Rozdział I. Założenia wstępne dla zrównoważonego przetwarzania informacji ze źródeł publicznych w czasach *big data*

Wojciech Rafał Wiewiórowski

§ 1. Przetwarzanie informacji w świecie *big data*

Nie ma żadnej przyjętej powszechnie ani nawet akceptowanej szeroko definicji pojęcia *big data*. Wszystkie popularne wyjaśnienia bazują na pojęciach niedookreślonych i ocennych. W zasadzie można przyjąć, że od początku ery społeczeństwa informacyjnego – niezależnie od tego, jak daleko w przeszłość cofniemy historię tego równie niedookreślonego pojęcia – stawiano tezę, że informacji jest za dużo, by człowiek lub nawet grupa ludzi rozumnie ją przetwarzali. Muszą posługiwać się w tym celu urządzeniami, które przetwarzają informację tak, jak ludzie je do tego przygotowali, ale również w sposób od owej woli ludzkiej niezależny. Od co najmniej dziesiątek lat mówi się o swoistej samoświadomości urządzeń (sieci, systemów), choć oczywiście sposób opisywania takiej „samowiedzy sieci" był zupełnie różny w czasach przed Alta Vistą, w czasach Alta Visty i w czasach dzisiejszych, opisywanych jako era Google'a[1].

W tej sytuacji jedną z dróg określenia, czym różni się dzisiejsze *big data* od innych obserwowanych przez lata fenomenów w przetwarzaniu wielkoskalowych zbiorów danych, jest stwierdzenie, że w przypadku *big data* liczy się nie tyle skala dostępnych zasobów, co łatwość ich łączenia z nowymi zasobami nieznanymi w momencie, gdy określamy pierwotny cel i sposoby przetwarzania danych. W tym przypadku bowiem nie zakładamy od początku, że lista

[1] *J. Gleick*, Informacja, s. 390–392.

możliwych do osiągnięcia celów i zasobów jest znana w momencie ustanawiania celu początkowego, lecz przyjmujemy, że mogą się one zmieniać w zależności od cech merytorycznych ewentualnych zasobów odnalezionych w przyszłości[2]. Oznacza to tym samym, że przy operacjach na zbiorach, określanych pojęciem *big data*, nie mamy zamiaru równoważenia celów i sposobów przetwarzania danych, a możemy co najwyżej – choć chyba bardzo rzadko w ogóle przychodzi to do głowy przetwarzającemu – oceniać zrównoważenie w trakcie działania.

Poszukując wspólnego mianownika w formułowanych na całym świecie definicjach, można stwierdzić, że termin *big data* – niekiedy tłumaczony na język polski jako „moc danych"[3] – odnosi się do praktyki łączenia dużych zasobów informacyjnych, obejmujących zróżnicowane i pierwotnie rozproszone dane, oraz do analizowania tychże zasobów przy pomocy rozbudowanych algorytmów, celem wydobycia zeń informacji, która może być wykorzystana do celowego działania. Termin odnosi się nie tylko do rozwiązań technologicznych służących do zbierania i przechowywania dużych ilości danych, lecz również – a może przede wszystkim – do analizy, rozumienia i skutecznego wykorzystania pełnej wartości niesionej przez dane. Szczególny nacisk kładziony jest w tych procesach na automatyczne przetwarzanie danych przy pomocy specjalnie stworzonych do tego aplikacji. W założeniu „ewangelistów" analiza *big data* powinna prowadzić do szybszego i łatwiejszego podejmowania trafniejszych decyzji, lepiej dostosowanych do potrzeb ekonomicznych i społecznych. Najczęściej przytaczanym przykładem obszaru wykorzystania takich wielkich zasobów danych i przetwarzających je algorytmów jest nauka, ze szczególnym uwzględnieniem medycyny i farmacji[4].

Poprawne wykorzystanie w tym zakresie zasobów już istniejących i nowo tworzonych prowadzić powinno nie tylko do powstania nowych generalnych rozwiązań, ale też umożliwiać indywidualizację zastosowania wyników badań naukowych, lepiej dostosowując podejmowane działania – w tym leczenie – do potrzeb indywidualnego pacjenta. Podkreśla się tym samym humanistyczne aspekty rozwiązań *big data* skierowanych ostatecznie nie do „masy" bezimiennych danych, lecz do jednostki mogącej z nich skorzystać.

---

[2] O mało udanych próbach wsparcia „wielkiej fuzji" oraz o tradycyjnych wątpliwościach co do pojemności sieci *M. Castells*, Społeczeństwo sieci, s. 393–404.

[3] Tak np. w tytule XXI Forum Teleinformatyki „Moc danych – nowe źródła i nowe metody analizy i ochrony danych", zorganizowanego w Warszawie w dniach 24–25.9.2015 r.

[4] *M. Ebeling*, Healthcare and Big Data, s. 27–48; *M. Ford*, Świt robotów, s. 157–158; *D. DiEuliis*, *J. Giordano*, Neurotechnological Convergence, s. 71–80.

Informacja przetwarzana przy pomocy aplikacji korzystających z rozwiązań *big data* w większości nie jest spersonalizowana *per se.* Dane generowane przez sensory badające zjawiska naturalne i pogodowe czy zanieczyszczenia środowiska bądź monitorujące techniczne aspekty działania urządzeń z zasady nie odnoszą się do zidentyfikowanych lub możliwych do zidentyfikowania osób fizycznych. Prawdą jest jednak, że największą wartość, tak w działaniach gospodarczych, jak i zorientowanych na efekty społeczne, niesie monitowanie zachowania ludzi w grupie i indywidualnie. Efekty takiego monitorowania mogą być skutecznie wykorzystane do określenia działań predykcyjnych przewidujących, jakie potrzeby – życiowe, społeczne i gospodarcze – należało będzie zrealizować w przyszłości.

W niniejszej pracy całkowicie pomijany jest problem sprzecznego z prawem uzyskania dostępu do danych. Wszystkie poniższe rozważania przeprowadzono przy założeniu, że wszystkie przetwarzane przez administratora dane zostały uzyskane zgodnie z prawem, a administrator ma prawo przetwarzać dane. Techniki zbierania i analizowania ogromnych zasobów danych mają bowiem z założenia pomagać nam oceniać świat, nawet jeśli na początku tego procesu nie wiemy, jakie cele w tym „ocenianiu świata" mamy zamiar osiągnąć.

Prawdziwą rewolucją jest więc nie tyle technika sama w sobie, a dostępne sposoby jej użycia. Podstawowym celem zdaje się jednak być predykcja zdarzeń w drodze wnioskowania na podstawie obserwacji indywidualnego i zbiorowego zachowania użytkowników w czasie. Oznacza to, że będziemy oceniać nieznany z początku zasób, pod kątem nieznanych z początku cech, by osiągnąć nieznane z początku rezultaty.

Z pozoru nie różni się to od założeń przyjmowanych przez wielu naukowców na początku każdych studiów empirycznych. Nie do końca wiemy, co mamy, nie do końca wiemy jak „to" będziemy przetwarzać, nie wiemy też, jaki będzie efekt. Różnicą w przypadku przetwarzania „zasobu", jakim jest informacja (a nie np. ruda miedzi), jest charakter tego zasobu. Nie jest to bowiem zasób niezależny od obserwatora. Jest natomiast w sposób bardzo podstawowy powiązany z prawami osób, których może dotyczyć.

Dodatkowo, rozważając problem zrównoważonego przetwarzania informacji i zrównoważonego zarządzania nią, nie możemy zapominać, że w przy-

padku *big data* – nie licząc znaczenia, jakie ma ogromny rozmiar zasobu – odchodzimy od dwóch założeń, jakie mają empiryczne badania naukowe:

1) nie twierdzimy, że wyniki będą w całości poprawne, zastępując to założenie predykcją ewentualnych korelacji;
2) odchodzimy od wyników jednostkowych na rzecz masowego prawdopodobieństwa poprawności informacji wyjściowej.

W tej sytuacji dla jakiegokolwiek myślenia o zrównoważeniu przetwarzania informacji musimy wyraźnie oddzielać informację wejściową od wyjściowej. Tylko to bowiem umożliwi nam – lub przyszłym badaczom tego samego zasobu – powrócenie „do źródeł" celem odtworzenia, od czego zaczęliśmy i umożliwienia nam zastosowania innych technik niż te, których dotychczas używaliśmy, do dalszych prac nad zasobem informacyjnym. Jest to szczególnie trudne w przypadku *big data* z dwóch podstawowych powodów. Po pierwsze, informacja wejściowa „dochodzi" do nas w trakcie przetwarzania. Następuje to tak w związku z uzupełnianiem zasobów o nową informację, której oczekiwaliśmy – np. o takie same dane, jak dotychczas zebrane, lecz pochodzące z kolejnych okresów rozliczeniowych – jak i o nowe dane, o których włączeniu do badań zadecydowaliśmy później. Po drugie zaś, przy przetwarzaniu danych przy pomocy technik *big data* informacja wytworzona w czasie przetwarzania – czyli oczekiwana wartość dodana badań – staje się natychmiast informacją wejściową dla kolejnych analiz[5]. W tej sytuacji tworzy się swego rodzaju oscylator informacji, w którym informacja wyjściowa jednego procesu staje się natychmiast informacją wejściową nowego procesu, utrudniając tym samym powrót do źródłowego zasobu.

§ 2. Informacja i jej cechy

Nim przejdziemy do kolejnego etapu rozważań nad zrównoważeniem przetwarzania informacji, warto zatrzymać się nad samą informacją i jej cechami, które będą wpływać na ocenę możliwości jakiegokolwiek zarządzania informacją w dużym przedziale czasu.

Podstawowym celem, jaki przyświeca wszystkim podmiotom, które chcą wykorzystywać informację dla celów innych niż te, dla których pierwotnie została ona zgromadzona, jest przetworzenie jej i zastosowanie w jak najbardziej

[5] O wartości danych na kolejnych etapach przetwarzania *V. Meyer-Schoenberger, K. Cukier*, Big Data, s. 137 i 146.

obiektywnej formie. Takie działanie zaczyna się *de facto* od rozbicia informacji na swoiste kwanty informacyjne (niekiedy odpowiadające pojęciu danych).

Nie wydaje się możliwe stworzenie jednej uniwersalnej definicji pojęcia informacji. Termin ten bywa używany w różnych znaczeniach, zależnie od kontekstu. Rozumiany jest zazwyczaj intuicyjnie, a wszelkie próby definiowania mogą służyć jedynie przedstawieniu poglądu danego użytkownika[6]. Nawet sam użytkownik owego pojęcia może posługiwać się nim w zdecydowanie różnych znaczeniach. Intuicyjnie przyjmujemy, że informację stanowi każdy czynnik, który zwiększa naszą wiedzę o otaczającej rzeczywistości[7]. W matematycznej teorii informacji oraz w cybernetyce informacja jest jednym z terminów pierwotnych, czyli wyrażeń przyjętych bez potrzeby ich definiowania. W obu dziedzinach informacja uważana jest za termin semantycznie prosty, służący do definiowania innych, nazywanych terminami złożonymi[8].

W najszerszym ujęciu – stosowanym przez cybernetyków, jeśli zostaną zmuszeni do definiowania pojęcia z założenia mającego cechę pierwotności – za informację uznaje się każdy czynnik przekazany człowiekowi lub urządzeniu automatycznemu, niezależnie od treści, sposobu przekazania, utrwalania i przechowywania, który może zostać wykorzystany dla bardziej sprecyzowanego, celowego działania. Można więc za *G. Szpor* powiedzieć, że informacją jest wszystko, co zmniejsza entropię[9].

Wyjątkowo często pojawiającym się błędem jest zamienne stosowanie terminów „dane" i „informacje". O ile w wielu dziedzinach wiedzy można utożsamiać posiadanie danych z posiadaniem informacji, o tyle, gdy mówimy o nowych technologiach informacyjnych, nie możemy zapominać, że w informatyce pojęcia te – choć bliskoznaczne – nigdy nie stanowią synonimów. Dane są wartościami przechowywanymi w bazie, rozumianymi wręcz jako wartości danego pola[10]. W przypadku rejestrów publicznych prowadzonych w postaci elektronicznej są to poszczególne atrybuty w relacyjnej bazie danych. Informacją zaś są takie dane, które zostały przetworzone w sposób, który uwidacznia ich znaczenie i tym samym czyni je użytecznymi.

---

[6] *M. Hetmański*, Świat informacji, s. 32–38.

[7] O problemie tworzenia uniwersalnego – interdyscyplinarnego – znaczenia pojęcia „informacja" na potrzeby prawa pisze *G. Szpor*, Pojęcie informacji, s. 7–19.

[8] Przegląd podstawowych problemów rozważanych przez teoretyków informacji pochodzących z różnych środowisk naukowych przedstawia *S. Kurek-Kokocińska*, Informacja. Zob. również *B. Stefanowicz*, Informacja, s. 17.

[9] *G. Szpor*, Pojęcie informacji, s. 7–19.

[10] *M. Hetmański*, Epistemologia informacji, s. 124.

Gdy teoretyk informacji czyta ustawę o ochronie danych osobowych[11], wydaje mu się, że jest ona typowym przykładem mieszania tych dwóch pojęć. Wszak w rozumieniu ustawy za „dane osobowe" uważa się wszelkie informacje dotyczące zidentyfikowanej lub możliwej do zidentyfikowania osoby fizycznej[12]. Obiektywnie należy jednak przyznać, że twórcy ustawy – świadomie lub nieświadomie – słusznie uznali, że wartość informacyjną mogą posiadać każde „dane", a ograniczenie dodane do ustawy, mówiące, że informacji nie uważa się za umożliwiającą określenie tożsamości osoby, jeżeli wymagałoby to nadmiernych kosztów, czasu lub działań, słusznie subiektywizuje rozróżnienie tych dwóch pojęć. Polska ustawa o ochronie danych osobowych prawidłowo przyjmuje bowiem, że informacja o osobie składa się z kwantów – minijednostek – którymi są dane, a dane te uzyskują status danych osobowych w zależności od zasobu informacyjnego, który przetwarza dana osoba.

Bogdan Stefanowicz przyjmuje, że informacja ma pewne cechy, które determinują jej dalsze zastosowanie:

- jest niezależna od obserwatora,
- przejawia cechę synergii,
- jest różnorodna,
- jest zasobem niewyczerpalnym,
- może być powielana i przenoszona w czasie i przestrzeni,
- można ją przetwarzać, nie powodując jej zniszczenia,
- ta sama informacja może mieć różne znaczenie dla różnych użytkowników,
- każda jednostkowa informacja opisuje obiekt tylko ze względu na jedną jego cechę[13].

Zatrzymajmy się na chwilę, by ocenić cechy informacji przetwarzanych przy zastosowaniu nowoczesnych technologii w procesie przetwarzania informacji pierwotnie zbieranej jako informacja publiczna, szczególnie skupiając się na przy zastosowaniu technik *data miningu* dla przetwarzania rozproszonych informacji z wielu wielkoformatowych baz danych (tzw. *big data*)[14].

Synergią informacji nazywamy efekt zwielokrotnienia siły informacji przez zwiększenie jej zasobu. Łączne rozpatrywanie dwóch lub więcej komunika-

[11] Ustawa z 29.8.1997 r. o ochronie danych osobowych (t.j. Dz.U. z 2016 r. poz. 922), dalej jako: OchrDanychU.

[12] *M. Sakowska-Baryła*, Dostęp do informacji publicznej, s. 72–74.

[13] *B. Stefanowicz*, Informacja, s. 17–19; *tenże*, Informacyjne systemy zarządzania, s. 25. Zob. również *J. Kisielnicki*, *H. Sroka*, Systemy informacyjne, s. 14.

[14] *T. Creigh*, *M.E. Ludloff*, Privacy and big data, s. 30.

tów wywołuje bowiem większy efekt niż rozpatrywanie tych informacji niezależnie od siebie. Komunikaty wzmacniają i uzupełniają się wzajemnie. Przy ich łącznym rozpatrywaniu można wyłowić cechy niedostrzegalne przy odczytywaniu tylko jednego z nich. Trudno o lepszy przykład takiej zależności niż przetwarzanie danych z wielu baz danych. Dane z każdej z nich mogą być w oczywisty sposób niepełne bez danych pochodzących z drugiej. Rozdzielenie w zespole osób badających dane pracy tak, by każda z osób zapoznała się z jedną z baz, nie musi prowadzić do oczekiwanego efektu. Dopiero bowiem połączenie wielu informacji spowoduje zaistnienie synergii, która doprowadzi do sformułowania ostatecznej oceny opartej o całość zasobu. W sytuacjach takich często mamy do czynienia z przetwarzaniem danych „anonimowych" (np. oznaczonych jedynie numerem IP, numerem IMEI, MAC-adresem, numerem telefonu czy pseudonimem) w większości baz i łączeniem ich z kluczem deanonimizującym, zawartym w zaledwie jednej z baz.

Co prawda *B. Stefanowicz* twierdzi, że informacja jest niezależna od obserwatora. Należy jednak pamiętać, że takie założenie jest w pewnym sensie idealistyczne. Przyjmuje bowiem, że przekazana informacja jest jedynym zasobem informacyjnym posiadanym przez obserwatora. W idealnym modelu może tak być. W praktyce sprawa znacząco się komplikuje. Kwanty informacyjne przekazane osobie (systemowi teleinformatycznemu) A wywołają skutki poprzez ich połączenie z zasobem informacyjnym już posiadanym przez osobę (system teleinformatyczny) A. Przy przekazaniu ich osobie (systemowi teleinformatycznemu) B efekt może być całkowicie inny. U podstaw teorii kwantowej w fizyce leży zasada nieoznaczoności (nieokreśloności) *Heisenberga*. Przy przetwarzaniu informacji możemy mówić o „zasadzie nieoznaczoności (nieokreśloności) informacyjnej". Nigdy nie wiemy, gdzie kwanty informacyjne znajdą zaczepienie u nowej osoby (w nowym systemie), a nie ma możliwości zbadania tego bez wpływu na obiekt badany.

Przy jakichkolwiek działaniach podejmowanych na zasobie informacyjnym nie można zapominać, że zarówno pojedyncza informacja, jak i cały ich zasób pełnią funkcje praktyczne, które podczas przetwarzania mogą zostać zatracone. Tautologią jest w zasadzie stwierdzenie, że informacja pełni funkcję informacyjną. Przy przetwarzaniu informacji w systemach nie można zapominać, że cecha ta powinna zostać zachowana nie tylko dla systemu jako całości, ale również dla pojedynczej informacji. Wyszukiwanie danych odpowiadających jakiejś zadanej przez użytkownika systemu komputerowego cesze nie może dawać wyniku jedynie sumarycznego. Powinno również umożliwić zapoznanie się z każdą z wyszukanych informacji. W innym przypadku infor-

macja sumaryczna nie może być ponownie zweryfikowana co do swojej poprawności, a przypisana jednej osobie, „żyje" już osobnym życiem wraz z tą osobą – nawet wówczas, gdy pierwotne przypisanie było błędne.

§ 3. Informacja w zasobach publicznych

Podmioty publiczne w wykonaniu przepisów ustawowych prowadzą obecnie około 270 zbiorów ewidencyjnych, które są rejestrami publicznymi w rozumieniu art. 3 pkt 5 ustawy z 17.2.2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne[15]. Ich podstawę prawną stanowi ponad 150 ustaw i ponad 150 rozporządzeń[16]. Wszystkie one po implementacji dyrektywy o ponownym wykorzystywaniu informacji sektora publicznego[17] w 2011 r. mogą być z zasady traktowane jako zbiory informacji publicznej. Przepisy szczególne mogą co najwyżej ograniczać sposób wykorzystania tego typu informacji i poszczególnych składających się na nią kwantów. Zasady dostępu do informacji publicznej i jej ponownego wykorzystywania odnoszą się do uzyskania informacji publicznej z jawnych rejestrów publicznych – w szczególności z rejestrów zawierających dane referencyjne dla infrastruktury informacyjnej państwa. Szczególne znaczenie dla polskiego organu ochrony danych ma ponowne wykorzystanie informacji publicznej zawierającej dane osobowe. Przy poprawnie zbudowanej infrastrukturze informacyjnej państwa dane referencyjne zgromadzone są tylko w rejestrze (np. w księgach wieczystych, czy w zintegrowanym systemie informacji o nieruchomości, łączącym informacje z ksiąg wieczystych i z ewidencji gruntów i budynków), który można nazwać referencyjnym, i dane takie nie są powielane w innych rejestrach, a co najwyżej uzupełniane o dane specyficzne dla owego rejestru pochodnego. To powoduje, że stworzenie zintegrowanego systemu informacji o nieruchomościach

15 T.j. Dz.U. z 2014 r. poz. 1114 ze zm., dalej w skrócie: InformatyzU.

16 Najnowsze całościowe opracowanie tego problemu można znaleźć w pracy: Systemy informacyjne administracji publicznej. Źródła danych dla badań statystyki publicznej. Część 1, Główny Urząd Statystyczny 2011 (427 systemów). Warto również odesłać do – obejmującego 268 aktów prawnych – opracowania *D. Chromicka*, Aneks. Dane identyfikujące w rejestrach publicznych w Polsce, s. 137–162. Dokument elektroniczny: http://frdl.mazowsze.pl/userfiles/2008%20Diagnoza%20barier%20informatyzacji%20JST.pdf. Zob. również: *P. Drobek*, Ochrona danych osobowych, s. 103–120.

17 Dyrektywa Parlamentu Europejskiego i Rady 2003/98/WE z 17.11.2003 r. w sprawie ponownego wykorzystywania informacji sektora publicznego (Dz.Urz. UE L Nr 34, s. 90), dalej jako: dyrektywa 2003/98/WE.

(nad którym od lat pracuje minister właściwy ds. administracji wspólnie z ministrem właściwym ds. sprawiedliwości) wprowadziłoby zupełnie nową jakość w zakresie informacji poddanej ponownemu wykorzystaniu[18].

Dane z rejestrów publicznych – stanowiących podstawowe zasoby infrastruktury informacyjnej państwa – pobierane są w postaci dokumentu elektronicznego, czyli stanowiącego odrębną całość znaczeniową zbioru danych uporządkowanych w określonej strukturze wewnętrznej.

Tak rozumiany dokument elektroniczny jest zupełnie innym pojęciem niż „dokument" bądź „dokument urzędowy", definiowany dla różnych gałęzi prawa. Przy rozważaniu dalszych kwestii należy pamiętać, że ze względu na to, że pozyskiwanie informacji do ponownego wykorzystania następować będzie przeważnie w postaci elektronicznej, przy regulowaniu tych czynności z zasady należy posługiwać się pojęciem „dokumentu elektronicznego".

Jednym z oczywistych powodów transpozycji do prawa polskiego dyrektywy Parlamentu Europejskiego i Rady 2003/98/WE w sprawie ponownego wykorzystywania informacji sektora publicznego jest zwiększenie transparentności życia publicznego w Polsce. Temu samemu celowi powinno służyć realizowanie idei „otwartej władzy publicznej". Trzeba jednak z przykrością stwierdzić, że działania te nie zostały poprzedzone żadnymi rozważaniami dotyczącymi rzeczywistego znaczenia pojęcia „otwartości". Przy braku takiej refleksji pojawia się tendencja do utożsamiania „otwartości" i „przeszukiwalności" informacji pojęciem jawności formalnej, utrwalonym w doktrynie prawa i orzecznictwie w Polsce[19]. Tymczasem bez zmiany znaczenia pojęcia jawności formalnej lub bez odróżnienia tego pojęcia od „otwartego dostępu do informacji" (dostępności) będziemy mieli do czynienia z poważnym naruszaniem zasad ochrony prywatności osób, których dane znajdują się w „otwartych zasobach", jeśli będziemy zawsze przyjmować, że każdy zasób jawny formalnie powinien być dostępny do dowolnego ponownego wykorzystania[20].

W tym miejscu rozważań warto zwrócić uwagę na pewną zmianę terminologiczną, którą świadomie zastosowano w niniejszym opracowaniu. Rozważania, jakie dotychczas toczyły się w doktrynie prawa na temat jawności danych z rejestrów publicznych, określano zawsze jako badania nad „jawnością reje-

[18] *W.R. Wiewiórowski*, Ponowne przetwarzanie informacji publicznej, s. 392–400.

[19] O podobnej dyskusji w Wielkiej Brytanii pisze *P. Birkinshaw*, Freedom of Information, s. 29–30.

[20] Niemiecką dyskusję na ten temat podsumowuje *M. Eifert*, Państwowe struktury informacyjne, s. 204–208. Należy zwrócić uwagę, że problematyka ta jest stale przedmiotem oceny Trybunału Sprawiedliwości UE.

strów". Nie kwestionując tego przyjętego w teorii prawa pojęcia, z zasady należałoby dziś mówić o „jawności danych rejestrowych" lub „jawności danych z rejestru". W większości przypadków mamy bowiem do czynienia nie tyle z jawnością samego rejestru, co z jawnością jego odbicia, jakim jest zestaw danych – a co za tym idzie informacji – przetwarzanych w systemie teleinformatycznym, który z założenia nie jest rejestrem samym w sobie. Jednocześnie dane przetwarzane w systemie teleinformatycznym są ze swej natury bardziej zatomizowane niż dokumenty przechowywane w rejestrze papierowym. Są po prostu wartościami pewnych pól stworzonych przez twórcę bazy i obsługiwanych wedle poleceń sformułowanych przez twórcę oprogramowania obsługującego daną bazę[21].

To zaś oznacza, że niezależnie od regulacji prawnej mogą być one – z technicznego punktu widzenia – selekcjonowane w dowolny dopuszczony przez obu twórców rozwiązania informatycznego sposób. Mogą być więc wydobywane z bazy w oderwaniu od innych danych, z którymi w dokumencie papierowym są logicznie powiązane, oraz mogą być automatycznie wiązane logicznie z innymi danymi, z którymi w dokumencie papierowym nie są powiązane logicznie. Truizmem jest stwierdzenie, że w rejestrze prowadzonym w postaci relacyjnej bazy danych kwestia uzyskanego zestawu danych z różnych rekordów w jednej tabeli jest tylko kwestią sposobu skonstruowania zapytania do bazy. Równie wielkim truizmem jest stwierdzenie, że przetwarzanie danych z różnych dokumentów elektronicznych wywoływanych przy pomocy różnych zapytań z tej samej lub z różnych baz danych jest tylko funkcją pomysłowości *data minera* i wydajności stosowanych przez niego narzędzi informatycznych[22].

§ 4. Wszechobecność danych

Kolejne etapy informatyzacji procesów gospodarczych i administracyjnych i rosnąca rola środków komunikacji elektronicznej w życiu codziennym powodowały, że już od lat 70. XX w. wzrastało przekonanie, że istnieją pod-

[21] O zmianach, jakie niesie dostęp do rejestrów drogą elektroniczną, pisze *G. Sibiga*, Udostępnianie danych z rejestrów publicznych, s. 225–235.

[22] Por. choćby w: *G.G. Chowdhury*, Introduction to Modern Information Retrieval, rozdział: User-centred models of information retrieval, s. 214–226 oraz rozdział: Intelligent information retrieval, s. 352–371.

mioty, które uzyskują bądź mogą uzyskać dostęp do lawinowo rozwijających się zasobów informacyjnych i mogą wdrożyć sposoby przetwarzania informacji, które – nieznane nieświadomym uczestnikom rynku i obywatelom – mogą prowadzić do podejmowania wobec nich środków, których nie są świadomi i które wręcz mogą prowadzić do dyskryminacji osób, grup społecznych czy przedsiębiorców. Pierwotnie organizacją, która była w naturalny sposób oskarżana o chęć świadomego, acz ukrytego, przetwarzania rozproszonych danych w sposób, który może naruszać prawa i wolności, było państwo. Szybko rozprzestrzeniało się przekonanie, że praktyki potężnych rządów i korporacji w zakresie przetwarzania danych redukują jednostkę do statusu przedmiotu opisanego danymi, co zagraża prawom podstawowym i wolnościom. Już w latach 70. i 80. możemy przywołać liczne wezwania do ograniczenia takich praktyk lub wprowadzenia różnych mechanizmów kontroli nad działaniami państwa, a wkrótce również nad działaniami podmiotów rynkowych. W tym znaczeniu możemy powiedzieć, że zastrzeżenia wobec możliwości naruszania wolności i praw informacyjnych lub manipulowania wielkoskalowymi zasobami danych nie są niczym nowym. Tym jednak, co wyróżnia obecną falę zintegrowanego przetwarzania informacji przy wykorzystaniu technologii komunikacyjnych określanych terminem *big data*, jest wszechobecność takich działań i ich siła.

Liczba urządzeń podłączonych do Internetu przewyższa liczbę ludzi żyjących na Ziemi. Jest to jednak dopiero początek procesu, który zmultiplikuje liczbę urządzeń, dostępną pamięć i pasmo transmisji. Przewiduje się, że Internet rzeczy oraz analiza dużych zbiorów danych zostaną dodatkowo wzmocnione poprzez powiązanie tych działań z systemami opartymi na sztucznej inteligencji, przetwarzania poleceń i treści zapisanych w języku naturalnym oraz z systemami przetwarzającymi informacje biometryczne. Choć sama idea zastosowania sztucznej inteligencji dla umożliwienia systemom uczenia się nie jest nowa, dla trzeciej dekady XXI w. będzie to już nie idea, lecz rzeczywistość. Instytucje publiczne i podmioty komercyjne są dziś w stanie wykroczyć poza „eksplorację danych" ku działalności, którą można by nazwać „eksploracją rzeczywistości"[23].

Tak w Polsce, jak i całej Europie panuje przekonanie, że konieczność wykorzystywania takich rozwiązań przez administrację publiczną nie idzie w parze możliwościami tworzenia i utrzymywania takich rozwiązań przez polskie i europejskie podmioty publiczne. Z pewnością instytucje publiczne – nawet

[23] *N. Bostrom*, Superinteligencja, s. 41–43.

te związane z utrzymaniem bezpieczeństwa publicznego – nie są dziś w stanie samodzielnie prowadzić centrów kompetencyjnych badających i wdrażających prawdziwie innowacyjne metody przetwarzania danych. Liderem jest tu z pewnością biznes. Jednocześnie panuje nie do końca słuszne przekonanie, że najbardziej innowacyjne rozwiązania informatyczne powstają dziś poza Europą.

§ 5. Rola Internetu rzeczy

Przełomem w zakresie wielkoskalowych rozproszonych zasobów informacyjnych przetwarzanych przy pomocy rozbudowanych systemów informacyjnych bazujących na automatycznie działających i samouczących się aplikacjach będzie rozwój Internetu rzeczy. Przewiduje się, że do 2020 r. łączność z siecią będzie standardową funkcją, a podłączonych będzie 25 miliardów przedmiotów (w 2015 r. było to 4,8 mld) – od przedmiotów związanych z telemedycyną po pojazdy, od inteligentnych liczników po całą gamę nowych – stacjonarnych i mobilnych – urządzeń wspierających inteligentne miasta.

Inteligentne środowiska[24], w których rozgrywać będą się procesy informacyjne XXI w. – takie jak choćby inteligentne miasto – nie są jednolitą strukturą stworzoną przez jednego architekta czy nawet jedną grupę projektową. Nie są nawet zespołem systemów zarządzanych centralnie przez centrum koordynacyjne na poziomie państwa, miasta czy sektora rynku. Podstawowym rozwiązaniem stają się zespoły z zasady otwartych systemów informacyjnych, które umożliwiają dynamiczne dołączanie do architektury stworzonej dla danego środowiska nowych komponentów. Część z systemów pozostaje zamknięta i dostępna jedynie dla głównych twórców danego kompleksu, lecz z założenia należy przyjmować, że i te systemy dążyć będą w najbliższej przyszłości do większej otwartości. Otwartość takich środowisk staje się podstawową wartością sama w sobie. Można zaryzykować stwierdzenie, że jedynie środowiska z zasady otwarte będą mogły skutecznie rozwijać się w sytuacji, gdy sama komunikacja w tych środowiskach oparta jest o otwarte sieci. Nie ma wątpliwo-

[24] Inteligentnym środowiskiem nazywamy obszar w świecie fizycznym, który jest bogato nasycony niewidocznymi sensorami, nadajnikami, czytnikami i innymi elementami przetwarzającymi dane, wbudowanymi w przedmioty codziennego użytku i stale połączonymi z siecią. Zob. *M. Weiser*, *R. Gold*, *J.S. Brown*, The origins of ubiquitous computing research, s. 693–696. Zob. również *Ch. Perera*, *P. Jayaraman*, *A. Zaslavsky*, *P. Christen*, *D. Georgakopoulos*, Context-Aware Dynamic Discovery, s. 215–218.

ści, że tworzone będą moduły zamknięte, ale będą one wyjątkiem od generalnej – chronionej przez prawo – zasady otwartości. Otwartość systemów nie będzie wszakże oznaczała całkowitego otwarcia dostępu do wszystkich przetwarzanych zasobów.

By przejawiać rzeczywistą „inteligencję", środowiska i działające w nich systemy informacyjne muszą stale reagować na zmieniające się otoczenie. To w naturalny sposób kieruje nas do uznania, będą one z zasady opierać się na infrastrukturze Internetu rzeczy (*Internet of Things*, IoT)[25].

Internet rzeczy można zdefiniować jako globalną infrastrukturę, w której przedmioty (rzeczy) jako takie lub połączone w jakikolwiek sposób z innymi przedmiotami lub osobami, otrzymują unikalne identyfikatory oraz są w stanie przekazywać dane i współdziałać z innymi systemami, wykorzystując możliwości współpracy. Internet rzeczy ewoluuje w zależności od szybkości i głębokości konwergencji technologii telekomunikacyjnych, systemów elektromechanicznych, mocy obliczeniowych serwerów, przetwarzania w chmurze oraz Internetu. W tym kontekście „rzeczy" definiuje się jako obiekty rzeczywiste lub wirtualne, które istnieją i poruszają się w czasie i przestrzeni oraz mogą być zidentyfikowane przez przypisane im numery identyfikacyjne, nazwy lub adresy lokalizacji[26]. Nieco prostszym językiem IoT definiuje się jako zespół rozwiązań umożliwiający ludziom i rzeczom łączenie się w dowolnym czasie i miejscu z czymkolwiek lub kimkolwiek, w idealnym świecie czyniąc to dowolną drogą lub siecią przy użyciu dowolnej usługi[27].

---

[25] Samo pojęcie Internetu przedmiotów ma już ponad 15 lat, gdyż po raz pierwszy użyte zostało przez *K. Ashtona* w 1999 r. podczas prezentacji dla Procter & Gamble – *E. Anzelmo*, *A. Bassi*, *D. Caprio*, *S. Dodson*, *R. van Kranenburg*, *M. Ratto*, Discussion Paper on the Internet of Things, s. 8, dostępne na stronie: http://www.theinternetofthings.eu/content/discussion-paper-iot-research-institute-internet-society-berlin (dostęp 24.11.2015 r.).

[26] *G. Santucci*, Towards Connectobjectome, s. 7–11. Zob. również: *E. van der Zee*, *H. Scholten*, Spatial Dimensions of big data, s. 140–141.

[27] *„The Internet of Things could allow people and things to be connected Anytime, Anyplace, with Anything and Anyone, ideally using Any path/network and Any service"*. Zob. *O. Vermesan*, *P. Friess i in.*, Internet of Things, s. 12. Sama definicja została wyprowadzona z rozważań o *ubiquitous computing* w ITU Internet Reports 2005: The Internet of Things, s. 2 i n., który autor niniejszego rozdziału omawiał podczas 6. Europejskiej Konferencji Ministerialnej na temat eGovernment pt. „Borderless eGovernment Services for Europeans. egov2011PL" w referacie „From eGovernment to uGovernment: new challenges for the European policy agenda". Konferencja została zorganizowana przez Polską Prezydencję w radzie Unii Europejskiej oraz Komisję Europejską w Poznaniu w dniach 17–18.11.2011 r.

Internet rzeczy[28], obejmujący miliardy połączonych urządzeń, jest w stanie przechowywać wszelkiego rodzaje dane (takie jak temperatura, wilgotność, nagrania wideo, ruch, tętno). Podczas gdy istniejące urządzenia mierzą dane typowe dla środowiska, nowe urządzenia skupiają się bardziej na obserwacji zwyczajów użytkowników. Internet rzeczy łączy zarówno urządzenia istniejące dotychczas, a wyposażone dodatkowo w chip RFID[29], zapewniający im bytność w Internecie, jak i urządzenia utworzone specjalnie w celu połączenia z Internetem (na stałe lub jedynie okazjonalnie).

Pierwsza faza rozwoju systemów informacyjnych umożliwiających działanie Internetu rzeczy polegała na umożliwieniu unikalnej identyfikacji przedmiotów bez bezpośredniej interakcji z osobami, następnie przedmioty te uzyskały możliwość interakcji z innymi przedmiotami oraz przyjmowania i wydawania poleceń i tworzenia dodatkowej nowej informacji. W kolejnej fazie rozwoju na podstawie tak gromadzonej informacji urządzenia będą w stanie same podejmować decyzje co do sposobu działania innych urządzeń. Jednocześnie zaś większość urządzeń podłączonych od Internetu rzeczy uruchamiać będzie transmisję danych automatycznie, gdy znajdzie się w zasięgu sieci oportunistycznej lub poprzez podłączenie się urządzenia do „napotkanego" inteligentnego środowiska.

Internet rzeczy bazuje na stałym – lub przynajmniej bardzo częstym – przekazie ze strony szerokiej gamy urządzeń i dzięki temu jest idealnym środowiskiem dla rozbudowanej interakcji urządzeń i systemów w inteligentnych środowiskach. Do tak rozumianego środowiska należą nie tylko rozwiązania infrastrukturalne i urządzenia końcowe, ale również aplikacje i zasoby, które działając w tym środowisku, przejawiać będą również cechy obiektów.

---

[28] Więcej rozważań o prawnych aspektach Internetu rzeczy w: *W.R. Wiewiórowski*, Ochrona danych osobowych w świecie inteligentnych przedmiotów, s. 3–11.

[29] Oceną skutków stosowania technologii RFID dla ochrony prywatności zajmowała się nieformalna grupa robocza ds. RFID poddana nadzorowi Grupy Roboczej Art. 29. Efektami tego nadzoru były opinia Nr 5/2010 na temat propozycji sektora w sprawie ram oceny skutków w zakresie ochrony danych i prywatności w zastosowaniach RFID, przyjęta 13.7.2010 r. (krytyczna wobec pierwotnej propozycji) i opinia Nr 9/2011 na temat zmienionej propozycji sektora w sprawie ram oceny skutków w zakresie ochrony danych i prywatności w zastosowaniach RFID, przyjęta 11.2.2011 r. (akceptująca drugą wersję propozycji podmiotów rynkowych). Obydwa dokumenty dostępne na stronie: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/index_en.htm (sprawdzono 24.11.2015 r.). Zob. również *S. Spiekermann*, The RFID PIA, s. 323 oraz *L. Beslay*, *A.-C. Lacoste*, Double-take: getting to the RFID PIA Framework, s. 347–348. O generalnych kwestiach prawnych związanych z interoperacyjnością systemów opartych o RFID piszą *R.H. Weber*, *R. Weber*, Internet of Things, s. 3–5.

Będą więc jednocześnie elementem tworzącym środowisko i „przejawem" jego funkcjonowania. Dynamicznie tworzone i przekształcające się algorytmy korzystające z danych, które zbierane są i przetwarzane przez różne sensory, będą operowały tak, jakby były niezależnymi elementami architektonicznymi (klasycznym przykładem będą tu rejestry publiczne i aplikacje służące do ich obsługi).

Stałe korzystanie przez rozbudowane systemy informacyjne z danych pochodzących z Internetu rzeczy pokazuje skalę zmiany w procesach informacyjnych. Czujniki uczestniczące w wymianie informacji w Internecie rzeczy zapewniać będą stały i natychmiastowy dostęp do bardzo szczegółowych danych, do których urzędy statystyczne i badacze nie mogli nigdy dotrzeć. Sam dostęp do danych jest jednak tylko niewielkim fragmentem procesu informacyjnego; warunkiem *sine qua non* jego działania, ale nie jego *clue*. Sercem działania systemów informacyjnych staje się działanie algorytmów, które w sposób niejawny dla przeciętnego użytkownika przekształca informację wejściową zebraną przez czujniki w informację wyjściową, która automatycznie staje się informacją wejściową dla kolejnego algorytmu, który przekazuje wyniki swego działania kolejnemu algorytmowi działającemu w tym samym systemie lub w interoperacyjnym wobec niego systemie zewnętrznym. Interoperacyjność systemów powoduje zaś, że użytkownik nie jest zaś w stanie ocenić, które systemy są w danej chwili połączone i jak wykorzystują przetwarzaną informację. Ten łańcuch algorytmów staje się częścią systemu, który możemy z łatwością uznać za po części przynajmniej samosterowalny. Jako że niektóre z jego funkcjonalności uruchamiać się będą tylko w konkretnych warunkach, których osiągnięcie nie jest kontrolowane na bieżąco przez architekta czy administratora systemu, a jedynie przez narzędzia wbudowane w sam system. Tak rozumiane działania autonomicznie uruchamiane w ramach sytemu stają się częścią obliczeń nazywanych środowiskowymi i niewidocznych dla jego twórców i użytkowników[30].

§ 6. Chmura obliczeniowa

Istotną zmianą jest upowszechnienie się zastosowania środowiska umożliwiającego nie tylko zaawansowaną analitykę i możliwości eksploracyjne, gro-

[30] Rolę Internetu rzeczy w tworzeniu nowego ładu gospodarczego i społecznego opisuje *J. Rifkin*, Społeczeństwo zerowych kosztów, s. 20 i 83–87.

madzenie i analizę dużych zbiorów danych, lecz także przepływ danych z Internetu rzeczy – jakim jest chmura (mniej poprawnie nazywana „chmurą obliczeniową"). Zastosowanie chmur umożliwiło koncentrowanie danych z rozmaitych urządzeń wykorzystujących „Internet rzeczy", jednocześnie dając szansę korzystania z olbrzymich ilości danych przechowywanych w działającej na szeroką skalę infrastrukturze do przechowywania i przetwarzania danych na całym świecie, oraz łączność z takimi danymi.

Cloud computing nie jest nowym fenomenem i w oczywisty sposób nie trzeba dlań tworzyć całości otoczenia prawnego. Większość rozwiązań technologicznych istniało w świecie komunikacji elektronicznej już od wielu lat w postaci technik i rozwiązań praktycznych wspierających różne formy wirtualizacji. Również od dawna wprowadzano odpowiednie rozwiązania prawne, które zabezpieczały interesy użytkowników procesu wirtualizacji. Tym niemniej eksplozja usług chmurowych po 2008 r. zmieniła zasadniczo warunki, w których owe rozwiązania działają. Po pierwsze, większość umów zawieranych w zakresie usług chmurowych stała się umowami adhezyjnymi, co wcale nie zniechęca administracji jako takiej, jak i jej przedstawicieli – urzędników – w korzystaniu z niej przy realizacji zadań publicznych. Po drugie zaś, chmury „umiędzynarodowiły się" w sposób dotąd niespotykany, można wręcz powiedzieć, że z technicznego punktu widzenia w ogóle pominęły fakt występowania granic pomiędzy państwami i porządkami prawnymi[31]. Ten proces może być porównany jedynie z niektórymi rewolucjami w handlu międzynarodowym związanymi z pojawieniem się nowych możliwości dla od dawna istniejących środków transportu. O ile jednak prawo morskie, odpowiadające na jedną z takich rewolucji, przybierało dzisiejszą formę przez setki lat, bazując na zwyczajach handlowych z wolna wprowadzanych do prawa stanowionego poszczególnych krajów i ostatecznie sankcjonowanych w normach prawa międzynarodowego, o tyle w przypadku chmur proces ten ma zająć co najwyżej klika lat. Co więcej, o ile w przypadku transportu morskiego – mimo pojawiających się od czasu do czasu kryzysów – mamy do czynienia z technikami, które będą używane jeszcze przez lata, o tyle przy cloud computingu już dziś szukamy jego następców[32].

[31] Analiza SWOT dla europejskich usług chmurowych w: *L. Schubert*, *K. Jeffery*, *B. Neidecker-Lutz*, The Future of Cloud Computing Opportunities, s. 35–36.

[32] O problemie stałego wyważania pomiędzy precyzyjnością i elastycznością przepisów prawa nowych technologii piszą *R. Brownsword*, *M. Goodwin*, Law and the Technologies, s. 306–307.

Definicja przetwarzania w chmurze przygotowana przez *National Institute of Standards and Technology* (NIST) zwraca uwagę na możliwości nieograniczonego przestrzenią, wygodnego dostępu na żądanie na żądanie do konfigurowalnej i współdzielonej puli zasobów obliczeniowych (np. sieci, serwerów, pamięci masowej, aplikacji i usług), które mogą być dostarczane szybko i wymagają minimalnego wysiłku w kwestii zarządzania i interakcji z usługodawcą. Jednocześnie NIST podkreśla, że przetwarzanie w chmurze nie jest nową technologią[33]. Instytut definiuje również pięć głównych cech chmur oraz ich trzy modele usługowe i pięć modeli wdrożeniowych

Pięć głównych cech chmurowego modelu przetwarzania to:

1) *on-demand self-service* – klienci chmury powinni mieć możliwość skorzystania z usług w chmurze (np. obliczeniowych, pamięci i zasobów pamięci masowej), używając mechanizmu samoobsługi (np. portal internetowy), tak żeby nabywanie usług nie wymagało interwencji przez usługodawcę chmury;
2) *broad network access* – chmura powinna być dostępna z każdego miejsca (jeśli wymagane) na różnych urządzeniach, takich jak: smartfony, tablety, laptopy, komputery stacjonarne oraz na wszystkich innych typach czytników, istniejących obecnie lub w przyszłości;
3) *resource pooling* – chmury powinny udostępniać pule współdzielonych zasobów, które wykorzystywane są przez klientów chmury. Zasoby, takie jak: moc obliczeniowa, pamięć, sieć i dysk, przydzielone są do klientów korzystających z usług udostępnionej, wspólnej puli. Zasoby pobierane są z aktualnej lokalizacji, a klienci nie są świadomi lokalizacji tych zasobów;
4) *rapid elasticity* – chmury powinny zapewnić szybkie zastrzeganie i uwolnienie zasobów ze względu na zapotrzebowanie usług w chmurze. Powinno to odbywać się automatycznie, bez konieczności ingerencji człowieka. Ponadto, odbiorcy usługi chmury powinni odnosić wrażenie, że istnieje nieograniczona pula zasobów – usługa jest w stanie spełnić wymagania dla dowolnego scenariusza w przypadku użycia;
5) *metered Services* – model chmury, określany jako *pay-as-you-go*, powinien czasem umożliwiać ładowanie usług konsumenta chmury w oparciu o rzeczywiste wykorzystanie zasobów chmury. Wykorzystanie zaso-

[33] *P. Mell*, *T. Grance*, The NIST Definition of Cloud Computing, s. 2; *W. Jansen*, *T. Grance*, Guidelines on Security and Privacy, s. 3 oraz *L. Badger*, *T. Grance*, *R. Patt-Corner*, *J. Voas*, Draft Cloud Computing Synopsis, s. 2–1.

bów jest monitorowane, zgłaszane i kontrolowane przez dostarczyciela usług chmurowych i politykę serwisu, które zapewniają przejrzystość rozliczeń, zarówno z dostarczycielem usług chmurowych, jak i z konsumentem usług[34].

Istotne dla tych rozważań jest to, że wykorzystanie istniejącej puli (*resource pooling*) obejmuje pojęcie niezależności lokalizacji. Taka niezależność oznacza jednak, że użytkownik w tradycyjnej chmurze publicznej nie jest w stanie kontrolować dokładnej lokalizacji zasobu. Z punktu widzenia dzisiejszych norm prawnych z zakresu ochrony danych osobowych jest jednak istotne, by użytkownik mógł kontrolować lokalizację na nieco wyższym poziomie abstrakcji (np. kraj, stan, centrum przetwarzania danych lub w – przypadku istnienia wiążących reguł korporacyjnych – korporacja).

Większość rozważanych kwestii prawnych odnosi się w podobny sposób do wszystkich trzech kategorii przetwarzania danych w chmurze, tj. do infrastruktury jako usługi (*Infrastructure as a Service* – IaaS), platformy jako usługi (*Platform as a Service* – PaaS) i oprogramowania jako usługi (*Software as a Service* – SaaS), jednak podane dalej rozwiązania i sugestie będą brały po uwagę przede wszystkim modele IaaS i SaaS, jako że *Platform as a Service* jest z zasady używane do celu budowy i rozwijania oprogramowania przy pomocy platformy dostarczanej przez chmurę. Z jednej strony wymaga to rozważenia przede wszystkim kwestii z zakresu prawa własności intelektualnej (tak własności przemysłowej, jak i praw autorskich), co nie jest celem niniejszego opracowania. Z drugiej zaś strony usługi administracji publicznej z zasady nie polegają na działaniach deweloperskich w tym znaczeniu i ewentualne wykorzystywanie modelu PaaS przez podmioty wykonujące zadania publiczne ma raczej charakter pomocniczy. Pozostałe coraz popularniejsze „aa-skróty", takie jak BaaS (*Business as a Service*)[35], CaaS (*Communications as a Service* – integracja komunikacji głosowej, wizualnej, chatów, komunikatorów itp.), DaaS (*Data as a Service* – np. The Google® Geocoding API™), EaaS (*Email as a Service*), GaaS (*Government as a Service*)[36] powinny być traktowane jedynie jako określenia zbiorcze dla rodzajów usług chmurowych[37], a nie jako modele biznesowe[38].

---

34 Tłumaczenie za: *M. Ledwoch*, Przegląd architektury Chmur Prywatnych. Nieco inaczej zasady te przedstawili w *A. Mateos*, *J. Rosenberg*, Chmura obliczeniowa, s. 27–31.

35 *M. Zgajewski* (red.), Cloud computing w sektorze finansowym, s. 6.

36 *R. Cohen*, Introducing government as a service.

37 Zob. również: *D. Blankenhorn*, *V. Ristau*, *C. Beesley*, Cloud Computing for Govies, s. 18–19; *D. McClure*, Cloud Shopping Made Easy, s. 12–13; *D. McClure*, *C. Coleman*, Providing A Helping Hand, s. 13 oraz *D. Catteddu*, *G. Hogben*, Cloud Computing, s. 120–121.