedings of the 2015 Federated Conference on Computer Science and Information Systems*, Polskie Towarzystwo Informatyczne, IEEE Computer Society Press, 2015, s. 1563–1568.

[11] J. Wieczorkowski, P. Polak, *Big data: Three-aspect approach*, „Online Journal of Applied Knowledge Management" 2014, vol. 2, issue 2, s. 182–196.

skrajnych. Ogólnie naruszenie prywatności jest wyraźnie zauważalne, ponieważ średnia ocena to 3,1.

Tabela 1. Zagadnienia związane z poczuciem naruszenia prywatności ze średnią oceną odpowiedzi na pytania i odchyleniem standardowym

Nr	Zagadnienie	Średnia arytmetyczna	Odchylenie standardowe
1.	Korzystanie z usług przechowywania prywatnych plików w chmurze	**2,7**	1,09
2.	Możliwość dostępu nieuprawnionych osób/maszyn do prywatnych e-maili	**3,6**	1,16
3.	Gromadzenie informacji o zachowaniach użytkowników w Internecie (np. odwiedzane strony)	**3,7**	0,98
4.	Automatyczne śledzenie informacji o aktywności w portalach społecznościowych	**3,8**	1,05
5.	Gromadzenie informacji o płatnościach dokonywanych np. kartami płatniczymi	**3,7**	1,15
6.	Gromadzenie danych o zachowaniach konsumentów dzięki programom lojalnościowym	**2,7**	1,08
7.	Gromadzenie w systemach IT danych o korzystaniu z usług służby zdrowia	**2,7**	1,19
8.	Gromadzenie przez telekomy danych geolokalizacyjnych i bilingów telefonów komórkowych	**3,6**	1,15
9.	Gromadzenie danych o sieciowym wykorzystaniu urządzeń, np. o logowaniach do Wi-Fi	**2,9**	1,12
10.	Powszechny monitoring miejski i przemysłowy (kamery)	**2,6**	1,26
11.	Masowe wykonywanie zdjęć: satelitarnych, lotniczych, typu *street view*	**2,3**	1,21
12.	System identyfikacji pojazdów, np. w celu naliczania opłat za drogi płatne, wykrywania wykroczeń drogowych	**2,7**	1,24

Źródło: opracowanie własne.

Największe poczucie zagrożenia wywołują na zbliżonym poziomie: automatyczne śledzenie informacji o aktywności w portalach społecznościowych (3,8); gromadzenie informacji o zachowaniach użytkowników w Internecie (3,7); gromadzenie informacji o płatnościach dokonywanych np. kartami płatniczymi (3,7); możliwość dostępu nieuprawnionych osób/maszyn do prywatnych e-maili (3,6); gromadzenie przez telekomy danych geolokalizacyjnych i bilingów telefonów

komórkowych (3,6). Powyższe zagrożenia są związane z powszechnymi, dobrze znanymi, praktycznie nieuniknionymi we współczesnym świecie czynnościami.

Najmniejsze poczucie zagrożenia wynika z: masowego wykonywania zdjęć satelitarnych, lotniczych, typu *street view* (2,3) oraz z powszechnego monitoringu miejskiego i przemysłowego (2,6). Są to z kolei działania niezależne od ankietowanych, które dzieją się w sposób niewidoczny, częściowo zainteresowani są ich nieświadomi.

Dość niskie zagrożenie jest także związane z: korzystaniem z usług przechowywania prywatnych plików w chmurze (2,7) oraz gromadzeniem danych o zachowaniach konsumentów dzięki programom lojalnościowym (2,7). Można to interpretować jako zagrożenia dość łatwe do uniknięcia. Przykładowo, do programów lojalnościowych przystępują osoby, które nie traktują tego jako naruszenia prywatności, pozostałe osoby unikają takich programów. Możliwe jest także unikanie przechowywania prywatnych plików w chmurze.

Jedno ze szczegółowych pytań dotyczyło **danych wrażliwych** (stan zdrowia). Co ciekawe, zagrożenie wynikające z gromadzenia w systemach IT danych o korzystaniu z usług służby zdrowia nie jest oceniane wysoko (2,7). Być może jest to związane z zaufaniem do prawa, które znacząco ogranicza możliwość swobodnego przetwarzania danych wrażliwych. Jednocześnie jednak ankietowani byli to ludzie młodzi, mający zazwyczaj niewielkie problemy zdrowotne. I prawdopodobnie to może być główną przyczyną niskiej oceny tego aspektu naruszenia prywatności.

Zróżnicowanie odpowiedzi znacząco nie różni się w poszczególnych pytaniach, lecz interesujące jest to, że największe odchylenie standardowe występuje przy zjawiskach określanych jako związane z najmniejszym naruszeniem prywatności: powszechny monitoring miejski i przemysłowy (1,26), systemy identyfikacji pojazdów (1,24), masowe wykonywanie zdjęć satelitarnych, lotniczych, typu *street view* (1,21). Działania takie są więc oceniane dość niejednoznacznie. Jednocześnie najmniejsze zróżnicowanie odpowiedzi jest w przypadku jednego z działań powodujących najwyższe naruszenie prywatności – gromadzenia informacji o zachowaniach użytkowników w Internecie – z odchyleniem standardowym na poziomie 0,98.

Warte odnotowania jest **zróżnicowanie ocen na przestrzeni czasu**. Mimo że okres badania nie jest długi, ponieważ obejmuje kolejne trzy semestry, jest zauważalny stały wzrost znajomości zagadnienia *big data* (wynik pierwszej części ankiety). Jednocześnie niemal nie ma różnic w ocenie naruszenia prywatności. Średnia ocena w kolejnych semestrach to: 3,08; 3,09; 3,09. W przypadku poszczególnych pytań także nie widać wyraźnych trendów zmiany ocen.

Uzasadnione wydaje się badanie **zróżnicowania odpowiedzi według płci**, ponieważ może być to związane z innym funkcjonowaniem w sytuacjach zagrożenia. Grupa ankietowanych to 144 kobiety i 112 mężczyzn. Nie zauważa się istotnego zróżnicowania średniej łącznej oceny poczucia zagrożenia prywatności według płci (kobiety 3,09; mężczyźni 3,07). Ciekawe różnice występują jednak w przypadku poszczególnych pytań.

Mężczyźni w większym stopniu niż kobiety mają poczucie naruszenia prywatności związane z systemami identyfikacji pojazdów (mężczyźni 2,95; kobiety 2,54) oraz powszechnym monitoringiem miejskim i przemysłowym (mężczyźni 2,86; kobiety 2,43). Są to więc te same pytania, w przypadku których odnotowano większe odchylenie standardowe odpowiedzi. Można wnioskować, że zróżnicowanie takie wynika z faktu, że mężczyźni są prawdopodobnie częściej rejestrowani przez monitoring miejski w związku z agresywnym zachowaniem, a także przez urządzenia kontroli drogowej za wykroczenia drogowe.

Kobiety natomiast bardziej odczuwają naruszenie prywatności związane z gromadzeniem danych o zachowaniach konsumentów dzięki programom lojalnościowym (kobiety 2,81; mężczyźni 2,55). Wynika to być może z faktu, że kobiety prawdopodobnie przykładają większą wagę do programów lojalnościowych, dostrzegając jednocześnie zagrożenia stąd płynące. Kobiety bardziej odczuwają także naruszenia prywatności związane z możliwością dostępu nieuprawnionych osób/maszyn do prywatnych e-maili (kobiety 2,75; mężczyźni 2,36). Autorowi trudno jednak wyjaśnić tę prawidłowość.

Badanie **korelacji pomiędzy wskaźnikiem wiedzy** o *big data* (uzyskanym na podstawie danych z pierwszej części ankiety) a poczuciem naruszenia prywatności w podziale na poszczególne pytania może prowadzić potencjalnie do nieoczywistych rezultatów, ponieważ z jednej strony większa wiedza o *big data* to większa świadomość różnych zagrożeń, z drugiej zaś – lepsze zrozumienie może przekładać się na mniejsze obawy. Wyniki badania tej korelacji rzeczywiście są niejednoznaczne. Ogólnie korelacja jest niewielka (dla wszystkich pytań łącznie +0,09), dla poszczególnych pytań w większości przyjmuje wartości bliskie zera (od –0,03 do +0,09). W przypadku dwóch pytań wykracza jednak poza podany przedział. Dotyczy to pytania o możliwość dostępu nieuprawnionych osób/maszyn do prywatnych e-maili (korelacja ze wskaźnikiem wiedzy +0,15) i o automatyczne śledzenie informacji o aktywności w portalach społecznościowych (korelacja +0,14). Można to wyjaśnić faktem, że osoby lepiej rozumiejące metody *big data* także lepiej rozumieją możliwości złożonych algorytmów realizujących takie działania i w konsekwencji są świadomi potencjalnych zagrożeń.

5. Podsumowanie

Badaną grupą byli ludzie młodzi, dla których Internet, telefonia komórkowa, powszechny monitoring są od zawsze elementem życia. Niemniej w kilku ostatnich latach znacznie zwiększyły się możliwości masowego przetwarzania danych osobowych (w szczególności danych nieustrukturyzowanych, ich analizy w czasie rzeczywistym – czyli elementów typowych dla metod *big data*), pociągając za sobą niebezpieczeństwo naruszenia dotąd oczekiwanego poziomu prywatności. Respondenci wyraźnie dostrzegają ten problem w przypadku działań podejmowanych zarówno przez organy państwa, jak i przez podmioty biznesowe. Najwyższy subiektywny poziom poczucia naruszenia prywatności dotyczy typowych działań, praktycznie nieuniknionych w codziennym życiu, takich jak analiza zachowań w Internecie oraz czynności z wykorzystaniem telefonów i kart płatniczych. W mniejszym stopniu obawy dotyczą zachowań, których można (jeszcze) uniknąć.

Tego typu badania powinny być kontynuowane ze względu na stosowanie coraz nowszych metod niosących zagrożenie powszechnej inwigilacji, a także na wchodzenie w dorosłe życie ludzi, dla których techniki *big data* będą normalnością. Autor planuje także poszerzyć badanie o analizę celów, które powodują akceptację naruszania prywatności (różne cele ogólnospołeczne i komercyjne), a także analizę faktycznie podejmowanych działań w celu zachowania prywatności.

Bibliografia

Boyd D., Crawford K., *Critical questions for big data in Information*, „Communication & Society" 2012, vol. 15, issue 5, s. 662–679.

Dopierała R., *Prywatność w perspektywie zmiany społecznej*, Zakład Wydawniczy Nomos, Kraków 2013.

Kołodziejczyk Ł., *Prywatność w Internecie: postawy i zachowania dotyczące ujawniania danych prywatnych w mediach społecznych*, Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 2014.

McKinsey Global Institute, *Big data: The next frontier for innovation, competition, and productivity*, 2011, http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation [dostęp 20.07.2016].

Polańska K., Wassilew A., *Analizy big data w serwisach społecznościowych*, „Nierówności Społeczne a Wzrost Gospodarczy" 2015, nr 4, cz. 2, s. 117–128.

Surma J., *The Privacy Problem in Big Data Applications: An Empirical Study on Facebook*, ASE/IEEE International Conference on Social Computing, 2013, s. 955–958.

Tabakow M., Korczak J., Franczyk B., *Big data – definicje, wyzwania i technologie informatyczne*, „Business Informatics" 2014, vol. 1(31), s. 138–153.

Ustawa z dnia 10 czerwca 2016 r. o działaniach antyterrorystycznych.

Źródła sieciowe

Albrecht J., Lauristin M., Jourová V., *Nowe przepisy o ochronie danych godne epoki cyfrowej*, „Gazeta Wyborcza", 15.04.2016, http://wyborcza.biz/biznes/1,100897,19924478, nowe-przepisy-o-ochronie-danych-godne-epoki-cyfrowej.html#BoxBizLink [dostęp 20.07.2016].

Pawełoszek I., Wieczorkowski J., *Big data as a business opportunity: an Educational Perspective*, „Annals of Computer Science and Information Systems", vol. 5, Proceedings of the 2015 Federated Conference on Computer Science and Information Systems, Polskie Towarzystwo Informatyczne, IEEE Computer Society Press, 2015, s. 1563–1568, https://fedcsis.org/proceedings/2015/pliks/365.pdf [dostęp 20.07.2016].

Wieczorkowski J., Polak P., *Big data: Three-aspect approach*, „Online Journal of Applied Knowledge Management" 2014, vol. 2, issue 2, s. 182–196, http://www.iiakm.org/ojakm/articles/2014/volume2_2/OJAKM_Volume2_2 pp182-196.pdf [dostęp 20.07.2016].

* * *

Big Data and Privacy: Privacy Concerns in the Virtual Word – Survey Results

Abstract

The term 'big data' includes the technological (new opportunities), business (new applications) and social (consequences) aspects. The article presents the social aspect – the issue of privacy and threats related to using big data technologies, especially personal data processing, video surveillance and monitoring Internet users' behavior during different activities. The aim of the paper is the identification of the subjective perception of privacy violation related to mass personal data processing. The author presents the results of a survey that has been recently conducted among students.

Keywords: big data, privacy, personal data, invasion of privacy, surveillance