Ochrona danych osobowych a Internet rzeczy, profilowanie i repersonalizacja danych

Paulina Leja[1]

Celem niniejszego opracowania jest opis Internetu rzeczy oraz repersonalizacji danych, a także skutków i zagrożeń, jakie ze sobą niosą w kontekście ochrony danych osobowych. Autorka odnosi przedstawiane kwestie do regulacji zawartych w rozporządzeniu Parlamentu Europejskiego i Rady (UE) 2016/679 z 27.4.2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych)[2].

Uwagi wstępne

Internet to ogólnoświatowy system połączeń między komputerami. Obecnie, jak podaje największy portal społecznościowy Facebook, 3,2 mld ludzi ma dostęp do Internetu i każdego roku zwiększa się on o 200–300 tys. osób[3]. Jednocześnie Internet odnosi się nie tylko do przestrzeni adresów IP przydzielonych serwerom i hostom, lecz także staje się instrumentem pozwalającym gromadzić dane o swoich użytkownikach.

Takie zjawisko wymusza pytanie o granice i środki ochrony danych osobowych, które stają się przedmiotem coraz bardziej szczegółowych regulacji, również ustawodawcy unijnego. Pogodzić on musi konieczność ochrony podstawowych praw i wolności osób fizycznych[4] w związku z czynnościami przetwarzania danych oraz interesy podmiotów, które te dane gromadzą.

Najważniejszym aktem prawa unijnego dotyczącym poruszanego zagadnienia jest RODO. Rozporządzenie to weszło w życie 25.5.2016 r., lecz jego bezpośrednie stosowanie, we wszystkich państwach członkowskich, rozpocznie się od 25.5.2018 r.

Rozporządzenie 2016/679

W preambule RODO znaleźć możemy zapis: „Szybki postęp techniczny i globalizacja przyniosły nowe wyzwania w dziedzinie ochrony danych osobowych. Skala zbierania i wymiany danych osobowych znacząco wzrosła. Dzięki technologii zarówno przedsiębiorstwa prywatne, jak i organy publiczne mogą na niespotykaną dotąd skalę wykorzystywać dane osobowe w swojej działalności. Osoby fizyczne coraz częściej udostępniają informacje osobowe publicznie i globalnie. Technologia zmieniła gospodarkę i życie społeczne i powinna nadal ułatwiać swobodny przepływ danych osobowych w Unii oraz ich przekazywanie do państw trzecich i organizacji międzynarodowych, równocześnie zaś powinna zapewniać wysoki stopień ochrony danych osobowych”[5]. Tym samym należy wywnioskować, że ustawodawca unijny świadomy jest postępującego rozwoju technologicznego, który niesie ze sobą zjawiska z zakresu zarządzania danymi osobowymi, często uregulowanymi prawnie w niewielkim stopniu, lub w ogóle. Wprowadzając pewne mechanizmy ochrony oraz normy celowościowe, ustawodawca stara się wyjść im naprzeciw. Odnosząc się do wybranych zjawisk występujących na skalę globalną, chciałabym przeprowadzić analizę zakresu ochrony, którą przyznaje europejska reforma o ochronie danych osobowych.

Zakresem zastosowania RODO jest przetwarzanie[6] danych osobowych w sposób całkowicie lub częściowo zautomatyzowany oraz przetwarzanie danych w sposób inny niż zautomatyzowany, o ile stanowią lub mają stanowić część zbioru danych[7]. Tytułem wyjaśnienia, termin „dane osobowe” – zgodnie z definicją zawartą w art. 4 RODO – oznacza informacje o zidentyfikowanej lub możliwej do zidentyfikowania osobie fizycznej. „Możliwość” zidentyfikowania opiera się na identyfikacji bezpośredniej lub pośredniej. Dane te mogą być wyrażone w dowolny sposób[8]. Aby dane osobowe można było uznać za zbiór, konieczne jest ich uporządkowanie według określonych kryteriów[9]. Kluczowe jest posiadania przez zbiór odpowiedniej struktury, co nie oznacza jednak uporządkowania poszczególnych elementów, ale jedynie ich dostępność[10].

[1] Autorka jest studentką prawa na Wydziale Prawa Administracji i Ekonomii Uniwersytetu Wrocławskiego.

[2] Dz.Urz. UE L Nr 119, s. 1; dalej jako: RODO.

[3] Zob. http://www.computerworld.pl/news/404591/Dostep-do-Internetu-ma-juz-3-2-mld-mieszkancow-Ziemi.html (dostęp z 30.3.2017 r.).

[4] Zob. art. 8 ust. 1 Karty praw podstawowych Unii Europejskiej (Dz.Urz. UE z C 2016 r. Nr 202, s. 1); art. 16 ust. 1 Traktatu o funkcjonowaniu Unii Europejskiej (Dz.Urz. UE C z 2012 r. Nr 326, s. 1).

[5] Zob. motyw 6 preambuły RODO.

[6] Więcej o przetwarzaniu zob. *P. Carey*, Data Protection, Oxford 2004, s. 20.

[7] *D. Wociór*, [w:] *D. Wociór* (red.), Ochrona danych osobowych i informacji niejawnych z uwzględnieniem rozporządzenia unijnego, Warszawa 2016, s. 5.

[8] *J. Barta*, *P. Fajgielski*, *R. Markiewicz*, Ochrona danych osobowych. Komentarz, Warszawa 2004, s. 391.

[9] *P. Kowalik*, [w:] *D. Wociór* (red.), Ochrona danych osobowych…, s. 45–46.

[10] *P. Fajgielski*, Ochrona danych osobowych w telekomunikacji – aspekty prawne, Lublin 2003, s. 47.

W RODO zawarto również normę traktującą o tym, że „osoby fizyczne powinny mieć kontrolę nad własnymi danymi osobowymi". Regulacja ta wraz z art. 6 – dającym legitymację legalności przetwarzania – stanowić będzie kanwę poniższych rozważań dotyczących wybranych zjawisk internetowych. Artykuł 6 RODO zawiera enumeratywną listę, z której podmiot musi spełniać co najmniej jeden wymóg, aby działania w zakresie przetwarzania danych osobowych były legalne. Mowa tu o:

1) zgodzie podmiotu na przetwarzanie danych osobowych;
2) niezbędności przetwarzania w celu wykonania umowy lub podjęcia działań przed zawarciem umowy, na żądanie osoby, której dane dotyczą;
3) niezbędności do wykonania obowiązków ciążących na administratorze danych;
4) niezbędności do ochrony żywotnych interesów osoby, której dane dotyczą, lub innej osoby fizycznej;
5) niezbędności do wykonania zadania realizowanego w interesie publicznym lub w ramach sprawowania władzy publicznej powierzonej administratorowi;
6) niezbędności do celów prawnie uzasadnionych interesów realizowanych przez administratora lub przez stronę trzecią, z tym że zastrzeżone są wyjątki, tj. występowanie interesów lub podstawowych praw i wolności osoby, które mają charakter nadrzędny wobec prawnie uzasadnionych interesów, które chce realizować administrator lub osoba trzecia, w szczególności kiedy dane dotyczą dziecka. Prawnie realizowane interesy muszą mieć podstawy w prawie unijnym lub prawie państwa członkowskiego, któremu podlega administrator.

W kontekście tych przesłanek oraz norm celowościowych zawartych w preambule RODO, chciałabym ocenić legalność oraz wskazać potencjalne zagrożenia, które powstały w obecnym stanie prawnym, w związku z wybranymi zjawiskami mającymi miejsce w Internecie, a w szczególności: Internetem przedmiotów i repersonalizacją danych.

Internet przedmiotów

Internet przedmiotów nazywany jest również Internetem rzeczy. W periodyku Forbes ukazał się artykuł[11], w którym *P. Prajsnar* – ekspert ds. analityki Big Data – wskazuje, że „wbrew pozorom Internet rzeczy wcale nie jest Internetem »rzeczy«, lecz danych. Gdyby nie dane, które stanowią paliwo IoT[12], byłby on tylko fabryką elektrośmieci". W najprostszym ujęciu Internet przedmiotów to zespół rozwiązań umożliwiający osobom i przedmiotom łączenie się poprzez sieć z innymi dowolnymi ludźmi lub przedmiotami, niezależnie od czasu czy miejsca[13]. W nieco szerszym ujęciu należy dodać, że chodzi o globalną infrastrukturę, w której rzeczy otrzymują jednostkowe identyfikatory, dzięki którym możliwe jest przesyłanie i identyfikacja danych, takich jak np. czas, lokalizacja, ruch, wilgoć, tętno, które pozwalają na współpracę z innymi systemami. Internet of Things wykorzystuje technologie, takie jak chmura[14], czyli dzięki mocom obliczeniowym serwerów[15] przetwarzane są dane z przedmiotów zaprojektowanych w celu bycia podłączonym do Internetu, lub chip RFID[16] (ang. *Radio-frequency identification*), który umożliwia przesyłanie danych za pomocą fal radiowych.

Przykłady zastosowania Internetu przedmiotów można dostrzec chociażby w transporcie (np. w możliwości śledzenia lotów[17]), logistyce (np. dzięki viatrack[18]) czy zarządzaniu infrastrukturą miejską[19]. Śledzenie aktywności pracowników czy przedmiotów w przemyśle jest jedną z wielu możliwości IoT, które zamiast na skalę masową, w ramach decyzji pracodawcy coraz częściej znajduje zastosowanie w indywidualnych przypadkach.

Użytkownicy smartfonów mają już możliwość pobrania na swój telefon aplikacji, które zbierają wiele szczegółów o naszej osobie. Przykładowo jedną grupę stanowią aplikacje dotyczące naszego zdrowia, a drugą – aplikację do zarządzania naszym domem, czyli służące automatyce domowej (tzw. *smart home*).

Urządzenia z zakresu ochrony zdrowia[20] (*quantifiedself*) wykorzystują możliwości smartfonów lub działają, jeżeli umocuje się je na ubraniu lub na ciele. Zaliczają się do nich wszelkiego rodzaju analizatory ciała (tzw. *Body trackers*), które mogą mierzyć jakość snu (np. *SleepBetter Runtastic*), liczbę kroków (np. *Fitbit*), analizować bieg (np. *runtastic Bieganie i Fitness),* temperaturę (*Body Temperature*). Kolejnym krokiem w rozwoju urządzeń z zakresu zdrowia jest mierzenie tętna (*HeartRate Monitor*) czy ilości składników mineralnych w ciele (*Vitastiq*)[21]. Rozwija się bowiem rynek osobistych czujników, który pozwala z jednej strony

[11] Zob. http://www.forbes.pl/czym-jest-internet-rzeczy-,artykuly,195983,1,1.html (dostęp z 30.3.2017 r.).

[12] IoT (ang. *Internet of Things*) – Internet rzeczy.

[13] *O. Vermesan*, *P. Friess*, Internet of Things Strategic Research Roadmap, Bruksela 2011, s. 12.

[14] *R. Surowiec*, Dane osobowe w chmurach, Rzeczpospolita 2011, Nr 168, s. 23, dostępne również: http://www.rp.pl/artykul/690616-Dane-osobowe--w-chmurach.html (dostęp z 27.3.2017 r.).

[15] *G. Santucci*, Towards Connectobjectome: The age when the totality of all objects become conneted, [w:] *I.G. Smith* (red.), Internet of Things, Halifax 2012, s. 7–11.

[16] *S. Spikermann*, The RFID PIA – developed by industry agreed by regulators, [w:] *D. Wright, P. De Hert* (red.), Privacy Impact Assesment, Berlin––Heidelberg–Nowy Jork 2012, s. 1–22.

[17] Zob. np. https://www.esky.pl/radar; http://lotradar.pl/ (dostęp z 30.3.2017 r.).

[18] Zob. np. http://www.viatrack.pl/index.php/zastosowania/duze-floty.html (dostęp z 30.3.2017 r.).

[19] Zob. aplikację: impk, http://pasazer.mpk.wroc.pl/jak-jezdzimy/mapa--pozycji-pojazdow (dostęp z 30.3.2017 r.).

[20] Więcej o aplikacjach w zakresie ochrony zdrowia: *E.M. Kwiatkowska,* Internet rzeczy. Czy będą nas leczyć komputery?, Internetowy Kwartalnik Antymonopolowy i Regulacyjny 2016, Nr 5, s. 25–27.

[21] *E. Mucha*, Technologie biometryczne, Przegląd Pololicyjny 2015, Nr 2, s. 190–203.

coraz dogłębniej poznać swoje ciało, jego funkcje i odruchy, a z drugiej strony kreuje coraz bardziej wrażliwe dane[22]. Część tych aplikacji niejako „przemyca" zgodę na publikowanie tych danych na portalach społecznościowych lub wymaga do ściągnięcia aplikacji dostępu do dokładnej lokalizacji urządzenia[23]. Natomiast wejście w posiadanie informacji o lokalizacji urządzenia w połączeniu ze zbieranymi danymi pozwala na niemal bezbłędną identyfikację użytkownika. Wydaje się, że regulacja unijna, mimo postulatów o szczególną ochronę danych wrażliwych, nie wdrożyła rozwiązań gwarantujących szczególną ochronę. Pozostawienie tego w gestii państw członkowskich umożliwia natomiast występowanie zjawiska *race to the bottom*[24], które jest niekorzystne dla użytkowników.

Urządzenia z zakresu „*smart home*" pozwalają, aby użytkownik „kontrolował wszystkie urządzenia w domu, od oświetlenia, przez rolety, po wentylację i ogrzewanie. Masz możliwość sterowania z każdego miejsca w budynku i poza nim"[25]– jak reklamuje się jeden z producentów rozwiązań z tego zakresu. Jednocześnie podaje on na swojej stronie bardzo przekonujące dane: 12% mniej zużycia wody rocznie, 24% mniej zużycia energii rocznie. Jednak fakt posiadania tych danych wiąże się z niczym innym jak ze zbieraniem i przetwarzaniem danych o użytkownikach swoich rozwiązań.

Rozważania dotyczące ryzyka dla ochrony prywatności i ochrony danych osobowych rozpocząć należy od kwestii legalności tych zjawisk. Użytkownik, który pobiera aplikację bądź synchronizuje dany czujnik ze swoim smartfonem, najczęściej spełnia chociażby jedną przesłankę, która legalizuje przetwarzanie danych przez twórców aplikacji. Z jednej strony warunkiem koniecznym do pobrania aplikacji jest wyrażenie zgody, a z drugiej strony w pewną wątpliwość poddać można, czy zgoda ta jest świadoma – czy pozwala użytkownikowi pozyskać wiedzę, na czyją rzecz udziela podmiotowej zgody i w jakim celu oraz jakie konkretnie dane będą przetwarzane[26]. Ponadto zgoda[27], jak podkreślają przepisy RODO, powinna być dobrowolna, konkretna, świadoma i jednoznaczna, co w przypadku pobierania aplikacji jest raczej stanem postulatywnym, odbiegającym od rzeczywistości.

Nieco dalej idącym, acz znacznie bardziej spektakularnym zagrożeniem, które nasuwa się przy rozważaniu tych zjawisk, zwłaszcza w przypadku rozwiązań z zakresu „smart home" – jest tzw. atak hakerski. Odzwierciedlenie tych lęków zostało zobrazowane już w popkulturze – np. w drugim sezonie serialu pt. Mr. Robot grupa hakerów, włamując się do systemu i ustawiając ekstremalne parametry w zakresie temperatury oraz aktywując wszelkie urządzenia elektryczne, wypędziła mieszkankę z jej własnego domu[28]. Potencjalne ryzyko związane z gromadzeniem i przetwarzania takich danych ma zatem o wiele dalej idące konsekwencje niż brak kontroli.

W. Wiewiórowski[29] wskazuje, że jako potencjalne ryzyka z tymi zjawiskami możemy uznać:

1) tworzenie i ujawnianie wzorców zachowań użytkowników[30];
2) oceny zachowania użytkownika, tworzenie wzorca normalności oraz stały nadzór mieszczenia się w normach[31];
3) wyciąganie negatywnych konsekwencji ze względu na odbieganie od normy, np. jeżeli zwiększamy stopniowo jasność w swoim domu – oznaczać to może, że psuje nam się wzrok, a taką informację można udostępnić potencjalnym reklamodawcom lub ubezpieczycielowi;
4) brak możliwości udzielania świadomej i dobrowolnej zgody na przesyłanie danych osobowych (wiele osób utożsamia wprowadzenie urządzeń w tryb offline z brakiem przesyłania danych. Wiele urządzeń zapewnia rejestrowanie i przechowywanie danych, które zostają przekazane administratorowi w momencie włączenia do sieci);
5) brak kontroli nad generowanymi danymi – związane głównie z wbudowanymi mikrofonami, chipami, czujnikami, które funkcjonują w sposób niezrozumiały dla użytkowników;
6) automatyczne podejmowanie decyzji przez urządzenia otrzymujące dane od nas;
7) długoterminowe przechowywanie danych;
8) repersonalizacja danych (omówiona poniżej).

Powołując się ponownie na wnioski *W. Wiewiórowskiego*, nie da się nie zgodzić, że kwestie bezpieczeństwa urządzenia oraz komunikacji pomiędzy urządzeniami są w konflikcie z interesami producentów, twórców aplikacji, sponsorów urządzeń lub aplikacji. Dane osobowe pozwalające na tworzenie wzorców, badanie nawyków i zachowań czy na reperso-

[22] *K. Krassowski*, Identyfikacja biometryczna – nasz przyjaciel czy wróg?, Studia Prawnoustrojowe, t. 23, 2014, s. 189–201.

[23] *R. Surowiec*, Dane osobowe…

[24] Zjawisko *race to the bottom* wyjaśnione zostało w piśmie *The Economist* http://www.economist.com/blogs/freeexchange/2013/11/labour-standards, jak również na blogu The Broker, http://www.thebrokeronline.eu/Blogs/Inclusive-Economy-Europe/The-race-to-the-bottom-explained (dostęp z 10.6.2017 r.).

[25] Hasło reklamowe: https://ampio.com.pl/ (dostęp z 30.3.2017 r.).

[26] *A. Dmochowska*, Unijna reforma przepisów ochrony danych osobowych – analiza zmian, Warszawa 2016, s. 19.

[27] Szerzej zob. *J. Kosuniak*, Odwoływalność zgody na przetwarzanie danych – doświadczenia przedsiębiorcy telekomunikacyjnego, [w:] *A. Mednis* (red.), Prywatność a ekonomia. Ochrona danych osobowych w obrocie prawnym, Warszawa 2013, s. 73–79.

[28] Zob. https://qz.com/733269/mr-robot-played-to-our-worst-technology-fears-with-a-mini-horror-movie-about-a-hacked-smart-home/ (dostęp z 30.3.2017 r.).

[29] *W.R. Wiewiórowski*, Ochrona danych osobowych w świecie Internetu przedmiotów, Dodatek do MoP 2014, Nr 9, s. 9.

[30] O tym, jak przejść od profilu cyfrowego do opisu osoby, pisze: *A. Roosendaal*, Digital Personae and Digital Profile as Representations of Individuals, [w:] *M. Bezzi, P. Duquenoy, S. Fischer-Huebner, M. Hansen*, *G. Zhang* (red.), Privacy and Identity Management for Life, Berlin–Heidelberg–Nowy Jork 2010, s. 227–223.

[31] *A. Welsh*, The Identity Theft Protection Guide, Nowy Jork 2004, s. 236–258.

nalizację stają się coraz cenniejszym towarem pozwalającym na wpływanie na zachowanie mas, a tym samym na zyski. W RODO ustanowiono wysokie standardy dotyczące zgody na przetwarzanie danych osobowych. Istotny jest także fakt, że przetwarzanie danych osobowych musi korespondować z wyrażoną zgodą również w aspekcie celowościowym[32]. Niemniej wydawać by się mogło, że system weryfikacji faktycznej korelacji między przetwarzaniem a celem producentów jest utrudniony ze względu na generalność tej normy. Wiele zagrożeń wiążących się ze zbieraniem danych pokazuje, jak daleko idące mogą być konsekwencje dostępu do danych użytkowników – mogą one dotyczyć płaszczyzny zarówno społecznej, jak i finansowej. Użytkownicy w pierwszej kolejności mogą być kategoryzowani jako atrakcyjni konsumenci, a w dalszej wiązać się to może z umożliwieniem dostępu do określonych dóbr, tworząc nową kategorię dóbr luksusowych w oparciu o schematy zachowań. Może to wykreować nowy wymiar kapitalizmu oraz konsumpcji opartej na ograniczeniach zamiast na dostępie. Tymczasem RODO wydaje się nie poruszać tych kwestii, ustanawiając jedynie normy umożliwiające dążenie do stanu postulatywnego, jakim jest ochrona jednostek. Przy ogromie potencjału korzyści ekonomicznych takie regulacje mogą okazać się niewystarczające.

Repersonalizacja danych

Repersonalizacja danych łączy się ściśle z tematem Internetu przedmiotów. Polega ona bowiem na tworzeniu z pozoru anonimowych zbiorów danych osobowych, które jednak przy stałym nadzorze nad osobą mogą prowadzić do tworzenia profili umożliwiających jej identyfikację[33]. Oznacza to, że dane, które zostały zebrane w czasie korzystania z Internetu rzeczy, mogą być przetwarzane w ogromnych zbiorach, a następnie odszyfrowane i przyporządkowane jednostkom.

Punkt 30 preambuły RODO zawiera zapis: „osobom fizycznym mogą zostać przypisane identyfikatory internetowe – takie jak adresy IP, identyfikatory plików cookie – generowane przez ich urządzenia, aplikacje, narzędzia i protokoły, czy też inne identyfikatory, generowane na przykład przez etykiety RFID. Może to skutkować zostawieniem śladów, które w szczególności w połączeniu z unikatowymi identyfikatorami i innymi informacjami uzyskiwanymi przez serwery mogą być wykorzystywane do tworzenia profili i identyfikowania tych osób".

Istnieją dwa podstawowe podejścia do tworzenia profili użytkowników, a mianowicie:

– profile predykcyjne[34] – tworzy się w drodze wnioskowania na podstawie obserwacji indywidualnego i zbiorowego zachowania użytkowników w czasie, w szczególności przez monitorowanie odwiedzanych stron oraz reklam, które użytkownik wyświetla lub na które klika. Bazując na tych informacjach próbuje się przewidzieć zachowanie takiej osoby;
– profile jawne – zestawy informacji dotyczące osób zebrane z różnych źródeł. Tworzy się na podstawie danych osobowych przekazywanych w ramach usługi sieciowej przez same osoby, których dane dotyczą, np. podczas rejestracji. Wspomniane metody można łączyć. Ponadto profile predykcyjne mogą stać się jawne później – kiedy osoba, której dotyczą dane, utworzy dane logowania dla danej strony internetowej[35].

Mimo że dane takie najczęściej są poddawane procesowi anonimizacji, wskazać należy potencjalne skutki repersonalizowania danych. *Y.-A. de Montjoye* wraz z grupą naukowców z MIT-u (*A. „Sandy" Pentland, L. Radaelli* i *V.K. Singh*)[36] przez trzy miesiące przyglądali się takim anonimizowanym danym dotyczącym kart kredytowych 1,1 mln osób z nieujawnionego kraju. W danych nie było ani nazwisk posiadaczy kart, ani informacji o kontach bankowych, z którymi karty są powiązane. Naukowcy chcieli sprawdzić, co można powiedzieć o ludziach na podstawie takich właśnie informacji. Okazało się, że całkiem sporo. Uczonym wystarczyły cztery różne fragmenty informacji, by zidentyfikować 90% posiadaczy kart płatniczych. Wykazano, że znajomość ceny produktu zwiększa ryzyko reidentyfikacji o 22%[37]. Wystarczyło połączyć dane z serwisów społecznościowych, na których użytkownicy oznaczali swoją obecność w restauracji lub chwalili się nowym ubraniem, z sumami transakcji z pobliskich miejsc, aby ustalić, kto jest posiadaczem karty płatniczej[38]. W innym eksperymencie *Y.-A. de Montjoye* wraz z *C. Hidalgo*, *V. Blondelem* i *M. Verleysenem* użyli 15-miesięcznych danych z 1,5 mln osób, aby pokazać, że cztery punkty – przybliżone miejsca i godziny – wystarczą do identyfikacji 95% osób fizycznych w bazie danych dotyczących mobilności. Opracowali formułę służącą do oszacowania niepowtarzalności

[32] Jak również legalności, celowości, adekwatności, czasowości, integralności i poufności danych. Szerzej zob. *A. Dmochowska*, Unijna reforma…, s. 11–17.

[33] *W.R. Wiewiórski*, Ochrona danych osobowych…, s. 10.

[34] Zob. http://www.computerworld.pl/news/395648/VII-Forum-Bezpieczenstwa-i-Audytu-IT-GIODO-o-analizie-predykcyjnej-i-profilowaniu.html (dostęp z 30.3.2017 r.).

[35] Zob. Opinia Grupy art. 29 Nr 2/2010 w sprawie internetowej reklamy behawioralnej przyjęta 22.6.2010 r., pkt 2.3., s. 8, https://piu.org.pl/public/upload/ibrowser/seminaria/Jakosc%20Danych%20IX/GIODO_W-Wiewiorowski_ZASADA_CELIOWOSCI_DATA_MINING.pdf, slajd 16 (dostęp z 30.3.2017 r.). O procesie zbierania danych o osobach fizycznych również: *K. Markowski,* Prywatność czy rzetelność obrotu gospodarczego?, [w:] *A. Mednis* (red.), Prywatność a ekonomia. Ochrona danych osobowych w obrocie prawnym, Warszawa 2013, s. 41–46.

[36] Zob. http://demontjoye.com/projects.html – strona *Y.-A. de Montjoye*; project: Unique in the Shopping Mall: Reidentifying credit card data (dostęp z 30.3.2017 r.).

[37] Zob. https://www.media.mit.edu/projects/on-the-reidentifiability-of-credit-card-metadata/overview/ (dostęp z 30.3.2017 r.).

[38] Zob. http://kopalniawiedzy.pl/dane-osobowe-karta-platnicza-anonimowosc,21822 (dostęp z 30.3.2017 r.).

śladów mobilności człowieka i wskazali, że nawet wtedy gdy dane są niskiej jakości, nie zapewniają one anonimowości[39].

Inni badacze – z University of Birmigham – dowiedli, że dzięki danym GPS zebranym w New Hampshire i taksówkach San Francisco są w stanie rozpoznać niemal 100% posiadaczy smartfonów[40].

W opinii 5/2014 w sprawie technik anonimizacji przyjętej przez Grupę Roboczą 10.4.2014 r. wskazuje się, że „wymagane jest zachowanie szczególnej ostrożności w postępowaniu ze zanonimizowanymi informacjami, zwłaszcza w każdym przypadku, gdy takie informacje wykorzystuje się (często w połączeniu z innymi danymi) do celów podejmowania decyzji". Jednocześnie w rozdziale pt. Analiza techniczna, niezawodność technologii i typowe błędy wskazuje się następujące zagrożenia:

„1) wyodrębnienie, które oznacza możliwość wydzielenia niektórych lub wszystkich zapisów identyfikujących określoną osobę fizyczną w zbiorze danych;
2) możliwość tworzenia powiązań, czyli zdolność do powiązania co najmniej dwóch zapisów dotyczących jednej osoby lub grupy osób, których dane dotyczą (w tej samej bazie danych lub w dwóch różnych bazach danych). Jeżeli atakujący może ustalić (np. w drodze analizy korelacji), że dwa zapisy przypisane są tej samej grupie osób fizycznych, ale nie może wyodrębnić poszczególnych osób w tej grupie, dana technika zapewnia ochronę przed wyodrębnieniem, ale nie przed możliwością tworzenia powiązań;
3) wnioskowanie, czyli możliwość wydedukowania ze znacznym prawdopodobieństwem wartości danego atrybutu z wartości zbioru innych atrybutów".

Grupa Robocza, podobnie jak *W. Wiewiórowski* w przypadku zagrożeń wynikających z Internetu rzeczy, wskazuje wiele konsekwencji wynikających ze zjawiska repersonalizacji. Identyfikacja osoby w grupie stanowi podstawę do określenia wzorców jej zachowania. Wskazuje również, że wdrażanie systemów zabezpieczeń, które chronią jedynie częściowo rekordy znajdujące się w bazach danych, doprowadzić może do dogłębnej analizy śladów pozostawianych przez użytkowników sieci[41].

Rozporządzenie 2016/679 nie reguluje *stricto* repersonalizacji danych. Natomiast w motywie 26 preambuły znaleźć można zapis: „zasady ochrony danych powinny mieć zastosowanie do wszelkich informacji o zidentyfikowanych lub możliwych do zidentyfikowania osobach fizycznych. Spseudonimizowane dane osobowe, które przy użyciu dodatkowych informacji można przypisać osobie fizycznej, należy uznać za informacje o możliwej do zidentyfikowania osobie fizycznej. Aby stwierdzić, czy dana osoba fizyczna jest możliwa do zidentyfikowania, należy wziąć pod uwagę wszelkie rozsądnie prawdopodobne sposoby (w tym wyodrębnienie wpisów dotyczących tej samej osoby), w stosunku do których istnieje uzasadnione prawdopodobieństwo, iż zostaną wykorzystane przez administratora lub inną osobę w celu bezpośredniego lub pośredniego zidentyfikowania osoby fizycznej. Aby stwierdzić, czy dany sposób może być z uzasadnionym prawdopodobieństwem wykorzystany do zidentyfikowania danej osoby, należy wziąć pod uwagę wszelkie obiektywne czynniki, takie jak koszt i czas potrzebne do jej zidentyfikowania, oraz uwzględnić technologię dostępną w momencie przetwarzania danych, jak i postęp technologiczny. Zasady ochrony danych nie powinny więc mieć zastosowania do informacji anonimowych, czyli informacji, które nie wiążą się ze zidentyfikowaną lub możliwą do zidentyfikowania osobą fizyczną, ani do danych osobowych zanonimizowanych w taki sposób, że osób, których dane dotyczą, w ogóle nie można zidentyfikować lub już nie można zidentyfikować".

Zapis ten daje podstawę do stosowania zasad legalizacji danych, które mają być anonimizowane, a później repersonalizowane. Norma ta zwraca uwagę na konieczność wzięcia pod uwagę zarówno postępu technologicznego, jaki nastąpił do momentu identyfikowania, jak i dostępną technikę[42]. Skutkiem tego regulacja ta buduje standardy i zakres ochrony niezależne od wiedzy prawodawcy, przenosząc ciężar potencjalnego dowodu w zakresie wiedzy na osoby depersonalizujące dane. Z drugiej strony, ze względu na dynamikę zmian oraz fakt, że informacje o sposobach depersonalizacji nie są powszechne, nieostrość regulacji zaciera granice ochrony.

Warto również wskazać, że repersonalizacja danych niesie ze sobą wiele różnych zagrożeń, niezależnie czy dane zostaną zdeanimizowane czy też nie. Obserwując pewne wzorce zachowań ludzkich, przedsiębiorcy mogą wykorzystywać nawyki ludzi, np. w sposobie odbierania treści ze stron internetowych, aby uniemożliwić lub znacznie utrudnić użytkownikom Internetu zapoznawanie się z konkretnymi treściami, poprzez odpowiednie rozmieszczenie informacji. Tworzenie wzorców zachowań pewnych grup społecznych może także doprowadzić do ograniczania możliwości wyboru osób fizycznych w ramach ich decyzji konsumenckich wyłącznie do ich dotychczasowych decyzji. Użytkownicy mogą być klasyfikowani, w zależności od swoich wyborów, przyporządkowani do konkretnych grup opartych na zbiorowych trendach i zamykani w świecie definiowanym, przez wybory z konkretnego, badanego odcinka czasu. Odpowiedzią na

[39] The privacy bounds of human mobility, http://demontjoye.com/projects.html (dostęp z 30.3.2017 r.).

[40] Zob. https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-015-0049-x oraz http://dl.acm.org/citation.cfm?doid=2426656.2426666 (dostęp z 30.3.2017 r.).

[41] Tzw. *browse fingerprints* czyli odciski palców w sieci, polegająca na anonimowym identyfikowaniu przeglądarki internetowej z dokładnością do 94%, szeroko można poczytać na technicznych stronach internetowych, np. https://github.com/Valve/fingerprintjs (dostęp z 18.11.2017 r.).

[42] Obecnie powstały już programy służące do depersonalizacji danych, dedykowane dla programów obsługujących migrację danych, np. http://ssisctc.codeplex.com/, dla programu SSIS (dostęp z 2.6.2017 r.).

anomalię użytkownika może być natomiast deanonimizacja danych, a następnie automatyczna korekta prezentowanych jednostce treści, dla danej jednostki, w oparciu o jej konkretne, zdeanimizowane już zachowania.

Rozporządzenie 2016/679 a Internet rzeczy i repersonalizacja danych

Powyżej opisane zostały potencjalne skutki oraz zagrożenia związane z operacjami na danych osobowych, które obiegają od klasycznego rozumienia ich przetwarzania. Rozporządzenie 2016/679 w motywie 39 i 4 preambuły przewiduje, że wszelkie przetwarzanie danych osobowych powinno być zgodne z prawem i rzetelne oraz że przetwarzanie danych osobowych należy zorganizować w taki sposób, aby służyło ludzkości. Faktem jednak jest, że zjawisko Internetu rzeczy i repersonalizacji danych nie zostało bezpośrednio uregulowane w RODO, stąd ustalenie granic legalności nasuwać może pewne problemy. Jak słusznie wskazuje *D. Lubasz*, RODO celowo nie definiuje nawet samego przetwarzania danych w sposób częściowo lub całkowicie zautomatyzowanych, ze względu na wprowadzenie neutralności technologicznej[43]. Właśnie gwałtowny postęp technologiczny wymusił wprowadzenie wielu klauzul generalnych, które pozwalają znaleźć punkty odniesienia do legalnych aspektów omawianych zjawisk, które interpretować trzeba z uwzględnieniem celów ustawodawcy. Same zapisy jednak nie dają jasnych odpowiedzi, jakie zabezpieczenia należy wprowadzić, żeby chronić jednostki, które nieświadomie mogą stać się podmiotem operacji na danych[44].

Ochrona osób możliwa jest przy konkretnym określeniu zagrożeń. Rozporządzenie 2016/679 odnosi się do ryzyka naruszenia praw lub wolności osób. W motywach od 75 do 81 preambuły możemy znaleźć szereg zapisów dotyczących tego zakresu. Ustawodawca unijny wskazuje, że naruszenie takie może prowadzić do uszczerbku fizycznego lub szkód majątkowych lub niemajątkowych. Za kluczowy w związku z przetwarzaniem danych osobowych wskazać należy motyw 75 preambuły, który zawiera katalog otwarty zagrożeń. Zawiera on w szczególności skutki, kiedy „przetwarzanie może poskutkować dyskryminacją, kradzieżą tożsamości lub oszustwem dotyczącym tożsamości, stratą finansową, naruszeniem dobrego imienia, naruszeniem poufności danych osobowych chronionych tajemnicą zawodową, nieuprawnionym odwróceniem pseudonimizacji lub wszelką inną znaczną szkodą gospodarczą lub społeczną; jeżeli osoby, których dane dotyczą, mogą zostać pozbawione przysługujących im praw i wolności lub możliwości sprawowania kontroli nad swoimi danymi osobowymi; jeżeli przetwarzane są dane osobowe ujawniające pochodzenie rasowe lub etniczne, poglądy polityczne, wyznanie lub przekonania światopoglądowe, lub przynależność do związków zawodowych oraz jeżeli przetwarzane są dane genetyczne, dane dotyczące zdrowia lub dane dotyczące seksualności lub wyroków skazujących i naruszeń prawa lub związanych z tym środków bezpieczeństwa; jeżeli oceniane są czynniki osobowe, w szczególności analizowane lub prognozowane aspekty dotyczące efektów pracy, sytuacji ekonomicznej, zdrowia, osobistych preferencji lub zainteresowań, wiarygodności lub zachowania, lokalizacji lub przemieszczania się – w celu tworzenia lub wykorzystywania profili osobistych". Warto również dodać, że RODO wprowadza również definicję legalną naruszenia ochrony danych osobowych, którą należy rozumieć jako naruszenie bezpieczeństwa prowadzące do przypadkowego lub niezgodnego z prawem zniszczenia, utracenia, zmodyfikowania, nieuprawnionego ujawnienia lub nieuprawnionego dostępu do danych przesyłanych, przechowywanych lub w inny sposób przetwarzanych[45].

Stworzenie tego katalogu oraz definicji konieczne było w celu nałożenia na osoby odpowiedzialne za przetwarzanie danych obowiązku oceny skutków oraz minimalizowania ryzyka. Jako kluczowe należy wskazać wdrożenie odpowiednich środków technicznych i organizacyjnych, RODO jednak przewiduje wiele innych wymogów[16] koniecznych do zastosowania w celu przetwarzania danych osobowych oraz w przypadku naruszeń[47]. Odpowiedzialność za przetwarzanie zgodne z prawem, rzetelne i przejrzyste dla osoby, której dane dotyczą[48], możliwe jest przy pełnym zrozumieniu mechanizmów funkcjonowania przetwarzania danych dla osób, których dane dotyczą. Potwierdza to ustawodawca, który wskazuje, że w motywie 60 preambuły zasady rzetelnego i przejrzystego przetwarzania wymagają, by osoba,

[43] *E. Bielak-Jomaa, D. Lubasz* (red.), RODO. Ogólne rozporządzenie o ochronie danych. Komentarz, Warszawa 2017, *passim*.

[44] Konsumenci często nie zdają sobie sprawy z potencjalnych zagrożeń związanych z przetwarzaniem danych osobowych. Powstał już sklep, w których płacić można danymi osobowymi: https://www.engadget.com/2017/09/07/data-dollar-store-london-ben-eine/czy https://www.lonelyplanet.com/news/2017/08/31/pay-personal-data-london-street-art-pop/(dostęp z 19.11.2017 r.). Zjawisku przyjrzał się już *P. Hustinx* w opinii z 26.3.2014 r., https://edps.europa.eu/sites/edp/files/publication/14-07-14_ph_for_ev_online_en.pdf (dostęp z 19.11.2017 r.). O tworzeniu waluty z danych osobowych zob. np.: https://www.kaspersky.com/about/press-releases/2017_data-dollar-the-new-currency-based-on-the-value-of-personal-data (dostęp z 19.11.2017 r.).

[45] Definicja legalna „naruszenia danych osobowych" znajduje się w art. 4 pkt 12 RODO.

[46] Szerzej *D. Lubacz*, Europejska reforma ochrony danych osobowych – nowe obowiązki administratorów w ogólnym rozporządzeniu o ochronie danych, [w:] *E. Bielak-Jomaa*, *D. Lubasz* (red.), Polska i europejska reforma ochrony danych osobowych, Warszawa 2016, s. 63–85.

[47] Regulacja zawarta jest w art. 34–36 RODO.

[48] Zasady dotyczące przetwarzania danych osobowych znajdują się w art. 5 RODO – szerzej np. *B. Kaczmarek-Templin,* Podstawy legalizacyjne przetwarzania danych osobowych w ogólnym rozporządzeniu o ochronie danych – wybrane zagadnienia, [w:] *E. Bielak-Jomaa, D. Lubasz* (red.), Polska i europejska reforma…, s. 102–126.

której dane dotyczą, była informowana o prowadzeniu operacji przetwarzania i o jej celach. Tym samym ustawodawca stwarza podstawy do uświadomienia jednostek o operacjach, które przeprowadza się na ich danych. Motyw 60 preambuły RODO stanowi, że każda osoba fizyczna powinna mieć prawo dostępu do zebranych danych jej dotyczących oraz możliwość łatwego wykonywania tego prawa w rozsądnych odstępach czasu, by mieć świadomość przetwarzania i móc zweryfikować zgodność przetwarzania z prawem. Analiza tego zapisu wraz z regulacją z motywu 59 preambuły stanowiącym, że należy przewidzieć procedury ułatwiające osobie, której dane dotyczą, wykonywanie praw przysługujących jej na mocy niniejszego rozporządzenia, w tym mechanizmy żądania i, gdy ma to zastosowanie, bezpłatnego uzyskiwania w szczególności dostępu do danych osobowych i ich sprostowania lub usunięcia oraz możliwości wykonywania prawa do sprzeciwu, daje obraz ram, jakie ustawodawca przewidział na ochronę jednostek[49]. Tak sformułowane normy dotyczące częstotliwości oraz łatwości dostępu i sprostowania, usunięcia czy sprzeciwu dotyczących danych osobowych, a także konieczność wprowadzenia procedur wydawać się mogą wystarczające w przypadku aspektów ujętych w RODO (np. przetwarzania danych w celach marketingowych). Natomiast w przypadku zbierania danych podczas korzystania z Internetu rzeczy oraz repersonalizacji danych, których skutkiem może być stworzenie dokładnych profili dotyczących niemalże każdego aspektu życia jednostki, można postulować o nieco dalej idące, bardziej szczególne zapisy. O wiele łatwiej w przypadku tych zjawisk o naruszenie ochrony danych osobowych, a wykorzystanie danych zgromadzonych w ten sposób może być o wiele korzystniejsze finansowo[50].

Postulat *delege ferenda* dla ustawodawcy unijnego

Ewolucja technologiczna, w tym rozwój Internetu wiąże się z licznymi wyzwaniami oraz konfliktami równorzędnych dóbr. Gromadzenie danych oraz przetwarzanie ich na szeroką skalę przez przedsiębiorców doprowadzić może do pewnych zmian w zakresie stosunków społeczno-gospodarczych, które wpłyną bezpośrednio na każdą jednostkę. Zgoda ustawodawcy na pewne zachowania przedsiębiorców oraz innych podmiotów czerpiących z przetwarzania danych osobowych korzyści finansowe wiąże się w pewnym stopniu z erozją praw człowieka, praw jednostek. Warto jednak wprowadzać rozwiązania, które stanowią kompromis pomiędzy interesami zarówno osób fizycznych, jak i przedsiębiorców. Wydawać by się mogło, że pewnym rozwiązaniem pozwalającym z jednej strony na rozwój działalności gospodarczej, a z drugiej na ochronę jednostek jest wprowadzenie pewnej formy kontroli społecznej sprawowanej wyłącznie przez zainteresowane osoby. Pełen dostęp w postaci *real time*[51] podglądu do przekazywanych konkretnym przedsiębiorcom informacji oraz sposób przetwarzania danych (np. przekazywanie je podmiotom trzecim) wraz z możliwością usunięcia części z nich może stanowić odpowiedź na pewien zakres przedstawionych problemów. Z jednej strony pozwoliłoby to zbudować świadomość społeczną oraz osób fizycznych o tym, że dane są gromadzone, a zgoda wcześniej wyrażona wiązałaby się z o wiele dalej idącą wiedzą o samym przetwarzaniu danych niż tą, która idzie za np. zaznaczeniem wymaganego okienka zgody w formularzach internetowych czy przy instalowaniu aplikacji. Z drugiej strony jednostki będą miały faktyczny wpływ na zakres informacji, na podstawie których będą otrzymywać spersonalizowane komunikaty lub oferty, co może stanowić dla użytkowników atrakcyjne rozwiązanie. Sami przedsiębiorcy zaś będą mogli budować swoje oferty oraz dostosowywać się do zachowań konsumentów przy ich faktycznej, a nie jedynie domniemanej zgodzie. Pozwoliłoby to również włączyć do tworzenia profili, na podstawie zbieranych informacji „czynnik ludzki" w postaci samych zainteresowanych, którzy mogliby, w oparciu o dostęp w poszczególnych aplikacjach, wpływać na informacje o nich gromadzone, a z drugiej strony strzec swojej prywatności. Wprowadzenie wymogu udostępnienia danych użytkownika nie stanowi też postulatu obciążającego finansowo przedsiębiorców w nadmiernym stopniu.

Podsumowanie

Niewątpliwie problematyka ochrony danych osobowych jest zagadnieniem niezwykle złożonym. Rozwiązania wprowadzone przez RODO pozwalają w pewnym stopniu pogodzić ochronę praw człowieka wraz z interesami osób czerpiących korzyści finansowe z operacji związanych z danymi osobowymi. Ze względu na rozwój technologiczny brak jest często dostatecznego zbadania zjawisk mających miejsce w Internecie oraz ich konsekwencji, a niezdefiniowanie problemu utrudnia postawienie prawnych norm oraz granic dopuszczalnego zachowania. Nie można nie zgodzić się, że

[49] Szerzej zob. *J. Kosuniak*, Odwoływalność zgody…, s. 73–79; *P. Litwiński*, Korporacyjne systemy raportowania nadużyć (whistle blowing hotlines) a ochrona danych osobowych, [w:] *A. Mednis* (red.), Prywatność a ekonomia. Ochrona danych osobowych w obrocie prawnym, Warszawa 2013, s. 113–127.

[50] O handlu danymi osobowymi np.: http://www.telegraph.co.uk/technology/news/12012191/How-much-is-your-personal-data-worth.html, (dostęp z 19.11.2017 r.); http://cyberprzestepczosc.info/handel-danymi-osobowymi/ (dostęp z 19.11.2017 r.)

[51] *Real time* [ang.] – system czasu rzeczywistego. System, w których odpowiedzi na zapytania urządzeń technicznych zależne są od chwili wypracowanego wyniku. Do istoty systemu należy równoległość w czasie zmian w środowisku oraz obliczeń realizowanych na podstawie stanu środowiska. Środowiskiem w powyższym kontekście nazywana jest baza danych lub jej część, zawierająca gromadzone dane użytkowników.

kwestie dotyczące Internetu rzeczy oraz repersonalizacji danych nie zostały wyczerpane przez ustawodawcę unijnego. Te dwa zjawiska posiadają potencjał usprawnienia funkcjonowania jednostek we współczesnym świecie, jednak niosą ze sobą zagrożenia, które mogą doprowadzić w dalszej perspektywie nawet do zmian społeczno-gospodarczych. Konieczne wydaje się obserwowanie tych zjawisk oraz reagowanie, zarówno legislacyjnie, jak i w dużej mierze dzięki orzecznictwu i doktrynie, które będą rozwijać generalne normy i normy celowościowe zawarte w RODO.

Słowa kluczowe: Internet rzeczy, IoT, repersonalizacja danych, RODO, rozporządzenie o ochronie danych osobowych, rozporządzenie 2016/679, ogólne rozporządzenie o ochronie danych, nowe technologie, aplikacje, deanonimizacja, anonimizacja danych, personalizacja danych.

Personal data protection and the Internet of Things, profiling and re-personalization of data

The aim of the article is describing the Internet of Things and re-personalization of data with effects and threats which they may cause in the context of protection of the data. The author of the article compares them to the regulation included in the Regulation 2016/679 of 27 April 2016 of the European Parliament and the Council of the European Union on the protection of legal persons with regard to the processing of personal data and on the free movement of such data, and repealing the Directive 95/46/EC (General Data Protection Regulation).

Key words: Internet of Things, IoT, re-personalization of data, Regulation (EU) 2016/679, General Data Protection Regulation, new technologies, application, anonymization of data, de-anonymization, personalization of data.