



1021/00/PL
WP207

Opinia 06/2013 w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego (ISP)

Przyjęta w dniu 5 czerwca 2013 r.

Grupa robocza została powołana na mocy art. 29 dyrektywy 95/46/WE. Jest ona niezależnym europejskim organem doradczym w zakresie ochrony danych i prywatności. Zadania grupy określone są w art. 30 dyrektywy 95/46/WE i art. 15 dyrektywy 2002/58/WE.

Obsługę sekretariatu zapewnia Dyrekcja C (Prawa Podstawowe i Obywatelstwo Unii Europejskiej) Dyrekcji Generalnej ds. Sprawiedliwości Komisji Europejskiej, B-1049 Bruksela, Belgia, biuro nr MO-59 02/013.

Strona internetowa: http://ec.europa.eu/justice/data-protection/index_en.htm

GRUPA ROBOCZA DS. OCHRONY OSÓB FIZYCZNYCH W ZAKRESIE PRZETWARZANIA DANYCH OSOBOWYCH

powołana na mocy dyrektywy 95/46/WE Parlamentu Europejskiego i Rady z dnia 24 października 1995 r.,

uwzględniając art. 29 i art. 30 ust. 1 lit. a) oraz art. 30 ust. 3 tej dyrektywy,

uwzględniając swój regulamin wewnętrzny,

PRZYJMUJE NINIEJSZĄ OPINIĘ:

I. Wprowadzenie

1.1. Przegląd dyrektywy ISP

W dniu 26 czerwca 2013 r. Unia Europejska przyjęła dyrektywę 2013/37/UE Parlamentu Europejskiego i Rady („dyrektywa zmieniająca dyrektywę ISP”) zmieniającą dyrektywę 2003/98/WE w sprawie ponownego wykorzystywania informacji sektora publicznego („dyrektywa ISP”)¹.

Celem dyrektywy ISP jest ułatwienie ponownego wykorzystania informacji sektora publicznego poprzez harmonizację warunków ich ponownego wykorzystania w całej Unii Europejskiej i usunięcie niepotrzebnych barier utrudniających ich ponowne wykorzystanie na rynku wewnętrznym.

W pierwotnym tekście dyrektywy ISP z 2003 r. zharmonizowano warunki ponownego wykorzystania, lecz nie wymagano od organów sektora publicznego udostępnienia danych w celu ponownego wykorzystania. Kwestia udostępnienia danych w celu ponownego wykorzystania była zasadniczo nieobowiązkowa: podjęcie decyzji w tej sprawie pozostawiono państwom członkowskim i organom sektora publicznego. W rezultacie wiele organów sektora publicznego w całej Europie po prostu zdecydowało się nie zezwalać na ponowne wykorzystanie ich informacji.

W tym kontekście jednym z kluczowych celów polityki dyrektywy zmieniającej dyrektywę ISP jest wprowadzenie zasady, zgodnie z którą wszystkie informacje publiczne (tzn. wszystkie informacje będące w posiadaniu sektora publicznego, które są publicznie dostępne na mocy prawa krajowego) są dostępne do ponownego wykorzystania zarówno w celach komercyjnych, jak i niekomercyjnych. W niektórych przypadkach mają zastosowanie wyłączenia z zakresu zmienionej dyrektywy ISP, w tym ze względu na ochronę danych².

W związku z tym w zmienionej dyrektywie ISP na organy sektora publicznego nakłada się obecnie obowiązek umożliwienia ponownego wykorzystania wszelkich informacji będących w ich posiadaniu. Jak zostanie wyjaśnione poniżej, dyrektywa nie nakłada jednak na organy sektora publicznego obowiązku publicznego ujawniania danych osobowych. W dyrektywie ISP nakłada się obowiązek ponownego wykorzystania informacji jedynie w przypadku, gdy informacje te są już publicznie dostępne na mocy prawa krajowego, i nawet wtedy wyłącznie pod warunkiem, że ich ponowne wykorzystanie nie naruszy zasad mających zastosowanie przepisów o ochronie danych.

1 Dz.U. L 175 z 27.6.2013, s.1.

2 Dalsze informacje na temat zakresu zmienionej dyrektywy ISP i przepisów dotyczących ochrony danych można znaleźć w pkt V poniżej.

Inne istotne nowe przepisy dyrektywy zmieniającej dyrektywę ISP rozszerzają zakres dyrektywy ISP w celu uwzględnienia bibliotek (w tym bibliotek uniwersyteckich), archiwów i muzeów.

W świetle powyższego dzięki zmienionej dyrektywie ISP istnieje możliwość znacznego zwiększenia dostępności informacji będących w posiadaniu organów publicznych.

1.2. Ponowne wykorzystywanie ISP i dane osobowe

Inicjatywy dotyczące ponownego wykorzystywania ISP zazwyczaj obejmują (i) udostępnienie całych baz danych (ii) w standardowym formacie elektronicznym (iii) każdemu wnioskodawcy bez stosowania postępowania sprawdzającego, (iv) bezpłatnie (lub z zastosowaniem niewielkiej opłaty) i (v) w celach komercyjnych lub niekomercyjnych bez żadnych warunków (lub w uzasadnionych przypadkach przy zastosowaniu nieograniczających warunków w ramach licencji)³.

Takie udostępnienie może przynieść korzyści, prowadząc do większej przejrzystości i innowacyjnego ponownego wykorzystania informacji sektora publicznego. Wynikająca z tego działania większa dostępność informacji wiąże się jednak z pewnym ryzykiem.

W celu ograniczenia tego ryzyka w każdym przypadku, który dotyczy danych osobowych, przepisy o ochronie danych muszą zapewnić pomoc w kierowaniu procesem wyboru w odniesieniu do kwestii, jakie dane osobowe mogą lub nie mogą być udostępniane do ponownego wykorzystania i jakie środki należy podjąć, aby zabezpieczyć dane osobowe. We wszystkich przypadkach, w których zagrożona jest ochrona prywatności i ochrona danych osobowych, konieczne jest stosowanie zrównoważonego podejścia. Z jednej strony przepisy w zakresie ochrony danych osobowych nie powinny stanowić zbędnej przeszkody w rozwoju rynku ponownego wykorzystywania ISP. Z drugiej strony niezbędne jest poszanowanie prawa do ochrony danych osobowych i prawa do prywatności. Należy podkreślić, że koncepcja otwartych danych skupia się na przejrzystości i odpowiedzialności organów sektora publicznego oraz na wzroście gospodarczym, a nie na przejrzystości pojedynczych obywateli.

Przy stosowaniu dyrektywy ISP i przepisów o ochronie danych do ponownego wykorzystywania danych osobowych organ sektora publicznego może podjąć jedną z trzech różnych decyzji:

1. decyzję o nieudostępnieniu danych osobowych do ponownego wykorzystania zgodnie z warunkami określonymi w dyrektywie ISP;
2. decyzję o przekształceniu danych osobowych do formy zanonimizowanej (najczęściej w formie zbiorczych danych statystycznych)⁴ i udostępnieniu do ponownego wykorzystania tylko takich zanonimizowanych danych;
3. decyzję o udostępnieniu danych osobowych do ponownego wykorzystania (w razie potrzeby podlegającą szczególnym warunkom i odpowiednim zabezpieczeniom).

³ Należy zauważyć, że zgodnie z art. 8 ust. 1 zmienionej dyrektywy ISP „warunki [licencji] nie ograniczają niepotrzebnie możliwości ponownego wykorzystywania i nie są stosowane do ograniczania konkurencji”.

⁴ Dalsze informacje na temat ponownego wykorzystywania zagregowanych i zanonimizowanych zbiorów danych uzyskanych z danych osobowych można znaleźć w pkt VI poniżej.

II. Cel opinii

2.1. Spójne wytyczne i najlepsza praktyka

Celem niniejszej opinii jest ułatwienie wzajemnego zrozumienia właściwych ram prawnych i przedstawienie przykładów spójnych wytycznych i najlepszych praktyk w zakresie sposobu wdrażania dyrektywy ISP (z późniejszymi zmianami) w odniesieniu do przetwarzania danych osobowych.

Celem niniejszej opinii nie jest próba harmonizacji podejść krajowych w odniesieniu do poziomu przejrzystości, przepisów krajowych dotyczących dostępu do dokumentów i dostępności informacji w ramach tych przepisów krajowych. Krajowe przepisy wykonawcze w odniesieniu do dyrektywy ISP i krajowa interpretacja dyrektywy 95/46/WE⁵ dotyczące ponownego wykorzystywania ISP czasami różnią się jednak w stopniu wykraczającym poza kwestie, które mogą być niezbędne, by uwzględnić różnorodność krajowych systemów dostępu i różnych poziomów przejrzystości.

W związku z powyższym zalecenia polityczne w zakresie prywatności z września 2012 r. sporządzone przez sieć tematyczną LAPSI wyraźnie ilustrują niepotrzebne rozbieżności w sposobach transponowania dyrektywy ISP w poszczególnych państwach członkowskich w odniesieniu do ochrony danych osobowych⁶. W samej dyrektywie ISP zawarto również ostrzeżenie, iż różnice i wątpliwości związane z prawodawstwami mogą nabrać jeszcze większego znaczenia wraz z dalszym rozwojem społeczeństwa informacyjnego, co już spowodowało znaczny wzrost transgranicznego wykorzystywania informacji⁷.

Brak spójnego podejścia może osłabić pozycję zainteresowanych osób fizycznych. Może to także wiązać się ze zbędnymi obciążeniami regulacyjnymi dla przedsiębiorstw i innych organizacji, które prowadzą działalność na poziomie transgranicznym, i tym samym stanowią przeszkodę w rozwoju wspólnego europejskiego rynku ponownego wykorzystywania ISP. Z jednej strony osoby, których dane dotyczą, muszą mieć pewność, że ich dane będą konsekwentnie chronione niezależnie od tego, czy zostały przekazane do innego państwa członkowskiego do celów ponownego wykorzystania. Z drugiej strony powinno się także unikać zbędnej złożoności i fragmentacji, aby umożliwić swobodny przepływ danych osobowych w całej Europie, co stanowi kolejny kluczowy cel dyrektywy 95/46/WE.

2.2. Potrzeba aktualizacji opinii nr 7/2003

Przyjęcie dyrektywy zmieniającej dyrektywę ISP następuje 10 lat po przyjęciu dyrektywy ISP w 2003 r. W tamtym czasie Grupa Robocza Art. 29 przyjęła opinię w sprawie kwestii ochrony danych w związku z ISP („Opinia nr 7/2003”)⁸. Chociaż główne zasady przedstawione w opinii nr 7/2003 są aktualne, zmiany technologiczne i inne zmiany w obszarze ISP oraz ochrony danych, w tym

⁵ Dyrektywa 95/46/WE Parlamentu Europejskiego i Rady z dnia 24 października 1995 r. w sprawie ochrony osób fizycznych w zakresie przetwarzania danych osobowych i swobodnego przepływu tych danych (Dz.U. L 281 z 23.11.1995, s. 31).

⁶ LAPSI jest europejską siecią tematyczną ds. aspektów prawnych informacji sektora publicznego sfinansowaną przez Komisję Europejską, zob. <http://www.lapsi-project.eu/>. Treść zalecenia politycznego jest dostępna na stronie internetowej http://www.lapsi-project.eu/lapsifiles/lapsi_privacy_policy.pdf.

⁷ Zob. motyw 7.

⁸ Zob. opinia Grupy Roboczej Art. 29 z 7/2003 w sprawie ponownego wykorzystywania informacji sektora publicznego i ochrony danych osobowych – zachowanie równowagi – przyjęta w dniu 12 grudnia 2003 r. (WP 83). Zob. również dwie poprzednie powiązane opinie Grupy Roboczej Art. 29: opinia nr 3/1999 w sprawie informacji sektora publicznego i ochrony danych osobowych przyjęta w dniu 3 maja 1999 r. (WP 20), a także opinia nr 5/2001 dotycząca sprawozdania specjalnego Europejskiego Rzecznika Praw Obywatelskich przyjęta w dniu 17 maja 2001 r.

proponowane zmiany legislacyjne w obu obszarach, uzasadniają obecne działania na rzecz zaktualizowania i uzupełnienia opinii z 2003 r.

Ponadto w niniejszej opinii można teraz również uwzględnić inne podjęte ostatnio i prowadzone na bieżąco działania w celu zapewnienia dalszych wytycznych, w szczególności:

- opinię Europejskiego Inspektora Ochrony Danych („EIOD”) z dnia 18 kwietnia 2012 r. dotyczącą pakietu Komisji w sprawie otwartego dostępu do danych⁹;
- opinię nr 3/2013 Grupy Roboczej Art. 29 dotyczącą zasady celowości¹⁰;
- prowadzone na bieżąco działania w podgrupie technologicznej Grupy Roboczej Art. 29 w zakresie technik anonimizacji¹¹;
- prowadzone w niektórych państwach członkowskich działania w zakresie anonimizacji i oceny ryzyka¹²; oraz
- istniejące orzecznictwo i praktykę w zakresie zrównoważenia ponownego wykorzystywania danych osobowych i ich ochrony w niektórych państwach członkowskich¹³.

III. Treść i struktura opinii

W opinii nr 7/2003 uwzględniono w szczególności zasadę celowości¹⁴, ale poruszono również inne kwestie, takie jak zgodne z prawem podstawy publicznego ujawniania i ponownego wykorzystywania ISP, szczególna ochrona zapewniana w odniesieniu do danych szczególnie chronionych, przekazywanie danych do państw trzecich, jakość danych oraz prawa przysługujące osobom, których dane dotyczą. Kwestie te są w dalszym ciągu istotne. Uwzględniając przeprowadzone już wcześniej działania, w niniejszej opinii jedynie aktualizuje się i uzupełnia wnioski przedstawione w opinii nr 7/2003, jeżeli jest to konieczne w świetle nowych zmian w prawodawstwie i zmian technologicznych.

W sekcji IV wyjaśniono, że obowiązek ponownego wykorzystywania na podstawie zmienionej dyrektywy ISP pozostaje bez uszczerbku dla wymogów dotyczących ochrony danych, i podkreślono znaczenie „ochrony danych w fazie projektowania i jako opcji domyślnej” oraz „oceny skutków w zakresie ochrony danych”, aby zapewnić uwzględnienie kwestii ochrony danych, zanim dane osobowe zostaną udostępnione do ponownego wykorzystania.

W sekcji V poprzez odpowiednie przykłady przedstawiono wytyczne dotyczące rodzaju danych, które mogą wchodzić w zakres dyrektywy ISP.

W sekcji VI skoncentrowano się w szczególności na sytuacjach, które są obecnie najbardziej powszechne, jeżeli chodzi o inicjatywy dotyczące ponownego wykorzystywania informacji sektora publicznego: tj. sytuacjach, w których zbiorcze dane statystyczne uzyskane z danych osobowych

⁹ Opinia EIOD z dnia 18 kwietnia 2012 r. w sprawie pakietu dotyczącego otwartego dostępu do danych, obejmującego wnioski dotyczący dyrektywy zmieniającej dyrektywę 2003/98/WE w sprawie ponownego wykorzystywania informacji sektora publicznego (ISP), komunikat w sprawie otwartych danych oraz decyzję Komisji 2011/833/UE w sprawie ponownego wykorzystywania dokumentów Komisji. Dostępna na stronie internetowej: http://www.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2012/12-04-18_Open_data_EN.pdf.

¹⁰ Opinia Grupy Roboczej Art. 29 nr 3/2013 w sprawie zasady celowości przyjęta w dniu 2 kwietnia 2013 r. (WP 203).

¹¹ Przewiduje się, że opinia w sprawie tego zagadnienia zostanie przyjęta w drugiej połowie 2013 r.

¹² Zob. na przykład, kodeks praktyk w zakresie anonimizacji: „Anonymisation: Managing data protection risk code of practice” wydany przez Biuro Rzecznika ds. Informacji w Zjednoczonym Królestwie w listopadzie 2012 r. oraz wytyczne w zakresie analizy ryzyka wydane przez francuski urząd ochrony danych w czerwcu 2012 r.

¹³ Zob. na przykład zalecenie polityczne LAPSI z września 2012 r. (s. 4–14).

¹⁴ Zob. art. 6 ust. 1 lit. b) dyrektywy 95/46/WE.

udostępnia się w postaci zbiorczej i zanonimizowanej. Te zbiorcze dane statystyczne dotyczą na przykład poziomu przestępczości, wydatków publicznych lub wyników w nauce osiągniętych przez dzieci w wieku szkolnym w różnych regionach geograficznych lub w różnych instytucjach oświatowych. Ponieważ jest to najpowszechniejszy scenariusz ponownego wykorzystania informacji sektora publicznego zawierających dane osobowe, znaczna część niniejszej opinii zostanie poświęcona właśnie temu scenariuszowi. W tym przypadku głównym zagadnieniem związanym z ochroną danych jest zapewnienie efektywnej agregacji i anonimizacji oraz ograniczenie do minimum ryzyka, że możliwe będzie ponowne zidentyfikowanie jakichkolwiek danych osobowych ze zagregowanych zbiorów danych.

W sekcji VII omawia się pokrótce sytuacje, w których dane osobowe są udostępniane publicznie i w związku z tym mogą potencjalnie być dostępne do ponownego wykorzystania. Chociaż obecnie nie jest to typowy scenariusz w odniesieniu do inicjatyw dotyczących ponownego wykorzystywania ISP, ważne jest uwzględnienie faktu, że organy sektora publicznego udostępniają publicznie coraz większą ilość danych osobowych, często za pośrednictwem internetu. Dotyczy to często danych osobowych, które mogą zostać bezpośrednio zidentyfikowane, takie jak na przykład informacje z rejestru gruntów na temat właściciela danej nieruchomości, oświadczenia o braku konfliktu interesów lub o wynagrodzeniu wypłacanym określonym urzędnikom bądź wydatkach członków parlamentu. Pojawia się tu pytanie, do jakiego stopnia, w jakim celu, pod jakimi warunkami i z zastosowaniem jakich zabezpieczeń można udostępnić przedmiotowe dane do ponownego wykorzystania. Ważne jest również wyjaśnienie, czy dane te podlegają przepisom dyrektywy ISP.

W tym kontekście istotne jest podkreślenie, że wszelkie informacje odnoszące się do zidentyfikowanej lub możliwej do zidentyfikowania osoby fizycznej, niezależnie od tego, czy są dostępne publicznie, stanowią dane osobowe. W związku z tym dostęp do danych osobowych, które zostały udostępnione publicznie, i ponowne wykorzystywanie tych danych (np. przez opublikowanie ich w internecie) nadal podlega właściwym przepisom o ochronie danych.

Niektóre inne szczegółowe scenariusze, takie jak przypadek wyników badań naukowych i sytuacja archiwów historycznych, które obecnie wchodzą w zakres dyrektywy ISP, zostaną krótko omówione w sekcji VIII i IX.

W sekcji X omawia się kwestię licencjonowania ISP w stosownych przypadkach oraz potrzebę integracji klauzuli o ochronie danych z licencjami.

Ponadto w pkt XI przedstawiono szereg wniosków i zaleceń.

IV. Nie wszystkie „dostępne publicznie” dane osobowe powinny zostać udostępnione do ponownego wykorzystania

4.1. Obowiązek ponownego wykorzystywania na podstawie dyrektywy ISP pozostaje bez uszczerbku dla wymogów dotyczących ochrony danych

Przyjęta w 2003 r. dyrektywa ISP nie nakładała na organy sektora publicznego obowiązku zezwolenia na ponowne wykorzystywanie ISP. Decyzja w sprawie zezwolenia na ponowne wykorzystywanie pozostawała w gestii zainteresowanych państw członkowskich lub organów sektora publicznego (z zastrzeżeniem krajowych ram prawnych w zakresie przejrzystości i dostępu). Opinia nr 7/2003 została przyjęta w świetle tego „braku obowiązku”. W sekcji 2 ppkt cc) opinii nr 7/2003 stwierdza się, że: „należy podkreślić, że nie można przywoływać dyrektywy w sprawie ponownego wykorzystywania jako obowiązku prawnego, który należy spełnić, ponieważ w przedmiotowej dyrektywie nie ustanawia się obowiązku ujawniania informacji osobowych”.

W przypadku dyrektywy zmieniającej dyrektywę ISP analiza jest bardziej złożona, lecz wniosek końcowy pozostaje taki sam.

W art. 3 ust. 1 zmienionej dyrektywy ISP stwierdza się, że: „z zastrzeżeniem ust. 2, państwa członkowskie zapewniają możliwość ponownego wykorzystywania dokumentów, do których niniejsza dyrektywa ma zastosowanie zgodnie z art. 1, do celów komercyjnych lub niekomercyjnych zgodnie z warunkami określonymi w rozdziałach III i IV”. Jeżeli nie można odmówić ponownego wykorzystania z powodów określonych w art. 1 (powodów związanych z krajowymi systemami dostępu i w szczególności również ze względu na ochronę danych osobowych), należy zezwolić na ponowne wykorzystanie.

Jednocześnie w motywie 21 dyrektywy ISP stwierdza się, że dyrektywa ISP „powinna być wykonywana i stosowana w pełnej zgodności z regułami odnoszącymi się do ochrony danych osobowych”. Ponadto w art. 1 ust. 4 określa się, że dyrektywa ISP „pozostawia nienaruszony i w żaden sposób nie wpływa na poziom ochrony osób fizycznych w odniesieniu do przetwarzania danych osobowych”.

Z łącznej lektury przedmiotowych przepisów wynika, że w przypadku, gdy chodzi o prawo do ochrony danych osobowych, „zasada ponownego wykorzystywania” nie jest stosowana automatycznie i nie ma charakteru nadrzędnego wobec właściwych przepisów o ochronie danych. Jeżeli istniejące dokumenty będące w posiadaniu organów sektora publicznego zawierają dane osobowe, ich ponowne wykorzystanie wchodzi w zakres stosowania dyrektywy 95/46/WE i w związku z tym w dalszym ciągu podlega odpowiednim przepisom o ochronie danych.

W związku z tym w przypadkach, w których ponowne wykorzystanie dotyczy danych osobowych, organ sektora publicznego nie może systematycznie powoływać się na konieczność zapewnienia zgodności z dyrektywą ISP jako zgodną z prawem podstawę udostępniania danych do ponownego wykorzystania¹⁵.

4.2. Znaczenie oceny skutków w zakresie ochrony danych przed otwarciem danych w celu ponownego wykorzystania

Uwzględniając potencjalne czynniki ryzyka przy ponownym wykorzystywaniu ISP, w szczególności fakt, że po publicznym udostępnieniu danych osobowych w celu ponownego wykorzystania sprawowanie skutecznej kontroli nad tymi danymi będzie znacznie utrudnione, Grupa Robocza Art. 29 podkreśla konieczność przestrzegania zasad dotyczących „ochrony danych już w fazie projektowania oraz ochrony danych jako opcji domyślnej”, a także konieczność odniesienia się do problemów związanych z ochroną danych na wczesnym etapie. W szczególności Grupa Robocza Art. 29 zdecydowanie zaleca, aby organ sektora publicznego przeprowadził dogłębną ocenę skutków w zakresie ochrony danych, zanim udostępni określone dane osobowe do ponownego wykorzystania. Państwa członkowskie powinny również rozważyć wprowadzenie obowiązku przeprowadzania tego rodzaju oceny skutków w ramach przepisów krajowych lub promowanie tej oceny jako najlepszej praktyki. W każdym razie, nawet jeżeli nie zostało to wyraźnie przewidziane w przepisach krajowych, przed ujawnieniem informacji i podjęciem decyzji o udostępnieniu ich w celu ponownego wykorzystania organy sektora publicznego powinny przeprowadzić dogłębną ocenę w celu ustalenia, czy określone dane osobowe mogą zostać udostępnione do ponownego

¹⁵ Grupa Robocza Art. 29 pragnie także doprecyzować, że z punktu widzenia ponownego użytkownika dyrektywa ISP sama w sobie nie stwarza zgodnej z prawem podstawy do przetwarzania danych. (W odniesieniu do podstaw prawnych zob. opinia 7/2003 oraz sekcja 7.5 poniżej.)

wykorzystania, a jeżeli mogą, należy ustalić, na jakich warunkach ponowne wykorzystanie jest dozwolone oraz jakie są wymagane konkretne zabezpieczenia służące ochronie danych,.

W ocenie powinno się wskazać między innymi podstawę prawną ujawnienia danych (i ewentualną podstawę prawną ich ponownego wykorzystania), przeanalizować zasady w zakresie celowości, proporcjonalności oraz minimalizacji danych, a także uwzględnić konieczność specjalnej ochrony danych wrażliwych. W trakcie przeprowadzania tej oceny należy starannie uwzględnić możliwy wpływ na osoby, których dane dotyczą.

Wspomniana ocena powinna ułatwić podjęcie decyzji o tym, które dane osobowe, jeżeli w ogóle, mogą być udostępniane do ponownego wykorzystania i z zastosowaniem jakich zabezpieczeń¹⁶. Należy podkreślić, że w proponowanym rozporządzeniu o ochronie danych¹⁷ zaleca się sporządzanie ocen skutków w zakresie ochrony danych jako kluczowego narzędzia pomocnego w zapewnianiu odpowiedzialności administratorów, a w niektórych przypadkach nawet wymaga się ich sporządzania¹⁸.

Tam, gdzie jest to możliwe, analiza przeprowadzana przed podjęciem decyzji o ponownym wykorzystaniu powinna opierać się na rzeczowej debacie z udziałem przedstawicieli różnych zainteresowanych stron, w tym administratora pragnącego udostępnić dane oraz tych, którzy żądają dostępu do danych i którzy w związku z tym mogą przedstawić kontekst dyskusji, a także przedstawicieli osób fizycznych, których przedmiotowe dane osobowe dotyczą (na przykład organizacje zajmujące się ochroną konsumentów, organizacje zajmujące się prawami pacjentów, związki zawodowe nauczycieli). W przypadkach, w których wynik nie jest oczywisty, właściwe organy ochrony danych i organy krajowe odpowiedzialne za dostęp do informacji publicznej mogą mieć możliwość przedstawienia wytycznych.

Państwa członkowskie powinny rozważyć także ustanowienie i wspieranie sieci wiedzy / centrów doskonałości, a tym samym umożliwić wymianę dobrych praktyk w zakresie anonimizacji i otwartych danych. Może być to szczególnie istotne dla mniejszych organów sektora publicznego,

¹⁶ W przypadku gdy ocena prowadzi do decyzji o odmowie udostępnienia do ponownego wykorzystywania samych danych osobowych oraz jeśli podjęta zostanie decyzja o udostępnieniu zanonimizowanych zbiorów uzyskanych z danych osobowych, powinno się przeprowadzić ocenę ryzyka ponownej identyfikacji. Dalsze informacje na temat anonimizacji i ponownej identyfikacji przedstawiono w sekcji VI niniejszej opinii.

¹⁷ W dniu 25 stycznia 2012 r. Komisja przyjęła pakiet dokumentów na potrzeby reformy europejskich ram ochrony danych. Pakiet ten obejmuje (i) „komunikat” (COM(2012)9 final), (ii) „proponowane rozporządzenie o ochronie danych” (COM(2012)11 final oraz (iii) „proponowaną dyrektywę o ochronie danych” (COM(2012)10 final).

¹⁸ W celu uzyskania dalszych wytycznych na temat sposobu przeprowadzania oceny skutków w zakresie ochrony danych zob. na przykład strony internetowe projektu PIAF dotyczącego ram oceny skutków w zakresie prywatności w odniesieniu do prawa o ochronie danych i prawa do prywatności (*A Privacy Impact Assessment Framework for data protection and privacy rights*): <http://www.piafproject.eu/Index.html>. PIAF jest projektem współfinansowanym przez Komisję Europejską, którego celem jest zachęcanie UE i jej państw członkowskich do przyjęcia postępowej polityki oceny skutków w zakresie ochrony danych jako środka mającego na celu zaspokojenie potrzeb i pokonanie wyzwań związanych z prywatnością i przetwarzaniem danych osobowych. Wytyczne dostępne są także w niektórych państwach członkowskich. Zob. na przykład podręcznik oceny skutków w zakresie ochrony danych (PIA) wydany przez rzecznika ds. informacji w Zjednoczonym Królestwie, dostępny na stronie internetowej: http://ico.org.uk/for_organisations/data_protection/topic_guides/privacy_impact_assessment; wytyczne w zakresie analizy ryzyka wydane przez francuski organ ochrony danych, wspomniane już w przypisie 12 powyżej; oraz wytyczne słoweńskiego rzecznika ds. informacji, szczególnie w odniesieniu do oceny skutków w zakresie ochrony danych w projektach dotyczących administracji elektronicznej, dostępne na stronie internetowej: https://webmail.europarl.europa.eu/exchweb/bin/redirect.asp?URL=https://www.ip-rs.si/fileadmin/user_upload/Pdf/smernice/PIASmernice__ENG_Lektorirano_10._6._2011.pdf.

którym może brakować wiedzy specjalistycznej do celów przeprowadzania anonimizacji, oceny skutków w zakresie ochrony danych oraz oceny i badania ryzyka ponownej identyfikacji¹⁹.

Ponadto ocena skutków jest także zdecydowanie zalecana przed wprowadzeniem nowych przepisów przewidujących podawanie informacji do wiadomości publicznej.

V. Zakres dyrektywy ISP: wyjątki ze względu na ochronę danych osobowych

W niniejszym punkcie przedstawiono wytyczne dotyczące zakresu stosowania dyrektywy ISP, a w szczególności wyjątków ze względu na ochronę danych.

5.1. Stosowanie ogólnych ram ochrony danych do ponownego wykorzystania ISP

W motywie 21 dyrektywy ISP stwierdza się, że dyrektywa ISP: „powinna być wykonywana i stosowana w pełnej zgodności z regułami odnoszącymi się do ochrony danych osobowych”. Ponadto w art. 1 ust. 4 określa się, że dyrektywa ISP: „pozostawia nienaruszony i w żaden sposób nie wpływa na poziom ochrony osób fizycznych w odniesieniu do przetwarzania danych osobowych”.

5.2. Wyjątki ze względu na ochronę danych osobowych

W dyrektywie ISP stwierdza się, że: „niniejsza dyrektywa nie ma zastosowania do: [...] dokumentów wyłączonych z dostępu na podstawie systemów dostępu państw członkowskich [...]”²⁰.

Oprócz tego w dyrektywie ISP, z późniejszymi zmianami, przedstawia się również wyjątki ze względu na ochronę danych. W art. 1 ust. 2 lit. cc) uwzględnia się następujące trzy sytuacje, z których wszystkie są wyłączone z zakresu stosowania dyrektywy ISP:

- dokumenty wyłączone z dostępu na podstawie systemów dostępu z powodu ochrony danych osobowych;
- dokumenty, do których dostęp jest ograniczony na podstawie systemów dostępu z powodu ochrony danych osobowych; i
- „części dokumentów dostępne na podstawie tych systemów, które to części zawierają dane osobowe, których ponowne wykorzystywanie zostało określone w przepisach jako niezgodne z prawem dotyczącym ochrony osób fizycznych w zakresie przetwarzania danych osobowych”.

5.3. Uwagi ogólne

Grupa Robocza Art. 29 podkreśla, że bez względu na „zasadę ponownego wykorzystania danych” sformułowaną w dyrektywie zmieniającej dyrektywę ISP, ponowne wykorzystanie danych do celów komercyjnych lub niekomercyjnych zgodnie z warunkami określonymi w dyrektywie ISP nie zawsze jest właściwe, jeżeli ISP, które mają być ponownie wykorzystane, zawierają dane osobowe. Decyzje dotyczące ponownego wykorzystywania danych osobowych zgodnie z warunkami

¹⁹ Przykładowo w Zjednoczonym Królestwie konsorcjum pod przewodnictwem Uniwersytetu w Manchesterze razem z Uniwersytetem w Southampton, Biurem Statystyk Krajowych (Office for National Statistics) i nowym, należącym do rządu Instytutem Otwartych Danych (Open Data Institute, ODI) prowadzi sieć anonimizacyjną Zjednoczonego Królestwa (UK Anonymisation Network, UKAN), aby umożliwić wymianę dobrych praktyk w zakresie anonimizacji w całym sektorze prywatnym i publicznym. Sieć ta obejmuje stronę internetową <https://webmail.europarl.europa.eu/exchweb/bin/redir.asp?URL=http://www.ukanon.net>, analizy przykładów, poradnie i seminaria.

²⁰ Zob. art. 1 ust. 2 lit. c) dyrektywy ISP.

określonymi w dyrektywie ISP należy podejmować indywidualnie w poszczególnych przypadkach; istnieje także potrzeba wprowadzenia dodatkowych środków prawnych, technicznych bądź organizacyjnych w celu ochrony zainteresowanych osób fizycznych.

Ponowne wykorzystywanie publicznie dostępnych danych osobowych jest ograniczone i powinno być ograniczane przez:

- przepisy ogólne stanowiące część mających zastosowanie przepisów o ochronie danych;
- szczegółowe dodatkowe ograniczenia prawne (w stosownych przypadkach); oraz
- zabezpieczenia techniczne i organizacyjne, które wprowadzono w celu ochrony danych osobowych.

5.4 Dokumenty wylączone z dostępu

W ramach tego przepisu z zakresu stosowania dyrektywy ISP wyłącza się wszystkie dokumenty, które są wyłączone na podstawie systemów dostępu w zainteresowanym państwie członkowskim ze względu na ochronę danych osobowych.

W przeciwieństwie do przepisów o ochronie danych, które są w dużym stopniu zharmonizowane na podstawie dyrektywy 95/46/WE, dostęp do przepisów dotyczących informacji różni się znacząco w poszczególnych państwach członkowskich UE. Systemy dostępu zazwyczaj wymagają przeprowadzenia testu bilansującego, w którym porównuje się interesy chronione w ramach przepisów w zakresie prywatności i ochrony danych z korzyściami wynikającymi z zapewnienia otwartości i przejrzystości. Uwzględniając rozbieżności, wynik działań bilansujących może być różny w poszczególnych państwach członkowskich UE. Przykładowo organy podatkowe w niektórych państwach członkowskich mogą publikować określone części deklaracji podatkowych w zakresie podatku dochodowego składane przez podatników (z zastrzeżeniem środków prawnych, technicznych i organizacyjnych służących zminimalizowaniu ryzyka niewłaściwego wykorzystania), podczas gdy inne państwo członkowskie uznałoby to za informację, która podlega wyjątkowi i zasadniczo nie powinna być ujawniana.

W związku z tym przepisy krajowe muszą być zgodne z art. 8 europejskiej Konwencji o ochronie praw człowieka i podstawowych wolności („EKPC”) oraz art. 7 i art. 8 Karty praw podstawowych Unii Europejskiej („Karta praw podstawowych UE”). Oznacza to, iż zgodnie z orzeczeniem Europejskiego Trybunału Sprawiedliwości w sprawie *Österreichischer Rundfunk i Schecke*²¹, należy ustalić, że ujawnienie jest konieczne i proporcjonalne do celu zgodnego z prawem, do którego dąży się w ramach przepisów.

W każdym razie po wyłączeniu z dostępu danych osobowych zawartych w dokumencie na mocy przepisów odpowiedniego państwa członkowskiego (w tym w sytuacjach, w których w przepisach krajowych dotyczących przejrzystości i otwartości nie przewiduje się ogólnej dostępności przedmiotowych danych osobowych) będą one także wyłączone z zakresu stosowania dyrektywy ISP.

W celu zapewnienia pewności prawa i przejrzystości w odniesieniu do osób, których dane dotyczą, dobrą praktyką jest przyjęcie w miarę możliwości podejścia proaktywnego i określenie z góry

²¹ Zob. wyrok Europejskiego Trybunału Sprawiedliwości z dnia 20 maja 2003 r. w sprawach połączonych C-465/00, C-138/01 i C-139/01 *Rundfunk* oraz wyrok z dnia 9 listopada 2010 r. w sprawach połączonych C-92/09 i C-93/09, *Volker und Markus Schecke*..

danych osobowych, które mogłyby być udostępniane publicznie. Osoby, których dane dotyczą, mogą być następnie informowane podczas gromadzenia danych o tym, czy jakkolwiek część tych danych osobowych, które zostały przez nie dostarczone lub które będą dalej przetwarzane w procedurze administracyjnej, zostanie udostępniona publicznie w wyniku obowiązywania przepisów w zakresie dostępu do informacji publicznej.

5.5 Dokumenty, do których dostęp jest ograniczony

W ramach tego przepisu z zakresu stosowania dyrektywy ISP wyłącza się wszystkie dokumenty, do których dostęp jest ograniczony na podstawie systemów dostępu w zainteresowanym państwie członkowskim ze względu na ochronę danych osobowych. Podobnie jak w sytuacji opisywanej powyżej, systemy dostępu w poszczególnych państwach członkowskich mogą różnić się pod względem danych, które mogą podlegać ograniczonemu dostępowi, oraz rodzajów ograniczeń. Przykłady takich dokumentów obejmują:

- zbiory archiwów krajowych zawierające dane osobowe, które są dostępne jedynie na szczegółowych warunkach dostępu i z zastrzeżeniem dodatkowych zabezpieczeń (zob. sekcja IX poniżej);
- zbiory wyników badań naukowych zawierające dane osobowe, które są dostępne jedynie na szczegółowych warunkach dostępu i z zastrzeżeniem dodatkowych zabezpieczeń (zob. sekcja VIII poniżej);
- określone informacje znajdujące się w rejestrach publicznych, aktach sądowych lub innych dokumentach administracyjnych zawierające dane osobowe, które mogą być dostępne jedynie dla osób fizycznych lub organizacji wykazujących uzasadniony interes lub jedynie na innych szczegółowych warunkach dostępu i z zastrzeżeniem dodatkowych zabezpieczeń.

5.6 Części dokumentów dostępnych, w przypadku których ponowne wykorzystywanie jest niezgodne z prawem

W ramach tego przepisu z zakresu stosowania dyrektywy ISP wyłącza się

- części dokumentów,
- dostępne na podstawie krajowych systemów dostępu,
- które zawierają dane osobowe, „których ponowne wykorzystywanie zostało określone w przepisach jako niezgodne z prawem dotyczącym ochrony osób fizycznych w zakresie przetwarzania danych osobowych”.

W przepisie tym potwierdza się, że nawet jeżeli określone dokumenty zawierające dane osobowe są w pełni dostępne, ich ponowne wykorzystywanie może jednak być ograniczone ze względu na ochronę danych.

Grupa Robocza Art. 29 podkreśla, że przedmiotowy przepis dyrektywy ISP powinno się interpretować zgodnie z art. 1 ust. 4 dyrektywy ISP, który stanowi, że dyrektywa ISP: „pozostawia nienaruszony i w żaden sposób nie wpływa na poziom ochrony osób fizycznych w odniesieniu do przetwarzania danych osobowych”.

Grupa Robocza Art. 29 z zadowoleniem przyjęłaby jako dobrą praktykę przyjęcie szczegółowych przepisów krajowych, które jasno wskazują: (i) które dane są udostępniane publicznie; (ii) do jakich celów; oraz (iii) w stosownych przypadkach określają, w jakim stopniu i na jakich warunkach ponowne wykorzystywanie jest dozwolone. Brak szczegółowych przepisów nie oznacza jednak, że

dostępne publicznie dane osobowe mogą być zawsze ponownie wykorzystywane na podstawie dyrektywy ISP.

Natomiast w takich przypadkach w przepisach o ochronie danych (stosowanych wraz z innymi odpowiednimi przepisami, takimi jak przepisy dotyczące dostępu do dokumentów) określa się, czy dane osobowe mogą w określonych przypadkach być udostępniane do celów ponownego wykorzystywania i jeżeli tak, to z zastrzeżeniem jakich dodatkowych zabezpieczeń. Jeżeli wynik tej oceny jest pozytywny, zezwala się na ponowne wykorzystanie, z zastrzeżeniem szczegółowych zabezpieczeń dotyczących ochrony danych i wszystkich innych warunków określonych w dyrektywie ISP (o ile pozostają one bez uszczerbku dla przepisów o ochronie danych). Jeżeli wynik oceny jest negatywny, ponowne wykorzystanie nie będzie wchodzić w zakres dyrektywy ISP.

Poniższe przykłady mogą pomóc w zilustrowaniu, w jakich sytuacjach należy stosować przedmiotowe wyłączenie z zakresu dyrektywy ISP. W pierwszym przykładzie ograniczenia ponownego wykorzystywania są wyraźnie określone w przepisach.

- Przepisy podatkowe w danym państwie członkowskim mogą stanowić, że deklaracje podatkowe w zakresie podatku dochodowego wszystkich rezydentów danego państwa są publicznie dostępne w siedzibie organu podatkowego do wglądu innych rezydentów, na ich wniosek, bez konieczności wykazania uzasadnionego interesu prawnego. W przepisach wyraźnie określono również, że dane nie mogą być poddawane dalszemu przetwarzaniu, na przykład publikowane w internecie, łączone z innymi danymi lub poddawane dalszej redakcji. Organizacja pozarządowa wnioskuje o udzielenie dostępu do bazy danych deklaracji podatkowych i o prawo do jej ponownego wykorzystania w celu opublikowania wyżej wymienionych danych na swojej stronie internetowej. W tym przypadku dane podatkowe nie wchodzą w zakres dyrektywy ISP i na organ sektora publicznego nie nałożono obowiązku udostępnienia zbioru danych do ponownego wykorzystywania na podstawie dyrektywy ISP.

W wielu innych przypadkach ograniczenia prawne mogą jednak być mniej wyraźnie określone i mogą mniej kategoryczne pod względem ponownego wykorzystywania. W przypadku różnych rejestrów obywatelskich, handlowych i rejestrów ludności oraz innych baz danych zazwyczaj zezwala się społeczeństwu na wgląd w dane osobowe, coraz częściej w postaci cyfrowej za pośrednictwem internetu. Dostępność jest obwarowana określonymi zabezpieczeniami, w tym ograniczeniom technicznym w zakresie możliwości wyszukiwania i masowego pobierania danych. Użytkownicy mogą zostać poproszeni o wyrażenie zgody na warunki uzyskiwania dostępu.

- Przepisy podatkowe w państwie członkowskim mogą stanowić, że nazwiska rezydentów, którzy mają zaległości podatkowe powyżej określonego progu przez dłuższy czas, będą przez pewien okres publikowane na specjalnej stronie internetowej, z zastrzeżeniem dodatkowych zabezpieczeń technicznych, w tym ograniczeń w zakresie masowego pobierania danych i możliwości wyszukiwania. Tego rodzaju upublicznienie danych ma na celu zachęcanie do terminowego płacenia podatku dochodowego i służy jako dodatkowa (reputacyjna) kara wobec tych, którzy nie wywiązują się z tego obowiązku. Konsorcjum banków wnioskuje o dostęp do danych w celu ponownego wykorzystywania, aby wprowadzić je do swojego systemu informacji kredytowej.
- W przepisach szczegółowych w sektorze służby zdrowia w państwie członkowskim, z zastrzeżeniem zabezpieczeń, można umożliwić pacjentom sprawdzenie na specjalnej stronie internetowej, czy konkretny lekarz lub inny członek personelu medycznego ma zakaz wykonywania zawodu. Zastosowano zabezpieczenia techniczne, w tym ograniczenia w zakresie masowego pobierania danych i możliwości wyszukiwania. Organizacja praw

pacjentów ubiega się o dostęp do danych w celu ponownego wykorzystywania, aby utworzyć wielojęzyczną i bardziej przyjazną dla użytkownika stronę internetową, za pomocą której będzie można uzyskać dostęp do tych samych danych.

- Przepisy szczegółowe w państwie członkowskim mogą wymagać publikacji nazwisk darczyńców na rzecz partii politycznych powyżej określonego progu. Informacje, które mogą ujawniać opinie polityczne wspomnianych darczyńców, są podawane do wiadomości publicznej za pośrednictwem specjalnej strony internetowej. Zastosowano zabezpieczenia techniczne, w tym ograniczenia w zakresie masowego pobierania danych i możliwości wyszukiwania. Grupa aktywistów zwraca się z wnioskiem o masowy dostęp do danych w celu ponownego wykorzystywania na podstawie dyrektywy ISP, aby utworzyć nową stronę internetową z dodatkowymi funkcjami i lepszymi możliwościami wyszukiwania.
- Nazwisko i adres właściciela nieruchomości są dostępne publicznie w rejestrze gruntów państwa członkowskiego, lecz przeglądanie dostępnej publicznie bazy danych jest ograniczone, aby możliwe było jedynie wyszukiwanie określonej nieruchomości, a nie określonej osoby. Masowe pobieranie danych jest również ograniczone. Spółka prawa handlowego zwraca się z wnioskiem o masowy dostęp do danych w celu ponownego wykorzystania, aby utworzyć stronę internetową bardziej przyjazną dla użytkownika po bardziej konkurencyjnej cenie.
- W rejestrach przedsiębiorstw w danym państwie członkowskim dopuszcza się udostępnianie szeregu danych osobowych, łącznie z nazwiskami, adresami i wzorami podpisów członków zarządu, oraz informacje dotyczące własności określonych rodzajów spółek. Istnieją pewne ograniczenia w odniesieniu do możliwości wyszukiwania i liczby elementów, które mogą zostać pobrane. Informacje te są dostępne na specjalnej stronie internetowej pod warunkiem uiszczenia opłaty. Spółka prawa handlowego zwraca się z wnioskiem o masowy dostęp do danych w celu ponownego wykorzystania, aby utworzyć stronę internetową, na której zebrane zostaną informacje z kilku rodzajów rejestrów, i zaoferować rozszerzone informacje po bardziej konkurencyjnej cenie.

We wszystkich przypadkach zainteresowany organ sektora publicznego musi dokonać starannej oceny skutków w zakresie ochrony danych, aby podjąć decyzję, czy dane mogą zostać udostępnione do ponownego wykorzystywania na podstawie dyrektywy ISP, a jeżeli tak, to czy w przepisach o ochronie danych wymaga się jakichkolwiek szczególnych warunków i zabezpieczeń. „Zasada ponownego wykorzystywania” nie jest stosowana automatycznie i nie ma charakteru nadrzędnego wobec właściwych przepisów o ochronie danych.

Taka staranna ocena jest jeszcze istotniejsza ze względu na fakt, że na podstawie dyrektywy ISP organ sektora publicznego zasadniczo nie może rozważać, kim jest dany ponowny użytkownik wnioskujący o dostęp. Zgodnie z art. 10 (Brak dyskryminacji) „żadne stosowane warunki ponownego wykorzystywania dokumentów nie mogą dyskryminować porównywalnych kategorii ponownego wykorzystywania”. Ponadto zgodnie z art. 11 (Zakaz umów o wyłączności) „ponowne wykorzystywanie dokumentów jest otwarte dla wszystkich potencjalnych uczestników rynku [...]. Kontrakty i inne umowy między organami sektora publicznego będącymi w posiadaniu dokumentów i stronami trzecimi nie mogą udzielać praw wyłącznych”.

W związku z tym podczas podejmowania decyzji w sprawie zezwolenia na ponowne wykorzystanie organy sektora publicznego muszą uwzględnić zgodność dopuszczania ponownego wykorzystania w ramach otwartej licencji, nie tylko z wnioskodawcą, ale także ze wszystkimi, którzy zwracają się o udostępnienie danych. Wymaga to wysokiego stopnia pewności, że żaden z potencjalnych ponownych użytkowników nie będzie mógł wykorzystać udostępnionych danych osobowych w niewłaściwy sposób.

Dyrektywa ISP nie wyklucza możliwości, by nałożone w tym względzie warunki dopuszczały przetwarzanie wyłącznie w określonych celach. Organ sektora publicznego musi zatem odpowiedzieć na pytanie, czy ponowne wykorzystanie danych przez każdego „potencjalnego uczestnika rynku” do tych celów jest zgodne z celami określonymi przez dany organ sektora publicznego. Potencjalne ponowne wykorzystanie przez instytucje finansowe informacji dotyczących zapłaty podatków na przykład do celów sprawozdawczości kredytowej jest istotne z uwagi na fakt, iż nadal są one potencjalnymi ponownymi użytkownikami, zgodnie z definicją „każdej osoby”. W związku z tym, aby rozwiązać wątpliwości dotyczące ochrony danych, a w szczególności aby zapewnić przestrzeganie zasady celowości, organ sektora publicznego (lub prawodawca) musi mieć możliwość ograniczania w stosownych przypadkach celów ponownego wykorzystania danych.

VI. Ponowne wykorzystanie zagregowanych i zanonimizowanych zbiorów danych uzyskanych z danych osobowych

6.1. Jakie korzyści płyną z agregacji i anonimizacji dla ponownego wykorzystywania ISP?

Do tej pory celem inicjatyw ponownego wykorzystywania ISP realizowanych przez organy sektora publicznego poprzez „portale otwartych danych” lub inne platformy było zwykle udostępnianie do ponownego wykorzystania danych zagregowanych i zanonimizowanych, a nie danych osobowych jako takich. Takie podejście jest rzeczywiście bezpieczniejsze i powinno być promowane.

Przepisy o ochronie danych zazwyczaj nie zezwalają na publiczne ujawnianie przez organy sektora publicznego danych osobowych, które zostały zebrane w innym celu, zwykle o charakterze administracyjnym²². W związku z tym w takich przypadkach ich ponowne wykorzystanie w ramach inicjatyw ponownego wykorzystywania ISP również nie jest możliwe. Zwykle to dane statystyczne uzyskane z danych osobowych, a nie same dane osobowe są i powinny, co do zasady, być udostępniane do ponownego wykorzystania. Jest to najskuteczniejsze rozwiązanie pozwalające zminimalizować ryzyko niezamierzonego ujawnienia danych osobowych. Takie zanonimizowane i zagregowane zbiory danych powinny uniemożliwiać ponowną identyfikację osób fizycznych, a zatem nie powinny zawierać danych osobowych.

Podjęcie decyzji co do tego, jaki poziom agregacji może być odpowiedni i jakich szczególnych technik anonimizacji należy użyć, stanowi duże wyzwanie. Jeżeli agregacja i anonimizacja nie zostaną przeprowadzone skutecznie, wiąże się to z ryzykiem, że osoby fizyczne mogą jednak zostać ponownie zidentyfikowane na podstawie tych zbiorów danych. W związku z tym przepisy o ochronie danych odgrywają ważną rolę w ustaleniu „bezpiecznego” progu udostępnienia zanonimizowanych i zagregowanych danych w ramach inicjatywy ISP.

W dyrektywie 95/46/WE ustala się wysoką wartość progową anonimizacji.

Użyty w tym dokumencie termin „anonimizacja” odnosi się do danych, które nie mogą być już uznane za dane osobowe na podstawie art. 2 lit. a) dyrektywy 95/46/WE. W art. 2 lit. a) definiuje się „dane osobowe” jako: „wszelkie informacje dotyczące zidentyfikowanej lub możliwej do zidentyfikowania osoby fizycznej (»osoby, której dane dotyczą«); osoba możliwa do zidentyfikowania to osoba, której tożsamość można ustalić bezpośrednio lub pośrednio, szczególnie przez powołanie się na numer identyfikacyjny lub jeden bądź kilka szczególnych czynników

²² W stosownych przypadkach prawodawstwo w zakresie dostępu do informacji publicznej może oczywiście wymagać ujawnienia danych osobowych, a interes polegający na zapewnieniu przejrzystości i dostępności informacji w niektórych przypadkach może mieć charakter nadrzędny wobec problemów dotyczących ochrony danych i prywatności. Jest to obszar nadal rozwijający się, który może wywołać zmiany w przyszłości.

określających jej fizyczną, fizjologiczną, umysłową, ekonomiczną, kulturową lub społeczną tożsamość”²³.

Motyw 26 dyrektywy 95/46/WE także odnosi się do tej kwestii i stanowi, że: „w celu ustalenia, czy daną osobę można zidentyfikować, należy wziąć pod uwagę wszystkie sposoby, jakimi może posłużyć się administrator lub inna osoba w celu zidentyfikowania owej osoby”.

Należy podkreślić, że w ten sposób ustala się wysoki próg, co zostanie omówione w dalszej części niniejszej opinii. Przepisy o ochronie danych nadal mają zastosowanie, chyba że możliwe jest zanonimizowanie danych w stopniu pozwalającym na osiągnięcie tego progu. Oznacza to między innymi, że o ile przedmiotowy próg nie zostanie osiągnięty, publiczne udostępnianie informacji (i jakiegokolwiek dalsze ich wykorzystywanie) musi pozostawać w zgodzie z początkowym celem gromadzenia danych na podstawie art. 6 ust. 1 lit b) dyrektywy 95/46/WE. Dodatkowo musi również istnieć właściwa podstawa prawna do przetwarzania zgodnie z art. 7 lit a)–f) dyrektywy 95/46/WE (na przykład wyrażenie zgody lub konieczność zachowania zgodności z prawem). Jeżeli natomiast dane zostały zanonimizowane w rozumieniu art. 2 lit. a) i motywu 26 dyrektywy 95/46/WE, przepisy o ochronie danych nie będą już stosowane i ponowni użytkownicy mogą ponownie wykorzystywać przedmiotowe dane bez tych ograniczeń.

Raz jeszcze należy podkreślić, że „dane zanonimizowane”, w znaczeniu stosowanym w niniejszej opinii, odnoszą się do danych, które nie są już uznawane za dane osobowe. Dane zanonimizowane należy w szczególności odróżnić od danych, które zmodyfikowano przy użyciu różnych technik w celu zmniejszenia ryzyka ponownej identyfikacji osób fizycznych, których dane te dotyczą, ale które nie osiągnęły progu wymaganego w art. 2 lit. a) i w motywie 26 dyrektywy 95/46/WE²⁴. W wielu przypadkach techniki te są jedynie odpowiednie do ujawniania w ograniczonym zakresie w celu ponownego wykorzystania przez zweryfikowane osoby trzecie, ale nie do pełnego publicznego ujawniania i ponownego wykorzystania w ramach otwartych licencji.

Należy także podkreślić, że gdy dane zostaną udostępnione publicznie do ponownego wykorzystania, nie będzie już żadnej kontroli nad tym, kto może mieć do nich dostęp. Znacznie wzrośnie prawdopodobieństwo, że „każda inna osoba” będzie dysponować środkami i zastosuje te środki do celów ponownej identyfikacji osób, których dane dotyczą. W związku z tym i niezależnie od interpretacji motywu 26 w innych kontekstach, jeżeli chodzi o udostępnianie danych do ponownego wykorzystywania na podstawie dyrektywy ISP, Grupa Robocza Art. 29 pragnie jednoznacznie zaznaczyć, że należy dołożyć wszelkich starań w celu zapewnienia, aby ujawniane zbiory danych nie zawierały danych, które można ponownie zidentyfikować za pomocą środków, jakimi może posłużyć się każda osoba, w tym potencjalni ponowni użytkownicy, ale również inne podmioty, w których interesie może leżeć uzyskanie danych, w tym organy ścigania.

Dalsze wytyczne w zakresie anonimizacji i pojęcia danych osobowych

²³ W oświadczeniu wydanym dnia 27 lutego 2013 r. w sprawie „bieżących dyskusji nad pakietem reform w zakresie ochrony danych” Grupa Robocza Art. 29 podkreśliła, że: „osobę fizyczną można uznać za możliwą do zidentyfikowania, jeżeli w grupie osób można ją odróżnić od pozostałych członków grupy i w związku z tym traktować odmiennie. Oznacza to, że pojęcie możliwości zidentyfikowania wiąże się z wyróżnieniem”. W oświadczeniu wyjaśnia się również, że: „numery identyfikacyjne, dane dotyczące lokalizacji, adresy IP, identyfikatory online lub inne szczególne czynniki odnoszące się do osoby fizycznej należy uznać za dane osobowe”.

²⁴ W oświadczeniu z dnia 27 lutego 2013 r. podkreśla się, że: „jeżeli możliwe jest dotarcie do osoby fizycznej lub (pośrednio) zidentyfikowanie danej osoby fizycznej przy użyciu innych środków, przepisy o ochronie danych nadal mają zastosowanie”.

Dalsze wytyczne w zakresie anonimizacji i pojęcia danych osobowych można znaleźć w opinii 4/2007 sporządzonej przez Grupę Roboczą Art. 29 w sprawie pojęcia danych osobowych przyjętej w dniu 20 czerwca 2007 r. (WP 136). Grupa Robocza Art. 29 może również przedstawić dalsze wytyczne w zakresie technik anonimizacji w oddzielnym dokumencie w drugiej połowie 2013 r.

6.2 Wyzwania i ograniczenia związane z anonimizacją do celów ponownego wykorzystywania ISP

Przeprowadzenie anonimizacji staje się coraz trudniejsze ze względu na poziom zaawansowania nowoczesnych technologii komputerowych oraz powszechną dostępność informacji. Ponowna identyfikacja osób fizycznych stanowi coraz powszechniejsze i bardziej realne zagrożenie²⁵. W praktyce istnieje bardzo duża szara strefa, w ramach której administrator udostępniający dane może być przekonany, że zbiór danych jest zanonimizowany, natomiast osoba trzecia może nadal być w stanie zidentyfikować przynajmniej niektóre osoby fizyczne na podstawie tych danych, na przykład poprzez wykorzystanie innych informacji dostępnych publicznie lub innych informacji, do których ma ona dostęp.

Jednym z głównych czynników ryzyka jest coraz większa liczba danych online i offline, zarówno dostępnych publicznie, jak i skoncentrowanych w rękach organizacji gospodarczych, które to dane mogą następnie zostać wykorzystane do profilowania osób fizycznych do celów reklamy behawioralnej i do coraz szerszego wachlarza innych celów. Jeżeli porównuje się je z realiami „dużych zbiorów danych”, do których organizacje te mają już dostęp, uzyskane z danych osobowych i udostępnione do ponownego wykorzystania ISP może prowadzić do zwiększenia prawdopodobieństwa zidentyfikowania osób fizycznych lub dodatkowego rozbudowania ich profili, często bez ich wiedzy.

6.3. Kto i kiedy powinien przeprowadzić agregację i anonimizację?

Agregację i anonimizację należy przeprowadzić najwcześniej jak to możliwe. Czynności tych powinien dokonać administrator lub zaufana osoba trzecia działająca w imieniu administratora lub kilku administratorów (posiadająca niezbędne umiejętności specjalistyczne). Nie można pozwolić na przeprowadzanie anonimizacji przez ponownego użytkownika, na przykład w ramach warunku udzielenia licencji. Ponadto należy zapewnić, aby w przypadku gdy ewentualne przeprowadzenie agregacji i anonimizacji powierzono organizacji będącej osobą trzecią nie występował konflikt interesów oraz aby organizacja ta była wyraźnie odpowiedzialna za to, że dane osobowe będą stosowane jedynie do przeprowadzenia anonimizacji i że w tym celu zostaną zastosowane wszystkie konieczne zabezpieczenia. Osoba trzecia powinna także być w stanie zagwarantować, że dane osobowe, z których uzyskano zagregowane i zanonimizowane zbiory danych, zostaną usunięte, gdy tylko przestaną być potrzebne do tych celów.

²⁵ Zob. na przykład „Transparent government, not transparent citizens”, sprawozdanie przygotowane dla Kancelarii Rządu Zjednoczonego Królestwa przez Kierona O’Harę z Southampton University w 2011 r., w którym autor ostrzega przed możliwością identyfikacji osób fizycznych na podstawie zanonimizowanych danych, stosując między innymi „identyfikację puzzlową” (ang. *jigsaw identification*) i twierdzi, że nie istnieją żadne w pełni skuteczne rozwiązania techniczne dla problemu deanonimizacji. Sprawozdanie jest dostępne na stronie internetowej: <http://www.cabinetoffice.gov.uk/sites/default/files/resources/transparency-and-privacy-review-annex-b.pdf>. Zob. także „Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization” autorstwa Paula Ohma z University of Colorado Law School, 57 UCLA Law Review 1701 (2010 r.), dostępne na stronie internetowej: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006.

6.4 Ocena ryzyka ponownej identyfikacji

Przepisy o ochronie danych nadal mają zastosowanie, chyba że możliwe jest zanonimizowanie danych w rozumieniu art. 2 lit. a) i motywu 26 dyrektywy 95/46/WE.

Administratorzy danych powinni ocenić, czy daną osobę fizyczną można w sposób zasadny zidentyfikować na podstawie zanonimizowanego zestawu danych, który ma być udostępniany do ponownego wykorzystania, a także na podstawie innych danych. Innymi słowy – czy jakakolwiek organizacja lub osoba fizyczna mogłaby zidentyfikować którąkolwiek osobę fizyczną na podstawie ujawnianych danych – albo na podstawie samych tych danych, albo w połączeniu z innymi dostępnymi informacjami.

Jak wyjaśniono w pkt 6.1, celem niniejszej opinii nie jest przedstawienie całościowych i rozstrzygających wytycznych w zakresie sposobu oceny ryzyka ponownej identyfikacji. Nie ma ona również na celu podania rozstrzygającej definicji „anonimizacji” czy „zanonimizowanych danych”. Należy jednak przypomnieć, że czytelnik może znaleźć szczegółowe wytyczne w istniejących dokumentach (w tym dokumentach przywołanych w pkt 6.1); w podgrupie ds. technologii Grupy Roboczej Art. 29 toczą się również prace nad technikami anonimizacji, jak zauważono w sekcji 6.1 i 2.2.

Pamiętając o tym i nie dążąc do zapewnienia całościowej perspektywy, Grupa Robocza Art. 29 pragnie podkreślić niektóre z czynników/koncepcji, które są pomocne przy ocenie ryzyka ponownej identyfikacji, w tym w szczególności:

- jakie inne dane są dostępne ogółowi społeczeństwa lub innym osobom fizycznym lub organizacjom; oraz czy dane, które mają być podane do wiadomości publicznej, można powiązać z innymi zestawami danych;
- prawdopodobieństwo próby ponownej identyfikacji (niektóre rodzaje danych będą bardziej atrakcyjne dla potencjalnych intruzów niż inne); oraz
- prawdopodobieństwo, że próba ponownej identyfikacji powiedzie się, z uwzględnieniem skuteczności zaproponowanych technik anonimizacji²⁶.

Jakie są „inne” dostępne informacje?

Przy ustalaniu, czy daną osobę fizyczną można zidentyfikować pośrednio, należy rozważyć, czy identyfikacja jest możliwa za pomocą odnośnych danych (w naszym przypadku „zanonimizowanego” zestawu danych) lub na podstawie tych danych oraz *innych informacji*, które posiada lub może/prawdopodobnie będzie posiadać organizacja lub osoba fizyczna, które podejmą próbę ponownej identyfikacji.

„Inne informacje” niezbędne do ponownej identyfikacji mogą być informacjami, do których dostęp mają określone przedsiębiorstwa lub inne organizacje, w tym organy ścigania lub inne organy sektora publicznego, niektóre osoby fizyczne, lub które są dostępne każdemu, ponieważ na przykład opublikowano je w internecie. Typowym przykładem jest sytuacja, gdy publicznie dostępne dane – takie jak spis wyborców, książkę telefoniczną lub inne dane łatwo dostępne dzięki wyszukiwarce internetowej – można połączyć ze zanonimizowanymi (w niewłaściwy sposób) danymi, dzięki

²⁶ W celu uzyskania dalszych informacji na temat technik anonimizacji zob. zapowiadana opinia Grupy Roboczej Art. 29 poświęcona temu zagadnieniu.

czemu możliwa jest identyfikacja osoby fizycznej (np. na podstawie jej daty urodzenia i kodu pocztowego).

Ryzyko ponownej identyfikacji może wzrosnąć, gdy jedna osoba fizyczna lub grupa osób fizycznych posiada dużo informacji na temat innej osoby fizycznej, na przykład członka rodziny, kolegi z pracy, kontaktu na portalu społecznościowym, lekarza, nauczyciela, urzędnika organu ścigania lub przedstawiciela innej branży.

W takiej sytuacji liczy się jednak nie tylko to, czy osoba fizyczna dysponująca wstępną wiedzą może zidentyfikować daną osobę, której dane dotyczą, ale również to, czy dowie się ona czegoś nowego z informacji uzyskanych w ramach ponownej identyfikacji. Dwa poniższe przykłady ilustrują wagę tego rozróżnienia.

Przykłady pierwszy: statystyki dotyczące odrę. W jednym przypadku ze zanonimizowanych danych statystycznych można się dowiedzieć, że w mieście A w roku 2012 X mieszkańców chorowało na odrę. Nie jest dostępna żadna inna klasyfikacja czy informacja. Lekarz, który przyczynił się do powstania statystyk, dostarczając informacje na temat własnych pacjentów właściwym organom zdrowia, ma pełniejsze informacje na temat tych pacjentów w swoim gabinecie, objęte tajemnicą lekarską. Lekarz ten mógłby łatwo ponownie zidentyfikować kilku pacjentów ze statystycznego zestawu danych. Podobnie matka, która wiedziała, że jej dziecko chorowało na odrę w tym roku, mogłaby łatwo ponownie je zidentyfikować w przedmiotowym zestawie danych. Ani matka, ani lekarz nie uzyskaliby jednak ze zanonimizowanego, udostępnionego publicznie zestawu danych żadnych informacji, których by nie znali wcześniej.

Przykład drugi: nadużywanie środków odurzających i alkoholu, niegodziwe traktowanie w celach seksualnych a wyniki w nauce osiągnane w szkole. Przykład ten można skonstruować z sytuacją opisaną powyżej. Prowadzone są badania nad korelacjami między nadużywaniem środków odurzających i alkoholu przez rodziców, niegodziwym traktowaniem dzieci w celach seksualnych a ich wynikami w nauce. Dane uzyskane w trakcie badań, które rzekomo zostały „zanonimizowane”, są publikowane w dobrej intencji, jednak bez dokładnej oceny ryzyka ponownej identyfikacji.

Statystyki pokazują między innymi, że w Szkole A, do której chodzi ogółem 500 uczniów, w roku 2012 20 % uczniów (100 osób) żyło w gospodarstwie domowym, w którym co najmniej jeden rodzic jest alkoholikiem lub narkomanem. Z tego w 8 % przypadków (8 uczniów) dziecko było niegodziwie traktowane w celach seksualnych. W sprawozdaniu stwierdza się też, że żaden inny uczeń w Szkole A nie był niegodziwie traktowany w celach seksualnych.

Dane liczbowe pokazują również, że w 96 % przypadków (96 uczniów) dzieci, których rodzice byli alkoholikami lub narkomanami, miały znacznie gorsze wyniki w nauce („niskie wyniki” określone zgodnie z właściwymi normami kształcenia), jednak w tej konkretnej szkole tylko 50 % dzieci niegodziwie traktowanych w celach seksualnych (4 uczniów) miało znaczne trudności w nauce.

W szkole tej wszyscy wiedzą, że AA, bystry i pracowity chłopiec, ma trudne środowisko rodzinne, a jego matka jest alkoholiką. Często jest szykanowany przez niektórych kolegów z klasy. Ci sami koledzy odkrywają ze statystyk przedrukowanych w gazetce szkolnej, że AA musi mieścić się wśród 50 % dzieci traktowanych niegodziwie w celach seksualnych, które nie mają problemów z nauką („dobre wyniki”). Uzyskali więc oni nowe informacje (w tym przypadku szczególnie chronione) z zestawu danych, które nie zostały skutecznie zanonimizowane.

Ryzyko połączenia informacji w celu uzyskania danych osobowych wzrasta wraz z rozwojem technik łączenia danych i zwiększaniem mocy obliczeniowych komputerów, a także wraz z

publicznym udostępnianiem nowych potencjalnie „dopasowywalnych” informacji. Co roku moce obliczeniowe wzrastają dwukrotnie, a przechowywanie danych, dzięki dostępności usług przechowywania w chmurze, prawdopodobnie stanie się dobrem powszechnie dostępnym. Tym samym ryzyko ponownej identyfikacji poprzez połączenie danych jest nieprzewidywalne, ponieważ nie sposób nigdy określić z całkowitą pewnością, które dane są już dostępne lub mogą zostać udostępnione w przyszłości.

Mimo tej niepewności ryzyko ponownej identyfikacji można zazwyczaj przynajmniej do pewnego stopnia ograniczyć poprzez stosowanie zasady minimalizacji danych, tj. poprzez dopilnowanie, by ujawniane były tylko dane niezbędne do danego celu.

Prawdopodobieństwo udanej ponownej identyfikacji: test „zdeteminowanego intruza”

Test „zdeteminowanego intruza” jest nowym pojęciem, które należy jeszcze dokładnie przeanalizować. Pomocne może być ustalenie, czy:

- ktokolwiek miałby powód, by przeprowadzić ponowną identyfikację oraz
- czy ponowna identyfikacja może/ma szansę się powieść.

Test zdeteminowanego intruza wiąże się głównie ze sprawdzeniem, czy „intruz” byłby w stanie dokonać ponownej identyfikacji, gdyby miał motywację, by to uczynić. „Zdeteminowany intruz” to osoba (osoba fizyczna lub organizacja), która pragnie zidentyfikować osobę fizyczną, od której pochodzą zanonimizowane dane osobowe. Test ten ma na celu dokonanie oceny, czy zdeteminowany intruz osiągnie swój cel. Podejście to zakłada, że „zdeteminowany intruz” jest kompetentny i ma dostęp do zasobów proporcjonalnych do jego motywacji do dokonania ponownej identyfikacji.

Niektóre rodzaje danych będą atrakcyjniejsze dla „zdeteminowanego intruza” niż inne. Przykładowo intruz może być zasadniczo bardziej zmotywowany do ponownego zidentyfikowania danych osobowych, jeżeli dane te:

- mają znaczną wartość handlową (w tym na czarnym rynku lub poza Unią Europejską), a zatem mogą zostać kupione i sprzedane w celu odniesienia korzyści finansowej²⁷;
- mogą zostać użyte w celu egzekwowania prawa lub czynności wywiadu;
- zawierają warte publikacji informacje na temat osób publicznych;
- mogą zostać użyte do celów politycznych lub społecznych (np. w ramach kampanii przeciwko konkretnej organizacji lub osobie);
- mogłyby zostać użyte do celów osobistych w złej intencji (np. stalking, mobbing, zastraszanie lub po prostu skompromitowanie innych).
- mogą budzić ciekawość (np. chęć odkrycia przez osobę z sąsiedztwa, kto brał udział w wydarzeniu umieszczonym na mapie przestępstw).

Choć rozumowanie przez pryzmat możliwych motywów potencjalnych intruzów jest pomocne, Grupa Robocza Art. 29 podkreśla, że podejście to ma poważne ograniczenia:

²⁷ Dane takie mogą obejmować na przykład dane transakcyjne lub inne dane behawioralne, z których można wywnioskować indywidualne profile konsumentów, które z kolei można następnie wykorzystać do celów reklamowych lub dyskryminacji cenowej; informacje finansowe lub inne umożliwiające kradzież tożsamości; dane szczególnie chronione, które mogą zostać wykorzystane do szantażowania lub dyskryminacji osób fizycznych; informacje medyczne, które mogłyby być wykorzystane przez firmy ubezpieczeniowe, na przykład w celu odmowy ubezpieczenia w oparciu o wcześniej istniejące schorzenie; informacje umożliwiające wnioskowanie o zdolności kredytowej, które mogą zostać wykorzystane do oceny ryzyka kredytowego itp.

- metoda ta może mieć w pewnym stopniu charakter spekulacyjny;
- w przypadku braku oczywistych „czynników motywacyjnych”, takich jak opisane powyżej, metoda ta może prowadzić do fałszywych zapewnień i może sugerować, że dane osobowe, które są stosunkowo nieszkodliwe, mogą zostać udostępnione do ponownego użycia bez poddania ich skutecznej anonimizacji;
- intruzy mogą posiadać zaawansowane umiejętności, mogą być innowacyjni i „na prowadzeniu”, znajdując takie sposoby użycia ponownie zidentyfikowanych danych, które nie są oczywiste dla innych;
- wraz z rosnącą tendencją do przeprowadzania analizy „dużych danych” istnieje zwiększone ryzyko, że po ponownym zidentyfikowaniu pozornie nieszkodliwe dane mogą w połączeniu z innymi informacjami w ostateczności stanowić poważniejsze ryzyko.

6.5. Test na ponowną identyfikację

W niektórych okolicznościach może być trudno ustalić ryzyko ponownej identyfikacji, szczególnie gdy osoba trzecia może wykorzystać skomplikowane metody statystyczne w celu połączenia różnych zanonimizowanych danych. Stąd dobrą praktyką jest zastosowanie testu na ponowną identyfikację w ramach całościowej oceny przeprowadzanej, aby określić ryzyko ponownej identyfikacji, – jest to rodzaj testu „penetracyjnego” lub „pentestu”, mający na celu wykrycie i wyeliminowanie potencjalnych zagrożeń ponownej identyfikacji. Polega on na próbie ponownego zidentyfikowania osób fizycznych na podstawie zestawów danych, których upublicznienie jest planowane.

Pierwszy etap testu na ponowną identyfikację powinien polegać na przeanalizowaniu zestawów danych, które sektor publiczny opublikował lub zamierza opublikować. Kolejny etap powinien polegać na próbie określenia, jakie inne dane – dane osobowe lub dane o innym charakterze – są dostępne, a które można powiązać z powyższymi danymi, aby doszło do ponownej identyfikacji. W szczególności ukierunkowane „testy penetracyjne” powinny pomóc ocenić, jaki jest poziom zagrożenia w kontekście identyfikacji puzzlowej, tj. złożenia różnych informacji w jeden obraz w celu stworzenia pełniejszego profilu jakiejś osoby.

Oczywiście testu na ponowną identyfikację nie powinno się postrzegać jako panaceum, a jego stosowanie nie powinno prowadzić do fałszywego poczucia bezpieczeństwa. Po pierwsze, testowanie może być trudne do przeprowadzenia, ponieważ często wymaga znacznej specjalistycznej wiedzy technicznej i odpowiednich narzędzi oraz świadomości, jakie inne dane mogą być dostępne. Po drugie, administratorzy muszą być również świadomi, że ryzyko ponownej identyfikacji może z czasem ulec zmianie. Przykładowo obecnie dostępne są coraz bardziej skuteczne i przystępne cenowo techniki i narzędzia analizy danych, w związku z czym korelacja z innymi zestawami danych staje się coraz łatwiejsza w miarę generowania coraz większej ilości danych. Stąd organizacje powinny przeprowadzać okresowy przegląd swojej polityki w zakresie udostępniania danych oraz technik wykorzystywanych do anonimizacji danych. Ponadto decyzji nie powinno się opierać wyłącznie na bieżących zagrożeniach – trzeba brać pod uwagę także przewidywane przyszłe zagrożenia.

Po dokonaniu oceny zgodnie z pkt 6.4 w zakresie ryzyka ponownej identyfikacji oraz w stosownych przypadkach po przeprowadzeniu testu na ponowną identyfikację organ sektora publicznego może ustalić, czy zestaw danych można uznać za zestaw zanonimizowany, innymi słowy, czy nie zawiera on już danych osobowych w rozumieniu art. 2 lit. a) i motywu 26 dyrektywy 95/46/WE. Jeżeli tak,

zestaw danych można udostępnić bez ograniczeń w zakresie ochrony danych²⁸. Z drugiej strony, jeżeli test się powiódł, danych tych nie można udostępniać jako danych zanonimizowanych (lub należy zaprzestać ich udostępniania) i należy je uznać za dane osobowe (a tym samym ich udostępnienie nie może być dopuszczalne lub może być dopuszczalne tylko pod warunkiem spełnienia wymogów omówionych w sekcji VII).

6.6 Wycofanie zestawów danych, których bezpieczeństwo zostało naruszone

W przypadku udowodnienia ponownej identyfikacji danych z otwartego zestawu danych organ sektora publicznego dostarczający ten zestaw danych musi być w stanie zaprzestać ich dostarczania lub usunąć zestaw danych z otwartej strony internetowej. W przypadku usunięcia zestawu danych ze strony internetowej, organ sektora publicznego musi również poinformować o tym ponownych użytkowników i wezwać ich do zaprzestania przetwarzania oraz do usunięcia wszystkich danych pochodzących z zestawu danych, którego bezpieczeństwo zostało naruszone. Ponieważ poinformowanie wszystkich użytkowników będzie trudne w ramach wymaganego przez dyrektywę ISP systemu otwartych licencji, organy publiczne muszą wdrożyć dostatecznie skuteczne środki, by rozwiązać ten problem. Choć wycofanie może być często zbyt późne, by zapobiec szkodom, jest ono niezbędnym krokiem, by pomóc złagodzić niekorzystne skutki dla osób, których dane dotyczą.

VII. Otwieranie danych osobowych do ponownego wykorzystania

7.1 Przykłady publicznie dostępnych danych osobowych ujawnianych przez organy sektora publicznego

Chociaż udostępnianie zestawu danych zanonimizowanych jest typowym scenariuszem dla inicjatyw ponownego wykorzystania ISP, w niektórych przypadkach organy sektora publicznego mogą również udostępniać dane osobowe do ponownego wykorzystania.

Wiele publicznie dostępnych rejestrów, takich jak rejestry gruntów lub rejestry przedsiębiorców, zawiera znaczną ilość danych osobowych, a także jest coraz częściej udostępnianych online w ramach inicjatyw administracji elektronicznej. Istnieje również wiele innych przykładów, w których prawodawcy w poszczególnych państwach członkowskich ustanowili podstawę prawną udostępniania danych osobowych osób fizycznych w internecie lub na wniosek o dostęp do dokumentów. Obejmują one na przykład²⁹:

- koszty, wynagrodzenia lub oświadczenia o braku konfliktu interesów niektórych urzędników publicznych lub beneficjentów pomocy państwa (na przykład dotacji rolniczych),
- nazwy organizacji lub nazwiska osób fizycznych wspierających finansowo partie polityczne,
- deklaracje podatkowe osób fizycznych³⁰,
- wyroki sądowe (z nazwiskami stron lub innych osób fizycznych, niekiedy usuwanymi lub zastępowanymi inicjałami w celu ograniczenia ryzyka ponownej identyfikacji),
- listy wyborców,
- wykazy sądowe (np. wokanda, zawierająca spis spraw, które będą toczyć się przed sądem w konkretnych dniach).

²⁸ Por. jednak pkt 10.3 „Warunki licencji dla zanonimizowanych zestawów danych”, a w szczególności konieczność wprowadzenia zabezpieczeń w celu dalszego zapewniania, by osoby fizyczne nie zostały ponownie zidentyfikowane.

²⁹ Zob. również przykłady przedstawione w pkt V przy omawianiu zakresu dyrektywy ISP.

³⁰ Zob. np. wyrok Trybunału Sprawiedliwości z dnia 16 grudnia 2008 r. w sprawie C-73/07 *Tietosuoja ja valtuutettu* przeciwko *Satakunnan Markkinapörssi Oy* i *Satamedia Oy*.

W każdym z tych przypadków organy sektora publicznego lub prawodawcy mogą samodzielnie rozważyć, czy chcą udostępnić przedmiotowe dane do ponownego użycia (na przykład w celu poprawy jakości usług publicznych, takich jak zapewnienie dostępu do rejestru przedsiębiorców lub gruntów). Potencjalni ponowni użytkownicy mogą również skontaktować się z organami sektora publicznego z wnioskiem o ponowne wykorzystanie danych. W niektórych innych przypadkach możliwa jest również sytuacja, w której ponowni użytkownicy zwyczajnie pobiorą dane osobowe, które są już dostępne online i wykorzystają je bez obowiązkowego kontaktu z organem sektora publicznego, który udostępnił informacje. We wszystkich tych trzech przypadkach ponowni użytkownicy będą oczywiście musieli dostosować się do przepisów o ochronie danych, ponieważ mają oni do czynienia z danymi osobowymi.

7.2 Różnice w krajowych systemach dostępu

Wymogi prawne w zakresie publicznego udostępniania niektórych danych osobowych różnią się znacznie w poszczególnych państwach członkowskich ze względu na różne tradycje kulturowe i prawne. W niektórych państwach członkowskich istnieje podstawa prawna do udostępniania określonych danych osobowych, podczas gdy inne państwa członkowskie zakazują udostępniania tych samych danych osobowych w takiej samej sytuacji. W dyrektywie ISP uznaje się i wyraźnie stwierdza, że jest ona oparta na istniejących systemach dostępu obowiązujących w państwach członkowskich i nie zmienia przepisów krajowych dotyczących dostępu do dokumentów³¹.

7.3 Konieczność oceny skutków w zakresie ochrony danych i odpowiednie zabezpieczenia

Co do zasady, jeżeli planowane jest udostępnienie danych osobowych do ponownego wykorzystania, absolutnie niezbędne jest zastosowanie ostrożnego podejścia. Grupa Robocza Art. 29 w szczególności zaleca, by przeprowadzić dogłębną ocenę skutków w zakresie ochrony danych osobowych przed publikacją zestawu danych (lub przed przyjęciem przepisów przewidujących wymóg upublicznienia), w której przeanalizuje się również możliwości i potencjalne skutki ponownego wykorzystania danych. Ogólnie należy unikać otwierania danych osobowych do ponownego wykorzystania w ramach otwartej licencji bez zastosowania jakichkolwiek ograniczeń technicznych i prawnych w zakresie ponownego wykorzystania.

7.4 Znaczenie systemu licencjonowania

Ponadto Grupa Robocza Art. 29 zaleca wprowadzenie ścisłego systemu licencjonowania, który należy odpowiednio egzekwować, aby mieć pewność, że dane osobowe nie będą wykorzystywane do niezamierzonych celów – na przykład do niezamówionych informacji handlowych lub w inny sposób, który osoby, których dane dotyczą, uznałyby za nieoczekiwany, niewłaściwy lub z innych przyczyn budzący sprzeciw.

7.5 Znaczenie solidnej podstawy prawnej zarówno dla upubliczniania, jak i ponownego użycia danych

Grupa Robocza Art. 29 ponownie podkreśla znaczenie ustanowienia solidnej podstawy prawnej dla publicznego udostępniania danych osobowych, z uwzględnieniem właściwych zasad ochrony danych, w tym zasady proporcjonalności, minimalizacji danych oraz zasady celowości.

Grupa Robocza Art. 29 zaleca, by przepisy przewidujące publiczny dostęp do danych, jasno określały cele udostępniania danych osobowych. Jeżeli tak się nie stanie albo cele te będą określone

³¹ Przy czym, jak wyjaśniono w pkt 5.4, przepisy krajowe muszą nadal być zgodne z art. 8 EKPC oraz art. 7 i 8 Karty praw podstawowych UE, zgodnie z właściwym orzecznictwem.

w sposób nieprecyzyjny i szeroki, ucierpi na tym pewność i przewidywalność prawa. W szczególności w odniesieniu do każdego wniosku o ponowne wykorzystanie danych organowi sektora publicznego i odnośnym, potencjalnym ponownym użytkownikom będzie bardzo trudno ustalić, jakie były zamierzone pierwotne cele upublicznienia, a tym samym, jakie dalsze cele byłyby zgodne z tymi pierwotnymi celami. Jak już wspomniano, nawet jeżeli dane osobowe zostały opublikowane w internecie, nie należy zakładać, że można je dalej przetwarzać w jakichkolwiek możliwych celach.

Każde dalsze ponowne wykorzystanie danych musi w takich przypadkach mieć odpowiednią podstawę prawną (np. wyrażenie zgody lub wymóg prawny) zgodnie z art. 7 lit. a) – f) dyrektywy 95/46/WE i być zgodne z wszystkimi pozostałymi przepisami o ochronie danych.

7.6 Zasada celowości

Skuteczne zastosowanie zasady celowości w przypadku ponownego użycia ISP stanowi znaczne wyzwanie. Z jednej strony sama koncepcja i siła napędowa innowacyjności stojąca za pojęciem „otwartych danych” i ponownym użyciem ISP sprowadza się do tego, że informacje powinny być dostępne do ponownego wykorzystania w nowych, innowacyjnych produktach i usługach, a tym samym w celach, które nie zostały wcześniej określone i nie sposób ich wyraźnie przewidzieć. Dyrektywa ISP również wymaga metod licencjonowania, które niepotrzebnie nie ograniczają ponownego wykorzystania danych.

Z drugiej strony zasada celowości jest kluczową zasadą ochrony danych, wymagającą, by dane osobowe, które zgromadzono do konkretnego celu, nie zostały następnie wykorzystane do innego celu niezgodnego z celem pierwotnym³². Zasada ta ma również zastosowanie do danych osobowych, które są publicznie dostępne. Sam fakt, że dane osobowe są publicznie dostępne w konkretnym celu, nie oznacza, iż takie dane osobowe są otwarte do ponownego wykorzystania w jakimkolwiek innym celu.

Przykładowo wydatki urzędników państwowych wyższego szczebla są udostępniane w internecie w celu zapewnienia przejrzystości, ale umożliwienie ponownego wykorzystania takich danych przez któregośkolwiek członka społeczeństwa do innych celów może być niezgodne z zasadą celowości.

Jak zostało to bardziej szczegółowo omówione w opinii 3/2013 Grupy Roboczej Art. 29 w sprawie zasady celowości (zobacz sekcja III.2.2 oraz załącznik 1), przeanalizowanie, czy dalsze przetwarzanie danych osobowych jest niezgodne z celami, dla których dane te zostały zgromadzone, wymaga wieloczynnikowej oceny. Należy w szczególności uwzględnić:

- a) związek między celami, dla których zgromadzono dane osobowe, a celami dalszego przetwarzania;
- b) kontekst, w jakim gromadzono określone dane osobowe, oraz uzasadnione oczekiwania osób, których dane dotyczą, co do ich dalszego wykorzystania;
- c) charakter danych osobowych oraz wpływ dalszego przetwarzania tych danych na osoby, których dane dotyczą;
- d) zabezpieczenia wprowadzone przez administratora w celu zapewnienia uczciwego przetwarzania danych oraz zapobieżenia niepożądanym skutkom dla osób, których dane dotyczą.

³² Dane można wykorzystać w sposób niezgodny z celami określonymi w chwili ich gromadzenia tylko w drodze wyjątku – z zastrzeżeniem ścisłych zabezpieczeń na mocy art. 13 dyrektywy 95/46/WE. Zob. pkt III.3 opinii 3/2013 Grupy Roboczej Art. 29 w sprawie zasady celowości.

Przedmiotowe kluczowe czynniki należy uwzględnić przy podejmowaniu decyzji o publicznym udostępnieniu danych osobowych oraz w każdym przypadku, gdy dane osobowe będą ponownie wykorzystywane. Poniżej przedstawiono kilka przykładów.

- Organ sektora publicznego publikuje w formie katalogu dane kontaktowe jego urzędników, w tym nazwisko, stanowisko oraz adres i telefon służbowy. Oczywiście, choć niewyrażonym wprost celem tego katalogu jest ułatwienie społeczeństwu ustalenia, do kogo należy się zwracać z urzędowymi zapytaniami i innymi urzędowymi sprawami. Ponowny użytkownik chce zebrać zawartość tego katalogu, połączyć go z adresami i telefonami domowymi pracowników (jeżeli są one publicznie dostępne na przykład w książce telefonicznej) i udostępnić zarówno domowe, jak i służbowe adresy i numery telefonów na interaktywnej mapie, aby pokazać, gdzie poszczególni urzędnicy żyją i pracują. To połączenie danych i ich ponowne użycie należy uznać za niezgodne z pierwotnym celem. Urzędnik, którego służbowe dane kontaktowe są udostępniane, aby umożliwić społeczeństwu skontaktowanie się z nim, nie może w uzasadniony sposób oczekiwać, że te informacje zostaną powiązane z innymi danymi, które wspomniany urzędnik udostępnił publicznie w innym celu, niezwiązanym z pracą.
- W niektórych państwach członkowskich zgodnie z przepisami krajowymi ogłoszenia o planowanym zawarciu związku małżeńskiego są publiczne i dostępne dla wszystkich. Takie zapowiedzi mają na celu poinformowanie o zamiarze zawarcia związku małżeńskiego przez zaręczoną parę i umożliwienie zainteresowanym osobom wyrażenia sprzeciwu. Fakt, że dane osobowe zawarte w publicznych ogłoszeniach o planowanym zawarciu związku małżeńskiego są dostępne dla wszystkich, nie zezwala jednak osobom trzecim na wykorzystanie takich informacji do przesyłania parze informacji handlowych. To dodatkowe wykorzystanie danych byłoby niezgodne z pierwotnym celem, biorąc pod uwagę cel publicznego udostępnienia zapowiedzi ślubnych, którym jest umożliwienie wyrażenia sprzeciwu wobec związku małżeńskiego, jak stanowią przepisy.

7.7. Cele komercyjne a cele niekomercyjne

W opinii 7/2003 podkreślono działalność handlową jako główny bodziec do ponownego wykorzystania ISP zestawiając ją z sytuacją udzielania dostępu do informacji, w którym to przypadku celem przepisów dotyczących dostępu do informacji publicznej jest zapewnienie obywatelom przejrzystości, otwartości i odpowiedzialności względem nich.

Ponadto w opinii 7/2003 podkreślono, że „w normalnej sytuacji [obywatele] wykorzystują informacje do własnych celów o charakterze niekomercyjnym”. Stwierdzenie to należy uaktualnić w świetle doświadczeń nabytych w związku z ponownym wykorzystaniem ISP. Doświadczenia z inicjatywami w zakresie otwartych danych pokazały, że ponowne wykorzystanie ISP może również znacznie przyczynić się do zwiększenia przejrzystości i odpowiedzialności oraz może prowadzić do lepszego wykorzystania usług publicznych. Rozróżnienie na ponowne wykorzystanie do celów komercyjnych i niekomercyjnych nie powinno być decydujące przy ocenie zgodności dalszego wykorzystania danych osobowych z pierwotnym celem. Ocena zgodności z pierwotnym celem nie powinna się opierać głównie na tym, czy model ekonomiczny potencjalnego ponownego użytkownika jest oparty na zysku.

Należy natomiast dokładnie ocenić, czy cele i sposób dalszego przetwarzania danych są zgodne z pierwotnymi celami zgodnie z kryteriami, o których mowa w pkt 7.6. W przypadku ponownego użycia ISP doprowadzi to w sposób nieunikniony do rozważenia szeregu scenariuszy przetwarzania danych, a nie tylko jednego.

7.8 Proporcjonalność i inne zagadnienia

Inną kluczową zasadą, o której mowa w dyrektywie 95/46/WE, jest zasada proporcjonalności³³. Istnieje wiele różnych metod i warunków publicznego udostępniania danych osobowych. Niektóre z nich mogą być bardziej niepożądane niż inne i mogą stanowić większe ryzyko. W konsekwencji niektóre z nich można uznać za proporcjonalne, podczas gdy innych za takie uznać nie można.

W odniesieniu do celu pojawia się problem, jak kontrolować dalsze przetwarzanie danych i zapewnić zgodność z innymi zasadami przepisów o ochronie danych, w tym między innymi z zasadą proporcjonalności. Gdy dane zostaną publicznie udostępnione, zwłaszcza w internecie, bardzo trudno skutecznie ograniczyć ich wykorzystanie i zapewnić zgodność z przepisami o ochronie danych.

Niektóre z wyzwań zapewnienia zgodności z przepisami o ochronie danych obejmują następujące kwestie:

- jak zapewnić aktualizowanie i dokładność danych, które są odłączone od pierwotnego źródła;
- jak dopilnować, by wykorzystanie danych osobowych pozostało ograniczone do funkcji przewidzianych w pierwotnym celu upublicznienia;
- jak zapewnić terminowe usunięcie danych, jeżeli upublicznienie danych osobowych było przewidziane tylko na pewien czas³⁴;
- jak wykonywać prawa osób fizycznych w przypadku danych osobowych udostępnionych do ponownego wykorzystania (w tym prawo do żądania poprawy, uaktualnienia lub usunięcia danych).

7.9 Prawne lub techniczne ograniczenia ponownego wykorzystania danych

Niekiedy przepisy lub projekt techniczny systemów ograniczają wykonywanie konkretnych operacji przetwarzania danych lub ustanawiają inne zabezpieczenia ograniczające wykorzystanie publicznych rejestrów (np. ograniczenie możliwości pobierania całej treści rejestru lub ograniczenie wyszukiwań przykładowo w oparciu o imię i nazwisko osoby fizycznej). W takim przypadku ponowne wykorzystanie danych powinno być zasadniczo możliwe tylko zgodnie z tymi konkretnymi ograniczeniami i warunkami.

W tym kontekście ważne jest dokładne sprawdzenie, jakie środki – zarówno prawne, jak i techniczne – można wdrożyć, by ułatwić rozwiązanie kwestii ochrony danych, w tym zagadnień, o których mowa w sekcji 7.8. Należy w szczególności sprawdzić, w jaki sposób ponowni użytkownicy będą mieli dostęp do danych – na przykład poprzez funkcję masowego pobierania danych albo ukierunkowany interfejs zawierający ograniczone funkcje dostępu podlegające określonym warunkom. W tym względzie zasadnicze znaczenie ma fakt, jakie dodatkowe środki bezpieczeństwa zostaną wdrożone, takie jak na przykład system weryfikacji „captcha”³⁵ zapobiegający

³³ Zob. art. 6 ust. 1 lit. c) dyrektywy 95/46/WE.

³⁴ Zob. na przykład wyrok Europejskiego Trybunału Sprawiedliwości w sprawach połączonych C 92/09 i C 93/09 *Volker und Markus Schecke GbR przeciwko Land Hessen*, pkt 31: „[n]ie ma możliwości wycofania danych z Internetu po upływie okresu dwóch lat przewidzianego w art. 3 ust. 3 rozporządzenia nr 259/2008”.

³⁵ CAPTCHA (ang. *Completely Automated Public Turing test to tell Computers and Humans Apart*) jest testem opartym na systemie zadania-odpowiedzi w celu rozróżnienia ludzi od automatycznych programów. CAPTCHA odróżnia człowieka od komputera poprzez wyznaczenie określonego zadania, które jest łatwe dla większości ludzi, ale jest trudniejsze dla obecnych programów komputerowych.

automatycznemu dostępowi do danych i minimalizujący ryzyko pobrania całej bazy danych. Wykorzystanie specjalnych środków technicznych mogłoby pomóc w ograniczeniu niewłaściwego wykorzystywania danych osobowych i negatywnych skutków dla osób, których dane dotyczą; skutki te w przeciwnym wypadku mogłyby stać się realne w wyniku nieograniczonego i bezwarunkowego dostępu ponownych użytkowników do całych zestawów danych.

Co ważniejsze, w wielu przypadkach może zachodzić konieczność zapewnienia, by ponowni użytkownicy byli w stanie dokonywać jedynie ukierunkowanych zapytań za pośrednictwem technologii mających na celu zapobieganie masowym pobraniom rejestrów danych, takich jak zaprojektowane na zamówienie interfejsy programowania aplikacji („API”). Może to przyczynić się do zapewnienia proporcjonalności wykorzystania danych i zmniejszenia ryzyka niewłaściwego wykorzystania całych baz danych. Ponadto takie indywidualne interfejsy mogą również ułatwić dopilnowanie, by dane były zawsze aktualizowane, a także by dane nie były już dostępne poprzez API po podjęciu takiej decyzji przez zainteresowany organ sektora publicznego. Z drugiej strony możliwe jest ograniczenie sposobów, w jakie ponowni użytkownicy mogą ponownie wykorzystać dane.

7.10. Dokładność, uaktualnianie i usuwanie

Innym szczególnym problemem jest to, co dzieje się, gdy dane osobowe są publikowane lub w inny sposób publicznie udostępniane tylko na określony czas. Artykuł 6 ust. 1 lit. e) dyrektywy 95/46/WE stanowi, że dane osobowe należy przechowywać w formie umożliwiającej identyfikację osób, których dane dotyczą, przez czas nie dłuższy niż jest to konieczne do celów, dla których dane zostały zgromadzone lub dla których są dalej przetwarzane. Motyw 18 dyrektywy ISP również stanowi, że jeżeli właściwe władze zdecydują o zaprzestaniu udostępniania niektórych dokumentów do ponownego wykorzystywania lub przestaną te dokumenty aktualizować, decyzje te powinny być znane publicznie tak szybko, jak to możliwe, w miarę możliwości przy wykorzystaniu środków elektronicznych.

Dopilnowanie, by dane zostały usunięte czy skasowane po ich upublicznieniu i udostępnieniu do ponownego wykorzystania jest jednak trudne, a czasami nawet niemożliwe.

W tym względzie pewnym rozwiązaniem – choć w żadnym razie nie idealnym – byłoby udostępnianie danych nie w formie umożliwiającej ich pobranie, a w formie indywidualnego interfejsu API oraz z zastrzeżeniem określonych ograniczeń i środków bezpieczeństwa, jak opisano powyżej.

VIII. Wyniki badań naukowych

Należy tu poczynić rozróżnienie z jednej strony na upublicznienie zanonimizowanych danych (zobacz sekcja VI), a z drugiej na ograniczony dostęp. Jest rzeczą oczywistą, że programy otwartych danych są oparte na publicznej dostępności danych. Wiele badań (zwłaszcza badań naukowych w celach komercyjnych lub niekomercyjnych, ale również badań o innym charakterze) odbywa się jednak poprzez ujawnienie danych w obrębie zamkniętej społeczności, tj. takiej, w której skończona liczba badaczy lub instytucji ma dostęp do danych oraz w której możliwe jest ograniczenie dalszego udostępniania lub wykorzystania danych i w której można zapewnić ich bezpieczeństwo.

Ograniczony dostęp jest szczególnie istotny przy przetwarzaniu danych osobowych (często w formie opatrzenia pseudonimem³⁶) pochodzących ze szczególnie chronionego materiału źródłowego lub gdy istnieje znaczne ryzyko ponownej identyfikacji. Udostępnianie danych z ograniczonym dostępem może nieść ze sobą pewne ryzyko, ale ryzyko to jest niższe i można je skuteczniej ograniczać, gdy dane są udostępniane w obrębie zamkniętej społeczności funkcjonującej według ustalonych reguł.

Problemem, przed jakim stają często osoby korzystające z danych do celów naukowych, jest fakt, że z jednej strony chcą one bogatych, szczegółowych danych, które będą mogły wykorzystać do swoich celów; a z drugiej strony chcą zagwarantować, by nie doszło do ponownej identyfikacji osób fizycznych. Z pewnej perspektywy jednostkowe dane opatrzone pseudonimami (na przykład poprzez zwykłe zakodowanie za pomocą klucza) mogą być bardzo cenne dla badaczy naukowych ze względu na szczegółowość na poziomie jednostkowym, a także dlatego, że opatrzone pseudonimami rejestry z różnych źródeł można stosunkowo łatwo do siebie dopasować. Oznacza to jednak również, że istnieje wysokie ryzyko ponownej identyfikacji: możliwość powiązania kilku zestawów danych (którym nadano pseudonimy lub ich nie nadano) z tą samą osobą fizyczną może być pierwszym krokiem do identyfikacji lub może umożliwić bezpośrednią identyfikację.

Dlatego przed jakimkolwiek upublicznieniem lub udostępnieniem do ponownego wykorzystania zestawów danych, które opatrzone pseudonimami, konieczny jest wyższy poziom kontroli i dodatkowe środki ostrożności. Ogólnie rzecz biorąc, im dane są bardziej szczegółowe, możliwe do połączenia i im bardziej jednostkowy mają charakter, tym bardziej ograniczony i kontrolowany powinien być dostęp do takich danych. Im dane są bardziej zagregowane i im trudniej jest je połączyć, tym większe jest prawdopodobieństwo, że mogą one zostać upublicznione i udostępnione do ponownego wykorzystania bez znacznego ryzyka.

Jest to złożony obszar podlegający ciągłym zmianom i niewłaściwe byłoby kategoryczne wykluczenie upublicznienia i ponownego wykorzystania wszystkich zestawów danych, które nie osiągają wysokiego progu anonimizacji opisanego w pkt VI. Grupa Robocza Art. 29 uznaje jednak jako praktyczną zasadę, że w ujęciu ogólnym ujawnianie na warunkach dyrektywy ISP zestawów danych o charakterze jednostkowym lub innych zestawów danych pociągających za sobą znaczne ryzyko ponownej identyfikacji często będzie niewłaściwe – przy czym zawsze pożądane jest przeprowadzenie zindywidualizowanej analizy i ostrożnej oceny.

Ponadto należy podkreślić, że jeżeli niektóre takie zestawy danych zostaną jednak upublicznione i udostępnione po dokładnej ocenie ryzyka i korzyści, udostępnienie i dalsze ponowne wykorzystanie danych musi nastąpić z zachowaniem pełnej zgodności z przepisami o ochronie danych (zobacz sekcja VII). Wynika to z faktu, że dane te mimo podjęcia pewnych (niekiedy bardzo znaczących) środków w celu ograniczenia ich ponownej identyfikacji nadal są uznawane za dane osobowe.

IX. Archiwa historyczne

Archiwa historyczne i muzea również posiadają określone cechy charakterystyczne wymagające specjalnych zabezpieczeń. W wielu przypadkach i w zależności od takich czynników, jak wiek i poziom szczególnej ochrony danych oraz kontekst ich gromadzenia, możliwe jest, iż bardziej właściwe niż cyfryzacja i udostępnianie danych do ponownego wykorzystania w internecie bez

³⁶ Zob. ponownie opinia 4/2007 w sprawie pojęcia danych osobowych, przyjęta w dniu 20.6.2007 r. (WP 136), w szczególności s. 12–21 (omówienie „danych opatrzonych pseudonimem”, „danych zakodowanych za pomocą klucza” i „danych anonimowych”, s. 18–21). Kwestię informacji „dotyczących” osoby fizycznej omówiono na s. 9–12. Jak odnotowano na s. 3, istotny jest również fakt, że Grupa Robocza Art. 29 zajmuje się obecnie opracowaniem kolejnych wytycznych w zakresie technik anonimizacji.

ograniczeń będą inne opcje – takie jak umożliwienie ograniczonego dostępu tylko pod warunkiem spełnienia wymogów poufności.

W odniesieniu do archiwów należy również podkreślić, że choć szczególna ochrona danych będzie co do zasady maleć z upływem czasu, niewłaściwe udostępnienie starych rejestrów zebranych na przestrzeni wielu dekad mogłoby nadal mieć bardzo szkodliwe skutki dla bezpośrednio zainteresowanej osoby fizycznej, ale również dla innych osób fizycznych, takich jak członkowie jej rodziny lub potomkowie. Jest tak zwłaszcza w przypadku danych szczególnie mocno chronionych. Dla przykładu ujawnione rejestry karne będą nadal wpływać negatywnie na określoną osobę fizyczną i utrudniać jej resocjalizację. Ponadto informacja, że zmarła osoba była agentem wywiadu lub współpracownikiem opresyjnego reżimu, pedofilem, przestępcą, cierpiała na stygmatyzującą chorobę psychiczną lub na chorobę dziedziczną, może mieć również negatywny wpływ na rodzinę (np. żyjącego małżonka, dzieci lub innych potomków) zmarłego. Próbkę DNA zmarłych osób, czasami przechowywane w archiwach szpitali publicznych, również mogłyby wymagać ochrony z tych samych względów. Stąd takie informacje, nawet jeżeli odnoszą się do zmarłych, mogą wymagać ochrony na mocy przepisów o ochronie danych lub ewentualnie innych przepisów chroniących prawa podstawowe.

Państwa członkowskie często posiadają specjalne przepisy regulujące dostęp do krajowych archiwów, archiwów okresów historii najnowszej cieszących się szczególnym zainteresowaniem (takich jak archiwa rejestrujące współpracę z opresyjnymi reżimami) i aktów wymiaru sprawiedliwości³⁷. W przepisach tych często wzywa się do wprowadzenia właściwych środków bezpieczeństwa i ograniczeń dostępu oraz innych zabezpieczeń służących zrównoważeniu wchodzących w grę interesów i zapewnieniu dostępności określonych danych osobowych do celów badań historycznych, przejrzystości i poszukiwań dziennikarskich, jednocześnie dopilnowując, by udostępnienia były w razie konieczności ograniczone w taki sposób, aby zapewnić poszanowanie życia prywatnego i rodzinnego oraz godności zainteresowanych stron.

W odniesieniu do „zasady celowości” należy zauważyć, że archiwa historyczne zazwyczaj przechowują informacje do celów badań historycznych. Cele te różnią się od pierwotnych celów, dla których gromadzono dane. Materiały, które ostatecznie trafiają do zbiorów archiwalnych, były początkowo tworzone w konkretnych celach administracyjnych przez różne organy sektora publicznego. Zazwyczaj po pewnym czasie, gdy dokumenty te nie są już niezbędne dla pierwotnych celów administracyjnych, przeprowadza się proces selekcji, a dokumenty uznane za dokumenty o „historycznej” wartości są przenoszone do archiwów historycznych. Pojawia się tu pytanie, jakie są cele, dla których dane osobowe przechowywane w archiwach powinny być dostępne do ponownego wykorzystania. W tym kontekście należy przeprowadzić dokładną ocenę – z uwzględnieniem potencjalnej wartości udostępnienia materiałów archiwalnych do ponownego wykorzystania, ale również jego potencjalnych skutków dla praw, wolności i godności zainteresowanych osób.

Zasadniczo można stwierdzić, że chociaż cyfryzacja niektórych rejestrów zawierających dane osobowe i udostępnienie ich do ponownego wykorzystania może być właściwe w pewnych sytuacjach, a niektóre dane można również ujawnić w formie zanonimizowanej, w innych przypadkach kluczowe znaczenie mają ograniczenia udostępniania i ponownego wykorzystania danych osobowych oraz odpowiednie środki bezpieczeństwa zapewniające ochronę takich danych.

³⁷ Inne przykłady mogłyby obejmować archiwa rejestrów stanu cywilnego, które w niektórych państwach członkowskich zawierają między innymi przyczynę śmierci, informację o zmianie płci, imię i nazwisko partnera (z którego można wywnioskować orientację seksualną osoby) lub fakt, że ktoś był adoptowany. Dostęp do tych archiwów również podlega szczególnym warunkom.

Dzięki całościowej ocenie skutków w zakresie ochrony danych powinno się zapewnić udostępnianie zbiorów archiwalnych do ponownego wykorzystania wyłącznie po wykluczeniu wszelkich potencjalnych negatywnych skutków dla osób fizycznych lub ograniczeniu takiego ryzyka do dopuszczalnego minimum. Sektor archiwów mógłby również rozważyć utworzenie kodeksów postępowaniu lub zmianę istniejących kodeksów w celu objaśnienia dobrej praktyki w tym zakresie.

X. Licencjonowanie danych osobowych do ponownego wykorzystania

10.1. Właściwe przepisy dyrektywy ISP

Motyw 15 dyrektywy ISP stanowi, że „warunkiem wstępnym rozwoju rynku informacyjnego w skali całej Wspólnoty jest zapewnienie, że warunki ponownego wykorzystywania dokumentów sektora publicznego są jasne i publicznie dostępne. Wszystkie stosowane warunki ponownego wykorzystywania dokumentów powinny być więc jasne dla potencjalnych użytkowników. Aby wspierać i ułatwiać wnioskowanie o ponowne wykorzystywanie, państwa członkowskie powinny zachęcać do tworzenia dostępnych w Internecie, tam gdzie jest to stosowne, wykazów dostępnych dokumentów”.

Motyw 26 dyrektywy zmieniającej dyrektywę ISP stanowi, że „w związku z wszelkimi przypadkami ponownego wykorzystywania dokumentu organy sektora publicznego mogą, w stosownych przypadkach, w drodze licencji, określać warunki [...]” oraz że „państwa członkowskie powinny w stosownych przypadkach zachęcać do korzystania z formatów otwartych przeznaczonych do odczytu komputerowego”.

Ponadto art. 8 ust. 1 stanowi, że „organ sektora publicznego może zezwolić na ponowne wykorzystywanie bez żadnych warunków lub może określić warunki, w uzasadnionych przypadkach w ramach licencji. Warunki te nie ograniczają niepotrzebnie możliwości ponownego wykorzystywania i nie są stosowane do ograniczania konkurencji”.

10.2. Licencjonowanie i ochrona danych

Licencje są centralną częścią systemu ISP. Mogą również wpływać na sposób przetwarzania danych osobowych i powinny znaleźć się wśród zabezpieczeń, które mają być stosowane przy udostępnianiu danych osobowych (lub zanonimizowanych danych, które pochodzą z danych osobowych) do ponownego wykorzystania. Licencje nie eliminują konieczności zachowania zgodności z przepisami o ochronie danych, jednak klauzula o ochronie danych przewidziana w warunkach licencji pomogłaby w zapewnieniu zgodności z przepisami o ochronie danych poprzez dodanie elementu „wykonalności”. Taka klauzula mogłaby również pomóc w podnoszeniu świadomości poprzez przypomnienie ponownym użytkownikom o ich obowiązkach jako administratorów.

W odniesieniu do treści licencji warto rozróżnić dwa odrębne warianty.

10.3. Warunki licencji dla zanonimizowanych zestawów danych

Po pierwsze, w odniesieniu do zanonimizowanych danych (tj. zestawów danych, które nie zawierają już danych osobowych), warunki licencji powinny

- przypominać, że zestawy danych zostały zanonimizowane;

- zakazywać posiadaczom licencji ponownej identyfikacji osób fizycznych³⁸;
- zakazywać posiadaczom licencji wykorzystywania danych do podejmowania środków lub decyzji wobec zainteresowanych osób fizycznych; oraz
- powinny również zawierać zobowiązanie posiadacza licencji do powiadomienia licencjodawcy w przypadku wykrycia, że osoby fizyczne mogą zostać lub zostały ponownie zidentyfikowane.

Jako alternatywę w stosunku do warunku licencji ponownym użytkownikom można przedstawić komunikat ostrzegawczy, wyświetlany w widocznym miejscu na portalu otwartych danych. Należy jednak promować przyjmowanie warunków licencji, ponieważ miałyby ono dodatkową zaletę wykonalności na podstawie umowy.

Wycofanie zestawów danych, których bezpieczeństwo zostało naruszone

Wszyscy pozostali użytkownicy internetu, w tym same osoby, których dane dotyczą, muszą mieć możliwość ostrzeżenia licencjodawcy o tym, że ponowna identyfikacja miała lub może mieć miejsce. Na wypadek wykrycia przez licencjodawcę zwiększonego ryzyka ponownej identyfikacji należy w licencji przewidzieć procedurę, w ramach której licencjodawca może „wycofać” zestaw danych, „których bezpieczeństwo zostało naruszone”. Innymi słowy, klauzula o ochronie danych powinna dawać licencjodawcy prawo do zawieszenia lub zakończenia dostępności danych (na przykład prawo do wyłączenia interfejsu API lub usunięcia pliku z platformy). Licencjodawca powinien dołożyć wszelkich zasadnych starań, aby wymóc na wszystkich ponownych użytkownikach usunięcie wszystkich zestawów danych bądź części tych zestawów danych, których bezpieczeństwo zostało naruszone (które mogą zostać ponownie zidentyfikowane). Starania te powinny obejmować umieszczenie widocznych powiadomień na stronach internetowych, takich jak portale otwartych danych lub fora/listy mailingowe/media społecznościowe odwiedzane przez grupy lub osoby fizyczne, które mogą ponownie wykorzystywać dane. Wymaganie rejestracji może być najbardziej skutecznym sposobem wycofania zestawów danych, ale nie należy zalecać rejestracji, jeżeli będzie ona wymagać gromadzenia nowych danych osobowych od ponownych użytkowników i jej ogólnym skutkiem byłoby zniechęcenie do korzystania ze stron internetowych ISP i innych usług.

10.4. Warunki licencji dla danych osobowych

W przypadku licencjonowania danych osobowych istnieje konieczność określenia granic wykorzystania takich danych. Głównym problemem jest tu dopilnowanie, by ponowne wykorzystanie danych było ograniczone do tego, co jest „zgodne z celami, dla których pierwotnie zgromadzono dane”³⁹. W tym celu warunki licencji muszą przynajmniej wyraźnie określać, dla jakich celów dane zostały pierwotnie opublikowane, i wskazywać, co byłoby uznane za wykorzystanie danych osobowych zgodne z pierwotnym celem, a co nie.

Należy jednak zauważyć, że nie powinno to „[ogranaczać] niepotrzebnie możliwości ponownego wykorzystywania” (art. 8 ust. 1 dyrektywy zmieniającej dyrektywę ISP). Często może to oznaczać, że ogólne warunki standardowych otwartych licencji nie są odpowiednie i że konieczne może być

³⁸ Ograniczone wyjątki mogą mieć zastosowanie np. w przypadkach testowania ponownej identyfikacji w dobrej wierze. Jednakże nawet w takich przypadkach wyniki testów powinny zostać przedstawione administratorowi danych i odnośnemu organowi sektora publicznego, a dane ponownie zidentyfikowane nie powinny być publikowane ani w inny sposób szerzej rozpowszechniane.

³⁹ Zob. ponownie opinia 3/2013 Grupy Roboczej Art. 29 w sprawie zasady celowości.

opracowanie specjalnych licencji dla niektórych danych osobowych, bądź też można użyć szablonów, które mogą zostać odpowiednio dostosowane.

Obecnie niektóre standardowe otwarte licencje (takie jak brytyjska otwarta licencja rządowa) wykluczają dane osobowe – nie są one w ogóle licencjonowane w ich warunkach.

10.5. W przypadku ponownej identyfikacji lub wykorzystania danych niezgodnie z pierwotnym celem powinno podejmować się zdecydowane środki w zakresie egzekwowania przepisów

Po opublikowaniu danych na podstawie licencji, takiej jak otwarta licencja rządowa, ich ochrona przed dalszym wykorzystaniem niezgodnym z celem pierwotnym, udostępnianiem lub zapewnienie ich bezpiecznego przechowania mogą być utrudnione. W tym kontekście niezwykle istotne jest monitorowanie ponownego wykorzystania danych i ściganie naruszeń – czy polegających na ponownej identyfikacji osób, których dane dotyczą, lub dalszym wykorzystaniu przez licencjodawcę danych w celu niezgodnym z celem pierwotnym.

Grupa Robocza Art. 29 przypomina znaczącą rolę organów sektora publicznego, podkreśla jednak również, że jeśli ponowny użytkownik gromadzi dane osobowe w procesie ponownej identyfikacji, taki ponowny użytkownik najprawdopodobniej zostanie uznany za przetwarzającego dane osobowe bezprawnie i może podlegać działaniom następczym ze strony organów ochrony danych. Zgodnie z proponowanym rozporządzeniem o ochronie danych działania takie obejmują wysokie grzywny.

XI. Wnioski

Podsumowując, Grupa Robocza Art. 29 po raz kolejny podkreśla, że ponowne wykorzystanie ISP może przynieść korzyści, prowadząc do większej przejrzystości i ponownego wykorzystania informacji sektora publicznego w sposób innowacyjny. Wynikająca z tego działania większa dostępność informacji wiąże się jednak z pewnym ryzykiem. W celu zapewnienia ochrony prywatności i danych osobowych osób fizycznych należy przyjąć zrównoważone podejście, a przepisy o ochronie danych muszą ułatwić kierowanie procesem wyboru w odniesieniu do kwestii, jakie dane osobowe mogą lub nie mogą być udostępniane do ponownego wykorzystania i jakie środki należy podjąć, aby zabezpieczyć te dane.

Niezależnie od „zasady ponownego wykorzystania danych” sformułowanej w dyrektywie zmieniającej dyrektywę ISP ponowne wykorzystanie danych do celów komercyjnych lub niekomercyjnych zgodnie z warunkami określonymi w dyrektywie ISP nie zawsze jest właściwe, jeżeli ISP, które mają być ponownie wykorzystane, zawierają dane osobowe. Często to dane statystyczne uzyskane na podstawie danych osobowych, a nie same dane osobowe, są i powinny być udostępniane do ponownego wykorzystania.

W pewnych sytuacjach może się jednak zdarzyć, że dane osobowe mogą zostać uznane za dostępne do ponownego wykorzystania na warunkach dyrektywy ISP, w stosownych przypadkach z zastrzeżeniem dodatkowych środków prawnych, technicznych lub organizacyjnych w celu ochrony zainteresowanych osób fizycznych. W odniesieniu do tych przypadków Grupa Robocza Art. 29 ponownie podkreśla znaczenie ustanowienia solidnej podstawy prawnej dla publicznego udostępniania danych osobowych, z uwzględnieniem właściwych zasad ochrony danych, w tym zasady proporcjonalności, minimalizacji danych oraz zasady celowości. W tym kontekście należy również ponownie podkreślić, że wszelkie informacje odnoszące się do zidentyfikowanej lub możliwej do zidentyfikowania osoby fizycznej, niezależnie od tego, czy są dostępne publicznie, stanowią dane osobowe. W związku z tym dostęp do danych osobowych, które zostały udostępnione

publicznie, oraz ponowne wykorzystywanie tych danych nadal podlega właściwym przepisom o ochronie danych.

W świetle tych rozważań Grupa Robocza Art. 29 zaleca, co następuje:

- zgodnie z zasadami „ochrony danych już w fazie projektowania oraz ochrony danych jako opcji domyślnej” przy rozważaniu publicznego udostępnienia ISP należy jak najwcześniej uwzględnić fakt, że niektóre ISP mogą zawierać dane osobowe;
- mając to na względzie zainteresowany organ sektora publicznego (lub ewentualnie prawodawca) powinien przeprowadzić ocenę skutków w zakresie ochrony danych, zanim ISP zawierające dane osobowe będą mogły być udostępnione do ponownego wykorzystania (lub przed przyjęciem przepisów zezwalających na upublicznienie danych osobowych, a tym samym czyniących je potencjalnie dostępnymi do ponownego wykorzystania); ocenę skutków w zakresie ochrony danych należy również przeprowadzić w sytuacjach, gdy do ponownego wykorzystywania będą udostępniane zanonimizowane zestawy danych uzyskane z danych osobowych;
- gdy zestawy danych zostały zanonimizowane, zasadnicze znaczenie ma ocena ryzyka ponownej identyfikacji oraz dobra praktyka w zakresie przeprowadzania testów na ponowną identyfikację;
- wyniki oceny mogłyby pomóc w określeniu odpowiednich zabezpieczeń minimalizujących ryzyko, w tym między innymi środków technicznych, prawnych i organizacyjnych, takich jak odpowiednie warunki licencji i środki techniczne uniemożliwiające masowe pobieranie danych, oraz odpowiednie techniki anonimizacji; wyniki te mogą również prowadzić do decyzji o rezygnacji z upublicznienia danych lub ich udostępniania do ponownego wykorzystania;
- warunki licencji ponownego wykorzystania ISP powinny zawierać klauzulę o ochronie danych w każdym przypadku, w którym przetwarzane są dane osobowe, w tym w sytuacjach, gdy do ponownego wykorzystania udostępnione zostaną zanonimizowane zestawy danych uzyskanych z danych osobowych;
- gdy w ocenie skutków w zakresie ochrony danych zostanie stwierdzone, że otwarta licencja nie wystarczy, by sprostać zagrożeniom dla ochrony danych, organy sektora publicznego nie powinny udostępniać danych osobowych na podstawie dyrektywy ISP. (Organ sektora publicznego może jednak nadal skorzystać ze swoich uprawnień, by rozważyć ponowne wykorzystanie danych poza warunkami i zakresem dyrektywy ISP, oraz może również wymagać od wnioskodawców wykazania, że kwestia jakiegokolwiek ryzyka dla ochrony danych osobowych została odpowiednio rozwiązana i że wnioskodawcy będą przetwarzać dane zgodne z właściwymi przepisami o ochronie danych.);
- w stosownych przypadkach organy sektora publicznego powinny dopilnować, by dane zostały zanonimizowane, a warunki licencjonowania wyraźnie zakazywały ponownej identyfikacji osób fizycznych i ponownego wykorzystania danych osobowych do celów, które mogą mieć wpływ na osoby, których dane dotyczą;
- państwa członkowskie powinny ponadto rozważyć ustanowienie i wspieranie sieci wiedzy/centrów doskonałości, a tym samym umożliwić wymianę dobrych praktyk w zakresie anonimizacji i otwartych danych.

Sporządzono w Brukseli dnia 5 czerwca 2013 r.

*W imieniu Grupy Roboczej
Przewodniczący
Jacob KOHNSTAMM*