Jędrzej Wieczorkowski[1]

Metody big data w zapewnieniu bezpieczeństwa publicznego a problem prywatności

1. Wstęp

Współczesne narzędzia informatyczne i metody statystycznej analizy dużych wolumenów danych, określane jako big data, mogą znacząco wspomagać działalność administracji publicznej, w tym w szczególności zapewnianie bezpieczeństwa publicznego. Zazwyczaj wiąże się to z masowym przetwarzaniem danych osobowych (umożliwiających identyfikację osoby fizycznej) i innych danych prywatnych (nieprzeznaczonych do dostępu publicznego), co rodzi niebezpieczeństwo naruszania prywatności. Problem prywatności poruszany jest w ostatnich latach bardzo często w kontekście dostępności danych osobowych i śladów aktywności użytkowników w internecie, szczególnie w zakresie marketingowego wykorzystywania tych danych np. przez profilowane reklamy. Popularnym tematem badań jest prywatność w kontekście informacji umieszczanych w ogólnodostępnym internecie, w tym w portalach społecznościowych.

Autor także od kilku lat prowadzi badania dotyczące poczucia zagrożenia prywatności, które wynika z masowego przetwarzania danych, lecz rozszerza je poza dane pochodzące wyłącznie z internetu, uwzględniając przykładowo monitoring, dane telekomunikacyjne, przestrzenne i inne, a także wykracza poza biznesowe zastosowanie tych danych, uwzględniając także cele ogólnospołeczne. W niniejszym artykule zaprezentowano wycinek tych badań dotyczący prywatności w kontekście działań zmierzających do zapewnienia bezpieczeństwa publicznego. Celem opracowania jest z jednej strony wskazanie możliwości wykorzystania metod big data w omawianym obszarze, z drugiej zaprezentowanie postaw społecznych związanych z akceptacją tych działań. W zakresie zastosowanych metod badawczych dokonano analizy literatury fachowej, a także

[1] Szkoła Główna Handlowa w Warszawie, Kolegium Analiz Ekonomicznych, Instytut Informatyki i Gospodarki Cyfrowej.

publikacji prasowych w celu lepszego zrozumienia społecznego kontekstu zjawiska. Posłużono się także ankietą z listą pytań zamkniętych.

2. Koncepcja big data i jej zastosowania w administracji publicznej

Wyjaśnienia pojęcia big data, rozumianego jako analiza dużych wolumenów danych, często szuka się w modelu kilku „V". Pierwotny model 3 V, choć początkowo nie dotyczył ściśle pojęcia big data, dość dobrze charakteryzuje późniejsze trendy w zjawisku przetwarzania danych masowych. Mówi się więc o dużym wolumenie (ang. *volume*), zmienności (ang. *velocity*) i różnorodności (ang. *variety*) danych[2]. Tak więc przykładowo według Gatner Group big data rozumiane jest jako zasoby informacyjne dużych rozmiarów, szybko zmieniające się i/lub charakteryzujące się dużą różnorodnością, które wymagają efektywnych kosztowo i innowacyjnych form przetwarzania, umożliwiających poprawę wglądu w dane, podejmowanie decyzji i automatyzację procesów[3]. W praktyce więc często rozwija się model „V", zwracając uwagę na kolejne cechy big data, takie jak duża wartość (ang. *value*), wiarygodność (ang. *veracity*) i potrzeba wizualizacji (ang. *visualization*) danych. Niemniej jednak najistotniejsze wydają się trzy cechy z pierwotnego modelu, z tym że duży wolumen jest pojęciem bardzo względnym i zmieniającym się wraz z rozwojem technologii. Szczególnie istotne są więc dwie pozostałe cechy, lecz w praktyce przy przetwarzaniu uznawanym za big data nie zawsze występują one jednocześnie. Można więc uznać za definicją Gartner Group, że wystarczające jest spełnienie co najmniej jednej z tych cech: zmienności (typowe przy przetwarzaniu strumieniowym np. obrazu wideo, a także danych z analizy aktywności w internecie w czasie rzeczywistym) lub różnorodności danych (np. wykorzystanie nierelacyjnych, słabo ustrukturyzowanych danych).

Zdaniem autora należy wyróżnić trzy podstawowe aspekty w rozumieniu pojęcia big data: techniczny – dotyczący możliwości właściwego IT i metod analitycznych; ekonomiczny (inaczej biznesowy) – dotyczący zastosowań, w omawianym kontekście zastosowań w administracji publicznej; społeczny – dotyczący

[2] D. Laney, *Application delivery strategies*, META Group, Stamford 2001.
[3] What is Big Data, http://www.gartner.com/it-glossary/big-data/ (30.11.2016).

konsekwencji społecznych, w szczególności zagrożenia utratą prywatności wynikającą z masowego przetwarzania danych osobowych.

Dużymi wolumenami danych charakteryzuje się statystyka publiczna oraz publiczne rejestry. Przetwarzanie takich danych jak dotąd jest przede wszystkim przetwarzaniem transakcyjnym i nie spełnia typowych wymogów big data związanych ze zmiennością i różnorodnością danych. Nie zmienia to faktu, że przykładowo rejestry publiczne mogą być jednym z istotnych źródeł danych dla zastosowań metod big data.

Klasyczne zastosowania metod big data dotyczą zaś zapewnienia bezpieczeństwa państwa. W szczególności obejmują typowe aspekty bezpieczeństwa publicznego i obronności realizowane przez policję i służby specjalne. Jednocześnie metody big data wykorzystywane są przez różne organy administracji publicznej i znajdują zastosowanie w takich obszarach jak bezpieczeństwo: transportu, cyberprzestrzeni, sanitarne i epidemiologiczne, klimatyczne i związane z klęskami żywiołowymi, finansowe[4]. Rozwiązania typu big data w administracji publicznej mogą opierać się na wielu źródłach – np. z wykorzystaniem platform integracyjnych ESB, jak przykładowe rozwiązanie wspomagające bezpieczeństwo obrotu żywnością[5]. Wśród typowych danych strumieniowych zastosowanie znajdują przykładowo dane z monitoringu miejskiego (wzbogacane np. automatycznymi systemami rozpoznawania twarzy) i drogowego (np. automatyczne sprawdzanie tablic rejestracyjnych)[6]. Bliską przyszłością jest zastosowanie do podobnych celów dronów. Już dziś powszechnie wykorzystuje się dane telekomunikacyjne (bilingi i lokalizacja). Możliwa jest inwigilacja obywateli na podstawie transakcji płatniczych (płatności kartami, przelewy). Ogromnym źródłem wiedzy dla policji i służb specjalnych jest internet, zarówno w zakresie danych ogólnie dostępnych, jak i ściśle prywatnych. Wykorzystanie powyższych danych wymaga poza stosownymi narzędziami, także możliwości prawnych lub procedur działania pozaprawnego.

---

4 *Demystifying Big Data: A Practical Guide to Transforming the Business of Government*, TechAmerica Foundation, Washington 2012.

5 T. Górski, W. Kuchta, *Zastosowanie magistrali usług ESB do przesyłania dużych wolumenów danych*, „Roczniki Kolegium Analiz Ekonomicznych" 2015, z. 38, s. 99–116.

6 Por. C. Stępniak, *Kierunki wykorzystania systemów monitoringu miejskiego w zarządzaniu rozwojem miast*, „Roczniki Kolegium Analiz Ekonomicznych" 2013, z. 29, s. 295–307.

3. Problematyka prywatności w kontekście bezpieczeństwa publicznego

Powyższe obszary zastosowań metod big data wiążą się z wykorzystywaniem danych różnego typu, w szczególności:

- nieosobowych,
- osobowych zanonimizowanych lub zagregowanych,
- szczegółowych danych osobowych i innych danych prywatnych.

Dane nieosobowe są przykładowo podstawą do wspomagania bezpieczeństwa klimatycznego, dane zanonimizowane i zagregowane – dla bezpieczeństwa epidemiologicznego. Jednakże dla typowego bezpieczeństwa publicznego podstawą jest przetwarzanie szczegółowych danych osobowych, a to jest związane z problematyką prywatności. Już w pierwszej dojrzałej koncepcji prywatności z 1890 r. Warren i Brandeis uznali za niedopuszczalne ujawnianie życia prywatnego, jeśli nie wiąże się to z interesem publicznym[7]. Problemem w praktyce pozostaje równowaga pomiędzy tym interesem publicznym a prywatnością. Na potrzeby rozważań o big data prywatność można rozumieć w szczególności jako kontrolę przepływu informacji prywatnej zgodnie z koncepcją Nisseubaum, w której problem kontroli dostępu do takich informacji rozważa się w określonym kontekście społecznym – które informacje, komu, kiedy i jakiej sytuacji powinny zostać przekazane[8].

Temat naruszania prywatności w kontekście masowego przetwarzania pochodzących z internetu danych osobowych i innych prywatnych na potrzeby bezpieczeństwa publicznego stał się głośny na świecie w 2013 r. w związku z aferą Edwarda Snowdena. Ponieważ ze zrozumiałych względów przekazywane przez niego informacje nie były potwierdzane przez stosowne urzędy, można jedynie opierać się na materiałach prasowych[9]. Wiadomo jednak, że uruchomiony już w 2007 r., oficjalnie w celu zapewnienia bezpieczeństwa USA i walki z terroryzmem, przez NSA (National Security Agency) program PRISM służy do

---

[7] S. Warren, L.D. Brandeis, *The Right to Privacy*, "Harvard Law Review" 1890, vol. IV, no. 5.

[8] H. Nissenbaum, *Privacy as Contextual Integrity*, "Washington Law Review" 2004, no. 79, pp. 101–139.

[9] W szczególności dotyczy to publikacji w „Washington Post": https://www.washingtonpost.com/ i w Guardian https://www.theguardian.com/uk (30.11.2016).
Ponadto część tych informacji zebrano w artykule: S. Iskierka, J. Krzemiński, Z. Weżgowiec, *Bezpieczeństwo i prywatność w sieci po ujawnieniu afery PRISM*, w: *Edukacja międzykulturowa w warunkach kultury globalnej*, Poznań 2014, s. 333–339.

kontrolowania komunikacji poprzez analizę dużych wolumenów danych przechwyconych z internetu. Kontrola dotyczy w szczególności danych pochodzących z e-maili, czatów, plików audio i wideo, fotografii, portali społecznościowych, różnych plików przechowywanych w chmurze, telefonii VoIP. Analizowane dane w bardzo dużej części pochodziły z serwerów takich firm jak Google, Microsoft, Yahoo, czyli o znaczeniu globalnym. Poza PRISM wykorzystywane są inne aplikacje do zbierania danych z infrastruktury teleinformatycznej oraz do łamania algorytmów zabezpieczających połączenia.

Podobne podejrzenia padły także na wywiad brytyjski i dotyczą programu Tempora realizowanego przez GCHQ (Government Communications Headquarters). Medialnie głośne stały się również powiązania rosyjskiego wywiadu z producentem oprogramowania antywirusowego Kaspersky[10]. Prawdopodobnie więc służby specjalne wielu krajów metodami zgodnymi z prawem lub pozaprawnymi naruszają prywatność obywateli w zakresie kontroli komunikacji internetowej i telefonicznej.

W Polsce brak jest powszechnie dostępnych informacji o programach realizowanych na podobną skalę, choć prawdopodobnie polskie służby dostarczały informacje amerykańskim odpowiednikom. Próbę szerszych badań skali oficjalnej inwigilacji w internecie prowadziła Fundacja Panoptykon w latach 2012 i 2013[11]. Liczba zapytań o dane użytkowników kierowanych oficjalną drogą przez różne organy państwa do usługodawców elektronicznych była wówczas stosunkowo niewielka. Ich struktura sugeruje, że dotyczyły głównie postępowań karnych. W kolejnych latach nastąpił wyraźny wzrost liczby zapytań o dane internetowe, lecz z drugiej strony wiadomo o znacząco większej liczbie zapytań policji i innych służb do operatorów o dane telekomunikacyjne: bilingi, dane abonenckie i geolokalizacyjne[12].

Jak zauważyła ówczesna Rzecznik Praw Obywatelskich I. Lipowicz ochrona prywatności postrzegana jest jako główny cel ochrony danych osobowych obok innych celów społecznych i państwowych[13]. Ustawa o ochronie danych osobowych ma jednak wiele wyłączeń swoich szczegółowych zapisów, jeśli ich

[10] Bloomberg.org, https://www.bloomberg.com/news/articles/2015–03–19/cybersecurity-kaspersky-has-close-ties-to-russian-spies (30.11.2016).

[11] K. Szymielewicz, M. Szumańska, *Dostęp państwa do danych użytkowników usług internetowych, Siedem problemów i kilka hipotez*, Fundacja Panoptykon, Warszawa 2013.

[12] *Nowe liczby, stare problemy?*, Panoptykon.org, https://panoptykon.org/wiadomosc/nowe-liczby-stare-problemy (30.11.2016).

[13] I. Lipowicz, *Nowe wyzwania w zakresie danych osobowych*, w: *Internet. Ochrona wolności, własności i bezpieczeństwa*, red. G. Szpor, Wydawnictwo C.H. Beck, Warszawa 2011.

działanie byłoby związane np. z zagrożeniem dla obronności lub bezpieczeństwa państwa, życia i zdrowia ludzi lub bezpieczeństwa i porządku publicznego[14]. Na początku 2016 r. w Polsce zmodyfikowano prawo, w szczególności Ustawę o Policji[15], ułatwiając działania służb naruszające prywatność (przez tzw. ustawę inwigilacyjną). Wprowadziła ona pojęcie danych internetowych. Nowe zapisy, jako bardzo niejednoznaczne i wprowadzające mechanizmy kontrolne trudne do wyegzekwowania, zostały skrytykowane przez niektóre środowiska i według stanu na koniec 2016 r. czekają na postępowanie w Trybunale Konstytucyjnym. Na podstawie zmienionych przepisów można przypuszczać, że władza szykuje grunt pod możliwości znacznego rozszerzenia inwigilacji obywateli.

Otwarte pozostają pytania:

- na ile inwigilacja obywateli opierająca się na masowej analizie danych prywatnych jest uzasadniona interesem społecznym?
- do jakiego stopnia system prawny zabezpiecza interes osób prywatnych i daje poczucie bezpieczeństwa w zakresie prywatności?
- czy obywatele ufają swojemu oraz innym państwom w zapewnieniu prywatności?

4. Koncepcja i wyniki badań dotyczących zagrożenia prywatności

W odpowiedziach na powyższe pytania pomagają wyniki badań prowadzonych przez autora od 2014 r. Opisana poniżej forma ankiety z zamkniętymi pytaniami sprawdza się w szczególności w badaniach obejmujących szeroki zakres tematyczny (jakim jest big data). Za pomocą szczegółowych pytań można zawęzić badany obszar do fragmentów istotnych z punku widzenia celów badania. Zdecydowano się przygotować dość szczegółową, lecz nie za obszerną listę zagadnień, których mają dotyczyć pytania. Wykorzystano w tym celu badanie wstępne przeprowadzone przez autora w 2013 r. Dotyczy ono potocznego rozumienia pojęcia big data, które zostało odtworzone na podstawie analizy artykułów

[14] Ustawa o ochronie danych osobowych z dnia 29 sierpnia 1997 r., z późniejszymi zmianami (DzU 1997 nr 133 poz. 883).

[15] Ustawa o Policji z dnia 6 kwietnia 1990 r., z późniejszymi zmianami (DzU 1990 nr 30 poz. 179).

opublikowanych przede wszystkim w prasie popularnej[16]. Analiza ta wykazała ścisły związek potocznego rozumienia pojęcia big data z jego społecznym aspektem, w szczególności z zagrożeniem prywatności. Wykorzystując wyniki tego badania, przygotowano w ankiecie 12 pytań dotyczących subiektywnego poczucia zagrożenia prywatności związanego z masowym przetwarzaniem prywatnych danych pochodzących z różnych źródeł oraz 9 pytań dotyczących celów takiego przetwarzania. Ankietowani wskazywali poziom poczucia zagrożenia swojej prywatności oraz zgody na jej naruszanie.

Jako respondentów wybrano studentów uczelni ekonomicznej. Oczywiście nie można w tym przypadku mówić o reprezentatywności takiej próby dla ogółu społeczeństwa. Jednak należy wziąć pod uwagę fakt, że ankietowani muszą spełniać pewne warunki, które uniemożliwiają w praktyce badanie na próbie w pełni reprezentatywnej. Badani muszą być dość dobrze zaznajomieni z nowoczesnymi technologiami i metodami przetwarzania danych, przede wszystkim w obszarze zastosowań (rozumieć pojęcie big data i jego ekonomiczny aspekt). Muszą także zdawać sobie sprawę z zagrożeń wykorzystania takich metod (aspekt społeczny). Studenci uczelni ekonomicznej, jako przyszli użytkownicy biznesowi, a obecnie zapewne również aktywni użytkownicy prywatni stykający się z omawianymi metodami, wydają się dość dobrą próbą na obecnym etapie badania. Ankietę przeprowadzono dotąd na 385 (liczba uzyskanych kompletnych ankiet) studentach Szkoły Głównej Handlowej w Warszawie w latach 2014–2016 w pięciu kolejnych semestrach. Ankieta jest anonimowa, w formie papierowej, a dzięki bezpośredniemu kontaktowi z prowadzącym badanie ma niemal 100% zwrot. Można także przypuszczać, że ankiety są wypełnione dość rzetelnie, gdyż przykładowo nie ma ankiet, w których w sposób automatyczny dla dużej serii pytań kolejnych powtarzano by takie same odpowiedzi. Liczba arkuszy niekompletnie wypełnionych nie przekracza 3%. Takie ankiety nie były brane pod uwagę przy analizie wyników.

Pierwsze pytanie polegało na ocenie w skali od 1 do 5 poziomu swojego subiektywnego poczucia naruszenia prywatności związanego ze wskazanymi zjawiskami. Oparto się więc na klasycznej skali Likerta, gdzie poziom 1 oznaczał brak poczucia naruszenia prywatności, poziom 5 – poważne jego naruszenie. W pytaniach poruszono problem niepożądanego dostępu na wielką skalę do danych pochodzących z internetu (chmura, poczta elektroniczna, śledzenie historii aktywności), a także m.in. gromadzenia historii płatności kartami,

[16] J. Wieczorkowski, P. Polak, *Big data: Three-aspect approach*, "Online Journal of Applied Knowledge Management" 2014, vol. 2, no. 2, pp. 182–196.

danych lokalizacyjnych i bilingów pochodzących od telekomów, danych przestrzennych, z monitoringu miejskiego i drogowego. Wymieniono jednak tylko rodzaje danych i dość ogólne hasła związane z możliwościami masowego przetwarzania danych osobowych, bez wskazania celu przetwarzania tych danych. Średnie arytmetyczne wyniki poczucia naruszenia prywatności dla poszczególnych pytań zawierają się w przedziale od 2,2 do 3,8, przy ogólnym średnim wyniku 3,1.

Z jednej strony najwyższy poziom poczucia zagrożenia związany jest ze śledzeniem aktywności w internecie. Przykładowo pytanie o „gromadzenie informacji o zachowaniach użytkowników w internecie (np. odwiedzane strony)" dało średni wynik 3,8 przy odchyleniu standardowym 0,99. Gromadzenie takich danych może być użyteczne z punktu widzenia zapewnienia bezpieczeństwa publicznego, lecz zapewne kojarzy się przede wszystkim z działaniami na potrzeby dostarczania spersonalizowanej reklamy.

Z drugiej strony jeden z najniższych średnich wyników otrzymało pytanie o „powszechny monitoring miejski i przemysłowy (kamery)" – 2,5, przy odchyleniu standardowym 1,24. Jednocześnie poziom zagrożenia prywatności związany z tą formą inwigilacji wyraźnie systematycznie maleje w kolejnych edycjach badania (od 2,72 do 2,14 w ciągu dwóch i pół roku). Zdecydowanie głównym celem monitoringu jest zapewnienie bezpieczeństwa publicznego. Można więc przypuszczać, że działania związane z zapewnieniem bezpieczeństwa są oceniane jako mniej naruszające prywatność niż te związane z celami marketingowymi. Zweryfikować taką tezę może pomóc analiza odpowiedzi z drugiej części ankiety, która poniżej zostanie omówiona bardziej szczegółowo.

Postawiono więc pytanie: „Do realizacji których z poniższych celów zgodziłbyś się na naruszenie swojej prywatności w skali od 1 do 5?". Poziom 1 oznacza brak zgody na naruszenie prywatności, poziom 5 – całkowitą zgodę. W celu doprecyzowania problemów (celów) przy każdym pytaniu podano przykład. Średnie wyniki odpowiedzi na poszczególne pytania zawierają się w przedziale od 2,2 do 4,0 przy ogólnym średnim wyniku 3,1. Widoczne jest więc znaczące zróżnicowanie poziomu akceptacji uzależnione od celu przetwarzania danych. Sześć pytań dotyczyło zapewnienia szeroko rozumianego bezpieczeństwa. Pytania te wraz ze średnią arytmetyczną odpowiedzi i odchyleniem standardowym zawarto w tabeli 1. Średni poziom akceptacji dla poszczególnych pytań z tej grupy mieścił się w przedziale od 2,8 do 4,0, przy średniej dla całej grupy 3,4.

Pozostałe zadane pytania (nieomówione tutaj tak szczegółowo) dotyczyły przetwarzania danych osobowych na potrzeby marketingowe – indywidualizacji reklam i oferty handlowej, np. na podstawie aktywności internetowej i programów

lojalnościowych. Wszystkie cele dotyczące potrzeb marketingowych otrzymały niższy poziom akceptacji naruszenia prywatności (średnia arytmetyczna dla poszczególnych pytań w przedziale od 2,2 do 2,7, dla całej grupy 2,5) niż którykolwiek z wyżej wymienionych celów dotyczących bezpieczeństwa.

Tabela 1. Poziom akceptacji naruszenia prywatności w zależności od celu wykorzystania danych

Nr.	Do realizacji których z poniższych celów zgodziłbyś się na naruszenie swojej prywatności (w skali od 1 do 5)?	Średnia arytmetyczna	Odchylenie standardowe
1	Zapewnienie bezpieczeństwa publicznego (np. wykorzystanie monitoringu w miejscach publicznych).	4,0	0,99
2	Wykrywanie przestępstw (np. analiza bilingów telefonicznych i danych geolokalizacyjnych).	3,6	1,03
3	Przeciwdziałanie terroryzmowi (np. częściowa kontrola poczty elektronicznej, plików przechowywanych w chmurze).	3,4	1,26
4	Wykrywanie naruszeń podatkowych (np. wykrywanie szarej strefy i analiza majątku z wykorzystaniem śledzenia ogólnodostępnych treści w internecie).	2,8	1,17
5	Poprawa bezpieczeństwa transportu (np. wykorzystanie monitoringu drogowego, fotoradarów itp.).	3,3	1,23
6	Poprawa funkcjonowania służby zdrowia i przeciwdziałanie zagrożeniom epidemiologicznym przy anonimizacji danych o zdrowiu pacjentów (np. dostęp do historii leczenia).	3,6	1,18

Źródło: opracowanie własne.

Wewnątrz pytań dotyczących bezpieczeństwa jednak także występują znaczące różnice. Najwyższy poziom akceptacji naruszenia prywatności dotyczy:

- zapewnienia bezpieczeństwa publicznego, np. przez wykorzystanie kamer monitoringu (4,0),
- poprawy funkcjonowania służby zdrowia i przeciwdziałanie zagrożeniom epidemiologicznym przy anonimizacji danych o zdrowiu pacjentów (3,6),
- wykrywania przestępstw, np. przez wykorzystanie danych od telekomów (3,6).

Wynik związany z pierwszym wymienionym celem jest logiczną konsekwencją stosunkowo niewielkiego poczucia zagrożenia prywatności wynikającego z pytania o monitoring z poprzedniej części ankiety. Warto także zwrócić

uwagę na pytanie o wykorzystywanie danych zanominowanych (na potrzeby służby zdrowia). Choć akceptacja jest dość wysoka, lecz być może nie aż tak, jak mogłoby to wynikać z faktu anonimizacji danych.

Niższy poziom akceptacji dotyczy:

- przeciwdziałania terroryzmowi, np. przez częściową kontrolę poczty elektronicznej i plików przechowywanych w chmurze (3,4),
- poprawy bezpieczeństwa transportu z wykorzystaniem monitoringu drogowego i fotoradarów (3,3),
- wykrywania naruszeń podatkowych z wykorzystaniem śledzenia ogólnodostępnych treści w internecie (2,8).

W szczególności interesujący jest niski poziom akceptacji związany z ostatnim celem (naruszenia podatkowe), tym bardziej że wskazano na treści ogólnodostępne. Jest to prawdopodobnie związane z polskimi uwarunkowaniami społecznymi, w których istnieje wysoka tolerancja obchodzenia podatków, a co za tym idzie pośrednio naruszania bezpieczeństwa finansowego państwa. Interesujący jest także dość niski poziom akceptacji dla urządzeń kontroli drogowej, związany najprawdopodobniej z powszechnością i tolerancją łamania przepisów drogowych i w konsekwencji naruszania bezpieczeństwa transportu.

Warto także zwrócić uwagę na poziom akceptacji naruszenia prywatności związanego z przeciwdziałaniem terroryzmowi przez kontrolę poczty elektronicznej i plików. Jest on wyraźnie niższy niż np. związany z monitoringiem, jednocześnie jest dość różnorodnie oceniany (najwyższe odchylenie standardowe). Zauważa się także ostatnio stopniowy spadek tej akceptacji w kolejnych badanych okresach (od 3,41 do 3,14 w ciągu dwóch i pół roku). Można to interpretować m.in. jako wynik zmian prawnych, które ostatnio znacznie zwiększyły uprawnienia organów państwa i w konsekwencji nagłośnienia sprawy oraz obaw przed nadużyciami w wykorzystywaniu niejednoznacznego prawa.

5. Podsumowanie i kierunki dalszych badań

Problem zagrożenia naruszenia prywatności związany z masowym przetwarzaniem danych osobowych i prywatnych jest wyraźnie zauważany przez ankietowanych studentów. Zdecydowanie wyższe poczucie naruszenia prywatności wynika z obaw przed śledzeniem aktywności w internecie (kontrola poczty, plików w chmurze, odwiedzanych stron, wpisów w portalach społecznościowych itd.) oraz śledzeniem wykorzystania telefonów komórkowych niż

z życia w stale monitorowanej przestrzeni (kamery). Ten ostatni element staje się częścią codzienności, na którą przestaje się zwracać uwagę. Istotny jest jednak cel, w którym podejmowana jest inwigilacja. Wyraźnie wyższy jest poziom akceptacji naruszenia prywatności w celach ogólnospołecznych, w szczególności zapewnienia bezpieczeństwa publicznego niż w celach marketingowych. Nie zmienia to faktu, że inwigilacja związana z zapewnieniem bezpieczeństwa także nie jest traktowana obojętnie. Zawsze pozostanie istotny dylemat zrównoważenia prawa do wolności i prywatności z zapewnieniem bezpieczeństwa publicznego.

Szczególnie istotne więc wydaje się budowanie zaufania do instytucji państwa i do systemu prawnego. Z jednej strony niestaranne budowanie prawa, z drugiej jego obchodzenie przez organy państwa może wpływać na zmniejszenie poziomu akceptacji wykorzystywania prywatnych danych do celów publicznych. W konsekwencji może to wpływać przykładowo na szersze wykorzystywanie mechanizmów zapewniających większą prywatność w internecie (takich jak zaawansowane szyfrowanie, tunel VPN, sieć TOR itp.), co utrudnia ważne działania uprawnionych organów państwa. Problem zaufania do instytucji dotyczy nie tylko poziomu państwa, należy bowiem zauważyć, że znacząca część usług internetowych wykorzystywanych przez polskich obywateli znajduje się poza jurysdykcją i faktyczną kontrolą państwa polskiego.

W dalszych planach badawczych autora pozostaje poszerzenie badanej grupy w celu zmniejszenia jej homogeniczności, rozszerzenie badań o działania faktycznie podejmowane przez respondentów mające na celu ochronę prywatności, a także wyjście poza wyłącznie metodę ankietową w celu pogłębienia zrozumienia postawy badanych osób.

Bibliografia

Demystifying Big Data: A Practical Guide to Transforming the Business of Government, TechAmerica Foundation, Washington 2012.

Górski T., Kuchta W., *Zastosowanie magistrali usług ESB do przesyłania dużych wolumenów danych*, „Roczniki Kolegium Analiz Ekonomicznych" 2015, z. 38.

Iskierka S., Krzemiński J., Weżgowiec Z., *Bezpieczeństwo i prywatność w sieci po ujawnieniu afery PRISM*, w: *Edukacja międzykulturowa w warunkach kultury globalnej*, red. N. Dębowska, M. Walachowska, N. Starik, Poznań 2014.

Laney D., *Application delivery strategies*, META Group, Stamford 2001.

Lipowicz I., *Nowe wyzwania w zakresie danych osobowych*, w: *Internet Ochrona wolności, własności i bezpieczeństwa*, red. G. Szpor, Wydawnictwo C.H. Beck, Warszawa 2011.

Nissenbaum H., *Privacy as Contextual Integrity*, "Washington Law Review" 2004, no. 79.

Stępniak C., *Kierunki wykorzystania systemów monitoringu miejskiego w zarządzaniu rozwojem miast*, „Roczniki Kolegium Analiz Ekonomicznych" 2013, z. 29.

Szymielewicz K., Szumańska M., *Dostęp państwa do danych użytkowników usług internetowych, Siedem problemów i kilka hipotez*, Fundacja Panoptykon, Warszawa 2013.

Ustawa o Policji z dnia 6 kwietnia 1990 r., z późniejszymi zmianami (Dz.U. 1990 nr 30 poz. 179).

Ustawa o ochronie danych osobowych z dnia 29 sierpnia 1997 r., z późniejszymi zmianami (Dz.U. 1997 nr 133 poz. 883).

Warren S.D., Brandeis L.D., *The Right to Privacy*, "Harvard Law Review" 1890, vol. IV, no. 5.

Wieczorkowski J., Polak P., *Big data: Three-aspect approach*, "Online Journal of Applied Knowledge Management" 2014, vol. 2, no. 2, http://www.iiakm.org/ojakm/articles/2014/volume2_2/OJAKM_Volume2_2pp182–196.pdf (30.11.2016).

Źródła sieciowe

Bloomberg.com, https://www.bloomberg.com (30.11.2016).

Nowe liczby, stare problemy, Panoptykon.org, https://panoptykon.org/wiadomosc/nowe-liczby-stare-problemy (30.11.2016).

The Guardian.com, https://www.theguardian.com/uk (30.11.2016).

Washington post.com, https://www.washingtonpost.com/ (30.11.2016).

What is Big Data?, http://www.gartner.com/it-glossary/big-data/ (30.11.2016).

* * *

Big Data Methods in Public Security and Privacy Issues

Abstract

The article presents, in the context of the big data phenomenon, the results of a survey on privacy threats. It focusses on public security. The survey analyses the level of acceptance of privacy violation resulting from mass data processing depending on the purpose of processing. The paper also describes the application of big data methods in public security and the overall concept of privacy.

Keywords: privacy, big data, public security, personal data