OCHRONA DANYCH OSOBOWYCH

Kontrola i postępowanie w sprawie naruszenia przepisów

Poradnik ze wzorami

Mirosław Gumularz, Patrycja Kozik

EDYTOWALNE WZORY NA STRONIE INTERNETOWEJ

WYDANIE 2

Wolters Kluwer

Materiały uzupełniające na stronie:
www.ochrona-danych-osobowych-naruszenia-przepisow.wolterskluwer.pl
będą dostępne do 30 czerwca 2024 r.

W razie problemów
ze znalezieniem zdrapki
z kodem aktywacyjnym
prosimy o kontakt
tel. 801 04 45 45

PIKTOGRAMY

wskazują ważne elementy
książki i ułatwiają
ich odnalezienie


Ważne

Przykłady


Podstawa prawna
Kontekst prawny


Pytania
Zadania


Rozwiązania
Odpowiedzi


Stanowisko stron
Pogląd


Orzecznictwo


Literatura


Historia

Nowe przepisy


Wiadomości dodatkowe


Informacje powiązane
Odesłanie

OCHRONA DANYCH OSOBOWYCH

Kontrola i postępowanie w sprawie naruszenia przepisów

Poradnik ze wzorami

Mirosław Gumularz, Patrycja Kozik

WYDANIE 2

Wolters Kluwer

WARSZAWA 2022

Stan prawny na 1 marca 2022 r.

Wydawca
Monika Pawłowska

Redaktor prowadzący
Livia Śpiewak

Opracowanie redakcyjne
Violet Design Wioletta Kowalska

Projekt okładek serii
Wojtek Janikowski, Przemek Dębowski

ISBN 978-83-8286-233-1 *ISBN PDF-a: 978-83-8286-586-8*
2. wydanie

Wolters Kluwer Polska Sp. z o.o.
Dział Praw Autorskich
01-208 Warszawa, ul. Przyokopowa 33
tel. 728 313 462
e-mail: PL-ksiazki@wolterskluwer.com

księgarnia internetowa www.profinfo.pl

SPIS TREŚCI

Część IV
Tabele

Część V
Wzory pism

WYKAZ SKRÓTÓW

Akty prawne

dyrektywa 95/46/WE – dyrektywa 95/46/WE Parlamentu Europejskiego i Rady z 24.10.1995 r. w sprawie ochrony osób fizycznych w zakresie przetwarzania danych osobowych i swobodnego przepływu tych danych (Dz.Urz. WE L 281, s. 31, ze zm.)

dyrektywa policyjna – dyrektywa Parlamentu Europejskiego i Rady (UE) 2016/680 z 27.04.2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych przez właściwe organy do celów zapobiegania przestępczości, prowadzenia postępowań przygotowawczych, wykrywania i ścigania czynów zabronionych i wykonywania kar, w sprawie swobodnego przepływu takich danych oraz uchylająca decyzję ramową Rady 2008/977/WSiSW (Dz.Urz. UE L 119, s. 89)

k.c. – ustawa z 23.04.1964 r. – Kodeks cywilny (Dz.U. z 2020 r. poz. 1740 ze zm.)

k.k. – ustawa z 6.06.1997 r. – Kodeks karny (Dz.U. z 2021 r. poz. 2345 ze zm.)

k.p. – ustawa z 26.06.1974 r. – Kodeks pracy (Dz.U. z 2020 r. poz. 1320 ze zm.)

k.p.a. – ustawa z 14.06.1960 r. – Kodeks postępowania administracyjnego (Dz.U. z 2021 r. poz. 735 ze zm.)

o.p. – ustawa z 29.08.1997 r. – Ordynacja podatkowa (Dz.U. z 2021 r. poz. 1540 ze zm.)

p.p.s.a. – ustawa z 30.08.2002 r. – Prawo o postępowaniu przed sądami administracyjnymi (Dz.U. z 2022 r. poz. 329)

pr. adw. – ustawa z 26.05.1982 r. – Prawo o adwokaturze (Dz.U. z 2020 r. poz. 1651 ze zm.)

pr. przed. – ustawa z 6.03.2018 r. – Prawo przedsiębiorców (Dz.U. z 2021 r. poz. 162 ze zm.)

RODO – rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z 27.04.2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych) (Dz.Urz. UE L 119, s. 1, ze zm.)

r.w.s.z.p. – rozporządzenie Rady Ministrów z 16.12.2003 r. w sprawie wysokości oraz szczegółowych zasad pobierania wpisu w postępowaniu przed sądami administracyjnymi (Dz.U. z 2021 r. poz. 535)

u.ABW – ustawa z 24.05.2002 r. o Agencji Bezpieczeństwa Wewnętrznego oraz Agencji Wywiadu (Dz.U. z 2020 r. poz. 27 ze zm.)

u.COVID – ustawa z 2.03.2020 r. o szczególnych rozwiązaniach związanych z zapobieganiem, przeciwdziałaniem i zwalczaniem COVID-19, innych chorób zakaźnych oraz wywołanych nimi sytuacji kryzysowych (Dz.U. z 2021 r. poz. 2095 ze zm.)

u.f.p. – ustawa z 27.08.2009 r. o finansach publicznych (Dz.U. z 2021 r. poz. 305 ze zm.)

u.o.d.o. – ustawa z 10.05.2018 r. o ochronie danych osobowych (Dz.U. z 2019 r. poz. 1781 ze zm.)

u.o.d.o.1997 – ustawa z 29.08.1997 r. o ochronie danych osobowych (Dz.U. z 2016 r. poz. 922 ze zm.)

u.o.d.z.p. / DODO – ustawa z 14.12.2018 r. o ochronie danych osobowych przetwarzanych w związku z zapobieganiem i zwalczaniem przestępczości (Dz.U. z 2019 r. poz. 125)

u.o.i.n. – ustawa z 5.08.2010 r. o ochronie informacji niejawnych (Dz.U. z 2019 r. poz. 742)

u.o.k.k. – ustawa z 16.02.2007 r. o ochronie konkurencji i konsumentów (Dz.U. z 2021 r. poz. 275)

u.s.d.g. – ustawa z 2.07.2004 r. o swobodzie działalności gospodarczej (Dz.U. z 2017 r. poz. 2168 ze zm.) – uchylona

u.z.n.u. – ustawa z 21.02.2019 r. o zmianie niektórych ustaw w związku z zapewnieniem stosowania rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobod-

nego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych) (Dz.U. poz. 730)

Inne

ABW – Agencja Bezpieczeństwa Wewnętrznego
CNIL – Commission Nationale de l'Informatique et des Libertés
ENISA – Europejska Agencja ds. Bezpieczeństwa Sieci i Informacji (EU Agency for Network and Information Security)
EROD – Europejska Rada Ochrony Danych; dawniej Grupa Robocza Art. 29
GGK – Główny Geodeta Kraju
GIODO – Generalny Inspektor Ochrony Danych Osobowych
ICO – International Commissioner's Office
IOD – inspektor ochrony danych
NSA – Naczelny Sąd Administracyjny
ONSAiWSA – Orzecznictwo Naczelnego Sądu Administracyjnego i Wojewódzkich Sądów Administracyjnych
Prezes UODO – Prezes Urzędu Ochrony Danych Osobowych (także jako Prezes Urzędu)
UODO – Urząd Ochrony Danych Osobowych
WSA – wojewódzki sąd administracyjny

Część I

PODZIAŁY DANYCH OSOBOWYCH OBJĘTYCH KONTROLĄ

1. Dane osobowe wrażliwe (szczególne kategorie danych), dane o karalności oraz dane osobowe zwykłe

Przepisy RODO wyróżniają następujące kategorie danych osobowych:

1) szczególne kategorie danych osobowych;
2) dane osobowe dotyczące wyroków skazujących i czynów zabronionych;
3) inne kategorie danych osobowych (dane osobowe zwykłe).

Szczególne kategorie danych osobowych w rozumieniu RODO to dane osobowe:

1) ujawniające pochodzenie rasowe lub etniczne;
2) ujawniające poglądy polityczne;
3) ujawniające przekonania religijne lub światopoglądowe;
4) ujawniające przynależność do związków zawodowych;
5) genetyczne (tj. dane osobowe dotyczące odziedziczonych lub nabytych cech genetycznych osoby fizycznej, które ujawniają niepowtarzalne informacje o fizjologii lub zdrowiu tej osoby i które wynikają w szczególności z analizy próbki biologicznej pochodzącej od tej osoby fizycznej);
6) biometryczne (tj. dane osobowe, które wynikają ze specjalnego przetwarzania technicznego, dotyczą cech fizycznych, fizjologicznych lub behawioralnych osoby fizycznej oraz umożliwiają lub potwierdzają jednoznaczną identyfikację tej osoby, takie jak wizerunek twarzy lub dane daktyloskopijne);
7) dotyczące zdrowia (tj. dane osobowe o zdrowiu fizycznym lub psychicznym osoby fizycznej, w tym o korzystaniu z usług opieki zdrowotnej, ujawniające informacje o stanie jej zdrowia);
8) dotyczące seksualności lub orientacji seksualnej osoby.

Należy zwrócić uwagę, że katalog szczególnych kategorii danych osobowych (danych wrażliwych) ma charakter zamknięty, chociaż niektóre z nich są określone w sposób bardzo ogólny, np. dane dotyczące seksualności. EROD wskazuje, że przetwarzanie szczególnych kategorii danych, np. o stanie zdrowia, może wynikać z możliwości wnioskowania tych informacji z danych zwykłych. Założenia lub wnioski dotyczące danych kategorii specjalnej, np. że dana osoba po odwiedzeniu strony głoszącej poglądy liberalne prawdopodobnie zagłosuje na daną partię, również stanowiłyby specjalną kategorię danych osobowych. Podobnie – jak wskazuje EROD – profilowanie może stworzyć specjalną kategorię danych, wnioskując z danych, które same w sobie nie są specjalną kategorią danych, ale stają się takie w połączeniu z innymi danymi. Na przykład można wywnioskować czyjś stan zdrowia na podstawie zapisów dotyczących zakupów żywności w połączeniu z danymi o jakości i wartości energetycznej żywności[1].

! Odrębną kategorią danych są dane dotyczące wyroków skazujących oraz czynów zabronionych lub powiązanych środków bezpieczeństwa[2]. Jak wskazuje art. 10 RODO, przetwarzania danych osobowych dotyczących wyroków skazujących oraz czynów zabronionych lub powiązanych środków bezpieczeństwa na podstawie art. 6 ust. 1 RODO można dokonywać wyłącznie pod nadzorem władz publicznych lub jeżeli przetwarzanie jest dozwolone prawem UE lub prawem państwa członkowskiego przewidującymi odpowiednie zabezpieczenia praw

[1] EROD, *Wytyczne 8/2020 dotyczące targetowania użytkowników mediów społecznościowych*, v. 2.0, 13.04.2021 r., nb 121, https://edpb.europa.eu/system/files/2021-11/edpb_guidelines_082020_on_the_targeting_of_social_media_users_pl_0.pdf (dostęp: 15.02.2022 r.).

[2] P. Litwiński, P. Barta, M. Kawecki [w:] *Rozporządzenie UE w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i swobodnym przepływem takich danych. Komentarz*, red. P. Litwiński, Warszawa 2018, art. 10, nt 3–4, przyjmują następujące rozumienie wskazanych elementów:

1) informacje o „naruszeniach prawa" – informacje o przypadkach popełnienia przestępstwa (naruszenia obowiązującego prawa) przez daną osobę fizyczną, ale niestwierdzonych wyrokami sądowymi (np. informacja o nałożeniu mandatu karnego, który ukarany przyjął),
2) powiązane środki bezpieczeństwa – środki zabezpieczające, które zostały wskazane w art. 93a § 1 k.k.

i wolności osób, których dane dotyczą. Wszelkie kompletne rejestry wyroków skazujących są prowadzone wyłącznie pod nadzorem władz publicznych.

Co istotne, zgodnie z treścią art. 10 RODO przedstawione kategorie danych, tj. dane dotyczące wyroków skazujących oraz czynów zabronionych lub powiązanych środków bezpieczeństwa, mogą być przetwarzane na podstawie art. 6 RODO (a nie art. 9 RODO, który dotyczy szczególnych kategorii danych osobowych). Wskazane kategorie danych mogą być przetwarzane wyłącznie pod nadzorem władz publicznych lub jeżeli przetwarzanie jest dozwolone prawem UE lub prawem państwa członkowskiego przewidującymi odpowiednie zabezpieczenia praw i wolności osób, których dane dotyczą.

Pozostałe kategorie danych osobowych to dane osobowe zwykłe. Dane osobowe zwykłe to wszystkie inne kategorie niż wskazane powyżej szczególne kategorie danych osobowych oraz dotyczące wyroków skazujących i czynów zabronionych.

2. Dane osobowe podane, zaobserwowane, pochodne oraz wywnioskowane

Dla zrozumienia, czym są dane osobowe, istotne jest wyróżnienie następujących kategorii[3]:

1) podane dane osobowe – dane osobowe świadomie podawane przez osoby fizyczne, np. przy wypełnianiu formularza online;
2) zaobserwowane dane osobowe – dane osobowe rejestrowane automatycznie, np. za pomocą plików cookies lub czujników online, lub telewizji przemysłowej umożliwiającej rozpoznanie twarzy;
3) pochodne dane osobowe – dane osobowe „tworzone" z innych danych w stosunkowo prosty i bezpośredni sposób, np. przy obliczaniu zdolności kredytowej klienta;

[3] ICO, *Big data, artificial intelligence, machine learning and data protection*, v. 2.2, 4.09.2017 r., s. 12, https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf (dostęp: 15.02.2022 r.).

4) wywnioskowane dane osobowe – dane osobowe tworzone za pomocą bardziej złożonych metod analitycznych (np. poprzez odszukiwanie korelacji między zestawami danych i wykorzystanie tych korelacji do kategoryzacji lub profilowania osób, np. obliczania wyników kredytowych lub przewidywania przyszłych wyników zdrowotnych). Dane wywnioskowane są oparte na prawdopodobieństwach i w związku z tym są mniej pewne niż dane pochodne.

Pojęcie danych osobowych obejmuje wszystkie wskazane kategorie. Potwierdza to wyrok Trybunału Sprawiedliwości z 20.12.2017 r., C-434/16, Peter Nowak przeciwko Data Protection Commissioner[4]. W przedmiotowej sprawie chodziło o odpowiedź na pytania:

1. Czy informacje zapisane przez osobę przystępującą do egzaminu zawodowego w odpowiedziach lub jako odpowiedzi na arkuszu egzaminacyjnym mogą stanowić dane osobowe w rozumieniu dyrektywy 95/46/WE?
2. Jakie czynniki są istotne dla określenia, czy w danym przypadku taki arkusz egzaminacyjny stanowi dane osobowe, i jakie znaczenie należy przypisać tym czynnikom?

We wskazanym orzeczeniu Trybunał Sprawiedliwości orzekł, że:

1) użycie w zawartej w art. 2 lit. a dyrektywy 95/46/WE definicji danych osobowych wyrażenia „wszelkie informacje" odzwierciedla przyświecający unijnemu prawodawcy zamiar przypisania temu pojęciu szerokiego zakresu znaczeniowego, nie ograniczając go do informacji wrażliwych czy też prywatnych, lecz rozszerzając go potencjalnie o informacje wszelkiego rodzaju, zarówno obiektywne, jak i subiektywne, w postaci opinii czy oceny, a jedynym warunkiem, które muszą one spełniać, jest to, aby „dotyczyły" danej osoby. Warunek ten jest zaś spełniony, gdy ze względu na swą treść, cel czy skutek dana informacja jest powiązana z daną osobą;
2) treść odpowiedzi egzaminacyjnych odzwierciedla poziom wiedzy i umiejętności osoby przystępującej do egzaminu w danej dziedzinie oraz, w odpowiednim przypadku, sposób jej rozumowania, wnioskowania i przyjęte przez nią krytyczne podejście. W przypadku

[4] EU:C:2017:994.

egzaminu pisanego odręcznie odpowiedzi te zawierają poza tym informacje odnoszące się do charakteru pisma;

3) w odniesieniu do naniesionych przez egzaminatora komentarzy dotyczących tych odpowiedzi należy stwierdzić, że stanowią one, tak jak udzielone na egzaminie przez przystępującą do niego osobę odpowiedzi, informacje dotyczące tej właśnie osoby. Treść tych komentarzy odzwierciedla bowiem wyrażoną przez egzaminatora opinię czy ocenę indywidualnych osiągnięć danej osoby na tym egzaminie, a w szczególności poziomu jej wiedzy i umiejętności w danej dziedzinie. To, że naniesione przez egzaminatora komentarze odnoszące się do odpowiedzi udzielonych na egzaminie przez osobę do niego przystępującą są również informacjami dotyczącymi egzaminatora, nie stoi na przeszkodzie uznaniu, że te komentarze, ze względu na ich treść, cel czy skutek, są powiązane z egzaminowaną osobą;
4) ta sama informacja może dotyczyć więcej niż jednej osoby fizycznej i stanowić w odniesieniu do tych osób dane osobowe w rozumieniu art. 2 lit. a dyrektywy 95/46/WE, pod warunkiem że osoby te zostały zidentyfikowane lub są możliwe do zidentyfikowania.

Konkludując, Trybunał Sprawiedliwości na zadane pytania odpowiedział, że pojęcie danych osobowych należy interpretować w ten sposób, że odpowiedzi udzielone przez osobę przystępującą do egzaminu zawodowego i ewentualne naniesione przez egzaminatora komentarze odnoszące się do tych odpowiedzi stanowią dane osobowe w rozumieniu tego przepisu. Wyrok ten zachował swą aktualność także na gruncie RODO.

Co istotne, w ramach tego samego celu, w zależności czy dochodzić będzie do przetwarzania danych podanych, czy wywnioskowanych lub pochodnych, inne mogą być podstawy przetwarzania danych osobowych. Zwraca na to uwagę EROD w opinii dotyczącej mediów społecznościowych[5]. Podobnie inny może być zakres realizacji praw. Dotyczy to przykładowo prawa do przenoszenia danych. Grupa Robocza Art. 29 (obecnie EROD), w kontekście prawa do przenoszenia danych, wskazała, że można wyróżnić różne kategorie danych w zależności od ich pochodzenia, aby określić,

[5] EROD, *Wytyczne 8/2020…*, nb 121.

czy są one objęte prawem do przenoszenia danych. W dalszej części Grupa wskazuje, że następujące kategorie danych można zaklasyfikować jako „dostarczone przez osobę, której dane dotyczą":

- dane aktywnie i świadomie dostarczone przez osobę, której dane dotyczą (np. adres e-mail, nazwa użytkownika, wiek itd.);
- zaobserwowane dane dostarczone przez osobę, której dane dotyczą, poprzez korzystanie z danej usługi lub danego urządzenia. Mogą one np. obejmować historię wyszukiwania danej osoby, dane o ruchu i dane dotyczące lokalizacji. Mogą również obejmować inne dane pierwotne, takie jak tętno monitorowane za pośrednictwem urządzenia do noszenia na ciele[6].

3. Dane uporządkowane oraz nieuporządkowane

3.1. Ogólne informacje

Nie każde przetwarzanie danych osobowych objęte jest regulacją RODO. Zgodnie z art. 2 ust. 1 RODO ma ono zastosowanie do przetwarzania danych osobowych w sposób całkowicie lub częściowo zautomatyzowany oraz do przetwarzania w sposób inny niż zautomatyzowany danych osobowych stanowiących część zbioru danych lub mających stanowić część zbioru danych. Pojęcie zbioru danych zgodnie z RODO obejmuje uporządkowany zestaw danych osobowych dostępnych według określonych kryteriów, niezależnie od tego, czy zestaw ten jest scentralizowany, zdecentralizowany, czy rozproszony funkcjonalnie lub geograficznie (art. 4 pkt 6 RODO).

Treść wskazanego przepisu art. 2 ust. 1 RODO może rodzić pewne wątpliwości interpretacyjne. Wykładnia językowa może prowadzić do dwóch rozstrzygnięć:

1) niezależnie od sposobu przetwarzania danych (zautomatyzowane lub niezautomatyzowane) RODO obejmuje wyłącznie przetwarzanie

[6] Grupa Robocza Art. 29, *Wytyczne dotyczące prawa do przenoszenia danych*, przyjęte 13.12.2016 r., ostatnio zmienione i przyjęte 5.04.2017 r., 16/EN WP 242, rev. 01, pkt III, https://uodo.gov.pl/pl/file/14 (dostęp: 15.02.2022 r.).

danych osobowych stanowiących część zbioru danych lub mających stanowić część zbioru danych;
2) RODO obejmuje przetwarzanie danych osobowych w sposób zautomatyzowany (np. systemy informatyczne), niezależnie od tego, czy dane osobowe są lub mają być uporządkowane według określonych kryteriów.

Powyższą kwestię ustawa o ochronie danych osobowych z 1997 r. rozstrzygała, rozróżniając przetwarzanie tradycyjne w uporządkowanych zestawieniach oraz w systemach informatycznych, niezależnie od tego uporządkowania. Zgodnie z art. 2 ust. 2 u.o.d.o.1997 ustawę stosowało się do przetwarzania danych osobowych:
1) w kartotekach, skorowidzach, księgach, wykazach i w innych zbiorach ewidencyjnych;
2) w systemach informatycznych, także w przypadku przetwarzania danych poza zbiorem danych.

Rozstrzygając powyższy problem na poziomie wykładni funkcjonalnej i historycznej, należy przyjąć drugi ze wskazanych wariantów interpretacyjnych. Odpowiada on bardziej potrzebie ochrony danych osobowych, chociaż oczywiście z punktu widzenia administratorów może rodzić problemy w zarządzaniu procesami przetwarzania danych. Powyższą kwestię nieco rozjaśnia motyw 15 RODO, zgodnie z którym, aby zapobiec poważnemu ryzyku obchodzenia prawa, ochrona osób fizycznych jest neutralna pod względem technicznym i nie zależy od stosowanych technik. Ochrona osób fizycznych ma zastosowanie do zautomatyzowanego przetwarzania danych osobowych oraz do przetwarzania ręcznego, jeżeli dane osobowe znajdują się lub mają się znaleźć w zbiorze danych. Zbiory lub zestawy zbiorów oraz ich strony tytułowe, które nie są uporządkowane według określonych kryteriów, nie powinny być objęte zakresem RODO.

Oznacza to, że w sytuacji, gdy dane osobowe są przetwarzane w sposób niezautomatyzowany (np. tradycyjne teczki z danymi), RODO będzie miało zastosowanie wyłącznie wtedy, gdy dane są lub mają tworzyć uporządkowany zestaw danych osobowych dostępnych według określonych kryteriów. Natomiast gdy dane będą przetwarzane w sposób

zautomatyzowany (np. w systemach informatycznych), niezależnie od wskazanego uporządkowania RODO znajdzie zastosowanie[7]. Z punktu widzenia definicji zbioru danych, a także wskazanego powyżej motywu 15 RODO należy zwrócić uwagę, że chodzi o uporządkowanie danych osobowych, które są dostępne według określonego kryterium. Chodzi więc o coś więcej niż np. uporządkowanie nośników danych bez porządkowania samych danych osobowych (np. uporządkowanie dokumentów według daty wpływu bez możliwości „wyszukania" danych osobowych).

Należy zwrócić uwagę, że zgodnie z art. 2 ust. 1 RODO przepisom RODO podlega przetwarzanie zautomatyzowane, natomiast samo RODO wyróżnia jeszcze kategorię przetwarzania w sposób wyłącznie zautomatyzowany w art. 22 ust. 1. Przepis ten wskazuje, że osoba, której dane dotyczą, ma prawo do tego, by nie podlegać decyzji, która opiera się wyłącznie na zautomatyzowanym przetwarzaniu, w tym profilowaniu, i wywołuje wobec tej osoby skutki prawne lub w podobny sposób istotnie na nią wpływa. *A contrario* trzeba przyjąć, że pojęcie zautomatyzowania z art. 2 RODO obejmuje przetwarzanie całkowicie lub częściowo zautomatyzowane. Co istotne, z punktu widzenia pojęcia zautomatyzowanego przetwarzania danych bez znaczenia pozostaje okoliczność, w jaki sposób dane zostały pozyskane (np. dane pozyskano w tradycyjny sposób i następnie przetwarzano w systemach informatycznych).

W kontekście zakresu przedmiotowego RODO istotne znaczenie ma wyrok TSUE z 10.07.2018 r., C-25/17, Tietosuojavaltuutettu[8]. Wskazane orzeczenie zostało wydane na gruncie dyrektywy 95/46/WE, jednak znajdzie zastosowanie także w odniesieniu do RODO. Po pierwsze, potwierdza ono, że problem uporządkowania (tj. przetwarzania w zbiorze danych) dotyczy wyłącznie „ręcznego" przetwarzania danych (niezautomatyzowanego) (zob. pkt 53 wyroku). Po drugie, dookreśla pojęcie zbioru danych, wskazując, że pojęcie zbioru obejmuje zestaw danych osobowych, o ile dane te są zorganizowane według określonych kryteriów umożliwiających w praktyce ich łatwe odnalezienie dla celów ich

[7] Tak np. P. Voigt, A. von dem Bussche, *The EU General Data Protection Regulation (GDPR). A Practical Guide*, Basel 2017, s. 9–10.

[8] EU:C:2018:551.

późniejszego wykorzystania. Aby zestaw taki był objęty tym pojęciem, nie jest konieczne, by zawierał kartoteki, szczególne rejestry lub inne systemy służące wyszukiwaniu (zob. pkt 62 wyroku).

Co istotne, RODO – tylko w kontekście zakresu przedmiotowego – odwołuje się do pojęcia zbioru danych[9]. Nie można jednak powiedzieć, że poza tą kwestią okoliczność, czy dane osobowe są przetwarzane w sposób uporządkowany, czy nie, nie ma znaczenia normatywnego. Zautomatyzowane przetwarzanie danych osobowych w nieuporządkowanych zestawieniach danych może jednak wpływać na zakres obowiązków administratora danych, o czym będzie mowa poniżej. W tym przypadku odszukanie (zidentyfikowanie) danych osobowych – chociażby w celu realizacji praw podmiotów danych – ze względu na brak kryterium umożliwiającego dostęp do danych może w pewnych sytuacjach być utrudnione. Wszystko jednak zależeć będzie od kontekstu. Można sobie wyobrazić sytuację, gdy administrator danych jest w stanie „odszukać" informacje przypisywalne do konkretnej osoby, nawet jeżeli nie zostały one uporządkowane w sposób umożliwiający dostęp do nich według określonych kryteriów (np. nieuporządkowane e-maile wyszukiwane poprzez wpisanie adresu nadawcy).

Co ważne, brak w RODO przepisu, który wymagałby od administratora danych porządkowania danych osobowych według kryteriów umożliwiających dostęp do tych danych. Jak się zdaje, brak takiego obowiązku potwierdza treść art. 11 ust. 1 RODO: „Jeżeli cele,

[9] Do pojęcia zbioru danych RODO odwołuje się także w motywie 31: „Organy publiczne, którym ujawnia się dane osobowe w związku z ich prawnym obowiązkiem sprawowania funkcji publicznej (takich jak organy podatkowe, organy celne, finansowe jednostki analityki finansowej, niezależne organy administracyjne czy organy rynków finansowych regulujące i nadzorujące rynki papierów wartościowych), nie powinny być traktowane jako odbiorcy, jeżeli otrzymane przez nie dane osobowe są im niezbędne do przeprowadzenia określonego postępowania w interesie ogólnym zgodnie z prawem Unii lub prawem państwa członkowskiego. Żądanie ujawnienia danych osobowych, z którym występują takie organy publiczne, powinno zawsze mieć formę pisemną, być uzasadnione, mieć charakter wyjątkowy, nie powinno dotyczyć całego zbioru danych ani prowadzić do połączenia zbiorów danych. Przetwarzając otrzymane dane osobowe, takie organy powinny przestrzegać mających zastosowanie przepisów o ochronie danych, zgodnie z celami przetwarzania".

w których administrator przetwarza dane osobowe, nie wymagają lub już nie wymagają zidentyfikowania przez niego osoby, której dane dotyczą, administrator nie ma obowiązku zachowania, uzyskania ani przetworzenia dodatkowych informacji w celu zidentyfikowania osoby, której dane dotyczą, wyłącznie po to, by zastosować się do niniejszego rozporządzenia". Klucz porządkujący dane może być traktowany właśnie jako dodatkowa informacja umożliwiająca identyfikację. Naszym zdaniem z RODO co do zasady nie wynika obowiązek porządkowania danych po to, by móc zidentyfikować osobę np. w poszczególnych dokumentach. Nie ma także obowiązku wdrażania systemów informatycznych do porządkowania danych lub odszukiwania danych pomimo braku uporządkowania. Należy jednak uwzględnić dwie kwestie. Po pierwsze, przepisy szczególne (np. dotyczące porządkowania akt osobowych) mogą w tym zakresie wprowadzać obowiązek. Po drugie, w pewnych sytuacjach ocena ryzyka, o której mowa w art. 25 RODO (*privacy by design*), może taki obowiązek aktualizować. Źródłem potencjalnego ryzyka w tym przypadku może być scenariusz naruszenia wymogu poprawności danych.

Natomiast jest kwestią otwartą, czy administrator powinien wdrożyć system do strukturyzowania (zidentyfikowanych/odszukanych) nieuporządkowanych danych osobowych w celu realizacji jednego z praw, tj. prawa do przenoszenia danych (w tym przypadku chodzi nie o problem identyfikacji osoby fizycznej, ale o kwestię realizacji prawa do przenoszenia danych)[10].

W tym zakresie jako niespójny systemowo może się wydawać art. 20 ust. 1 RODO, zgodnie z którym: „Osoba, której dane dotyczą, ma prawo

[10] W tym zakresie pomocny (chociaż nierozstrzygający) jest motyw 68 RODO: „Przysługujące osobie, której dane dotyczą, prawo do przesłania lub otrzymania swoich danych osobowych nie powinno nakładać na administratorów obowiązku prowadzenia lub wprowadzenia kompatybilnych technicznie systemów przetwarzania. Jeżeli określony zestaw danych osobowych odnosi się do więcej niż jednej osoby, której dane dotyczą, prawo do otrzymania danych osobowych nie powinno powodować uszczerbku dla praw i wolności innych osób, których dane dotyczą, na podstawie niniejszego rozporządzenia". Można przyjąć, że administrator danych ma obowiązek wydać dane w sposób ustrukturyzowany (jeżeli przesłanki prawa do przenoszenia danych zostaną spełnione), nawet jeżeli dane są przetwarzane w sposób nieuporządkowany, przy uwzględnieniu dostępnej mu technologii.

otrzymać w ustrukturyzowanym, powszechnie używanym formacie nadającym się do odczytu maszynowego dane osobowe jej dotyczące, które dostarczyła administratorowi, oraz ma prawo przesłać te dane osobowe innemu administratorowi bez przeszkód ze strony administratora, któremu dostarczono te dane osobowe, jeżeli:

a) przetwarzanie odbywa się na podstawie zgody w myśl art. 6 ust. 1 lit. a) lub art. 9 ust. 2 lit. a) lub na podstawie umowy w myśl art. 6 ust. 1 lit. b); oraz
b) przetwarzanie odbywa się w sposób zautomatyzowany".

Jak wynika z treści powyższego przepisu, obowiązek wydania ustrukturyzowanych danych dotyczy sytuacji, gdy są one przetwarzane w sposób zautomatyzowany. Jak to jednak zostało wskazane wyżej, dane osobowe przetwarzane w sposób zautomatyzowany podlegają RODO, nawet jeżeli nie są lub nie mają być częścią uporządkowanego zestawu. Pojawia się w związku z tym pytanie, czy administrator danych ma obowiązek porządkowania danych tylko po to, aby zrealizować prawo wynikające z art. 20 RODO (tj. gdy przetwarza dane w sposób zautomatyzowany, ale nieuporządkowany). Po pierwsze, można sobie wyobrazić odczytywalne maszynowo ustrukturyzowane dane, jednak nieuporządkowane i niestanowiące zbioru danych osobowych. Być może właśnie dlatego ustawodawca unijny w treści art. 20 RODO nie posłużył się kryterium uporządkowania. W dalszym ciągu można odwołać się do treści art. 11 ust. 2 RODO, zgodnie z którym, jeżeli w przypadkach, o których mowa w art. 11 ust. 1 RODO, administrator może wykazać, że nie jest w stanie zidentyfikować osoby, której dane dotyczą, w miarę możliwości informuje o tym osobę, której dane dotyczą. W takich przypadkach zastosowania nie ma m.in. art. 20 RODO, chyba że osoba, której dane dotyczą, w celu wykonania przysługujących jej praw dostarczy dodatkowych informacji pozwalających ją zidentyfikować.

! Jak się zdaje, nie każde uporządkowanie daje możliwość przyjęcia zbioru danych w rozumieniu RODO. Istotna z perspektywy definicji zbioru danych jest dostępność danych osobowych według określonych kryteriów. Przykładowo, uporządkowanie dokumentów dotyczących określonych spraw według wpływu wcale nie musi umożliwiać łatwego dostępu do konkretnych danych osobowych.

Powyższe oznacza, że RODO może znaleźć zastosowanie także do danych nieuporządkowanych, jeżeli są przetwarzane w sposób zautomatyzowany. Przykładem obszarów, w których może (w zależności od oceny stanu faktycznego) dochodzić do przetwarzania (zautomatyzowanego) danych nieuporządkowanych, mogą być:

- e-maile,
- pliki PDF,
- obrazy cyfrowe (zdjęcia akt),
- pliki wideo i audio.

Tak jak to zostało zaznaczone wcześniej, zautomatyzowane przetwarzanie danych nieuporządkowanych może – ale nie musi – rodzić problemy z możliwością przypisania określonych informacji do wyodrębnionej osoby fizycznej.

3.2. Uporządkowanie a ustrukturyzowanie

Kwestia ta będzie jeszcze wyjaśniana, jednak już w tym miejscu należy zwrócić uwagę, że zgodnie z cytowanym wyżej art. 20 ust. 1 RODO, jeżeli przetwarzanie odbywa się:

a) na podstawie zgody w myśl art. 6 ust. 1 lit. a lub art. 9 ust. 2 lit. a lub na podstawie umowy w myśl art. 6 ust. 1 lit. b RODO oraz
b) w sposób zautomatyzowany,

to osoba, której dane dotyczą, ma prawo otrzymać w ustrukturyzowanym, powszechnie używanym formacie nadającym się do odczytu maszynowego dane osobowe jej dotyczące, które dostarczyła administratorowi, oraz ma prawo przesłać te dane osobowe innemu administratorowi bez przeszkód ze strony administratora, któremu dostarczono te dane osobowe[11].

[11] ICO przyjmuje, że dane ustrukturyzowane to: „data where the structural relation between elements is explicit in the way the data is stored on a computer disk". Zob. ICO, *Guide to the General Data Protection Regulation (GDPR)*, v. 1.0.248, 2.08.2018 r., s. 131, https://ico.org.uk/media/for-organisations/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf (dostęp: 15.02.2022 r.).

W tym kontekście należy zwrócić uwagę, że art. 20 ust. 1 RODO:

1) nie odwołuje się do pojęcia zbioru danych (uporządkowanego zestawu), ale ustrukturyzowania. Istotne, że pojęcie zbioru odwołuje się do uporządkowania danych, natomiast pojęcie ustrukturyzowania – do określonego formatu. Może być tak, że dane osobowe przetwarzane w uporządkowanych zestawach – umożliwiających dostęp do danych – nie będą stanowić ustrukturyzowanego formatu, który będzie odczytywalny maszynowo;
2) odwołuje się do przetwarzania w sposób zautomatyzowany, ale niezależnie od tego, czy w zbiorach danych, czy nie (kwestia ta została omówiona powyżej).

Oczywiście realizacja prawa do przenoszenia danych z art. 20 RODO jest uzależniona od zidentyfikowania danych i weryfikacji tożsamości osoby fizycznej (zob. art. 11 i 12 RODO).

4. Dane identyfikowalne oraz nieidentyfikowalne

4.1. Dwie normy wynikające z art. 11 RODO

W praktyce często podział na dane osobowe uporządkowane albo nieuporządkowane mylony jest z innym podziałem danych osobowych: przetwarzanych w sposób wymagający i niewymagający identyfikacji. Kwestii tej dotyczy art. 11 RODO, obejmujący swoim zakresem dwie normy. Po pierwsze, zgodnie z art. 11 ust. 1 RODO administrator nie ma obowiązku zachowania, uzyskania ani przetworzenia dodatkowych informacji w celu zidentyfikowania osoby, której dane dotyczą, wyłącznie po to, by zastosować się do RODO. Przepis ten dotyczy sytuacji, gdy cele, w których administrator przetwarza dane osobowe, nie wymagają lub już nie wymagają zidentyfikowania przez niego osoby, której dane dotyczą. Normę tę należy traktować jako logiczną konsekwencję zasady minimalizacji danych i ograniczenia przechowywania danych osobowych i – jak się zdaje – nawet gdyby RODO nie zawierało art. 11 ust. 1, podobną normę należałoby wyinterpretować poprzez wykładnię art. 5 ust. 1 lit. c oraz e RODO.

Druga norma wynika z art. 11 ust. 2 RODO. Zgodnie z tym przepisem w przypadkach wskazanych w art. 11 ust. 1 RODO, tj. gdy administrator danych może wykazać, że bez dodatkowych informacji nie jest w stanie zidentyfikować osoby, której dane dotyczą, ma obowiązek w miarę możliwości poinformować o tym osobę, której dane dotyczą. W takich przypadkach zastosowania nie mają art. 15–20 RODO (dotyczące realizacji praw podmiotów danych i wymagające weryfikacji tożsamości), chyba że osoba, której dane dotyczą, w celu wykonania praw przysługujących jej na mocy tych artykułów dostarczy dodatkowych informacji pozwalających ją zidentyfikować.

Oczywiście często dane nieuporządkowane będą jednocześnie przetwarzane w sposób uniemożliwiający identyfikację osoby fizycznej bez dodatkowych danych (np. w przypadku nagrań monitoringu bez informacji o tym, kiedy i kto został nagrany). Są to jednak dwa odrębne wyróżnienia, które się krzyżują.

4.2. Artykuł 11 RODO a definicja danych osobowych

Rozróżnienie przetwarzania danych wymagających i niewymagających identyfikacji jest ściśle związane z samym pojęciem danych osobowych. Jak wskazano powyżej, dane osobowe to wszelkie informacje o zidentyfikowanych lub możliwych do zidentyfikowania osobach fizycznych. Można wręcz powiedzieć, że treść art. 11 RODO dookreśla definicje danych osobowych.

W obszarze przetwarzania informacji dotyczących „możliwych do zidentyfikowania" osób fizycznych mieści się m.in.:

1) pseudonimizacja (dotyczy przetwarzania danych osobowych w taki sposób, by nie można ich było już przypisać konkretnej osobie, której dane dotyczą, bez wykorzystania dodatkowych informacji, pod warunkiem że takie dodatkowe informacje są przechowywane osobno i są objęte środkami technicznymi i organizacyjnymi uniemożliwiającymi ich przypisanie zidentyfikowanej lub możliwej do zidentyfikowania osobie fizycznej);

2) sytuacja określona w art. 11 ust. 2 RODO, tj. gdy administrator może wykazać, że nie jest w stanie zidentyfikować osoby bez dodatkowych danych.

W powyższym kontekście należy zadać pytanie: czy art. 11 ust. 2 RODO dotyczy sytuacji, gdy te dodatkowe dane osobowe – umożliwiające identyfikację – znajdują się w organizacji administratora danych, czy tylko sytuacji, gdy znajdują się poza organizacją administratora danych i mogą zostać dostarczone m.in. przez podmiot danych lub pozyskane z innych źródeł? Ten pierwszy przypadek może dotyczyć sytuacji, gdy dane osobowe przetwarzane są poza zbiorem danych i w związku z tym brak znanego administratorowi kryterium umożliwiającego dostęp do danych. Jeszcze inaczej ujmując powyższe pytanie: czy problem z identyfikacją osoby fizycznej (regulowany treścią art. 11 ust. 2 RODO) można rozumieć jako dotyczący także samego „odszukania" danych (np. gdy wymagałoby to wielomiesięcznej analizy danych archiwalnych)?

Odpowiadając na sformułowane pytanie, należy wyjść od stwierdzenia, że art. 11 ust. 2 RODO dotyczy sytuacji, gdy administrator danych może wykazać, że nie jest w stanie zidentyfikować osoby bez dodatkowych danych, ale jednocześnie cały czas chodzi o osobę możliwą do zidentyfikowania. Chodzi więc o przetwarzanie danych osobowych. W przeciwnym razie art. 11 ust. 2 RODO, ograniczający pewne wymogi RODO, nie miałby sensu. Jeżeli informacje nie mają charakteru danych osobowych, to nie podlegają w ogóle RODO. Oznacza to, że treścią art. 11 ust. 2 RODO objęte są sytuacje „utrudnionej" identyfikacji, która jednak cały czas odnosi się do danych osobowych. Dodatkowo należy zwrócić uwagę, że art. 11 ust. 2 RODO posługuje się elastyczną formułą, wskazując, że chodzi o sytuacje, gdy „administrator może wykazać, że nie jest w stanie zidentyfikować osoby, której dane dotyczą".

Wydaje się, że można przyjąć, iż art. 11 ust. 2 RODO dotyczy także sytuacji, gdy administrator danych nie jest w stanie zidentyfikować osoby, bo np. musiałby tylko w tym celu stworzyć zbiór danych albo przeszukać wszystkie nieuporządkowane miejsca, gdzie dane osobowe konkretnej osoby mogą się znajdować. W tym przypadku może być pomocne odwołanie się do kryteriów wskazanych w motywie 26 RODO (zob. poniżej),

tj. definicji samych danych osobowych. Skoro ocena, czy dochodzi do przetwarzania danych osobowych, wymaga uwzględnienia m.in. kosztów i czasu potrzebnego do zidentyfikowania oraz aspektu technologicznego (bez rozróżnienia, czy chodzi o identyfikację na podstawie zasobów administratora, czy poprzez pozyskanie danych „z zewnątrz"), to tym bardziej uzasadnione jest uwzględnienie tych czynników przy wykładni art. 11 ust. 2 RODO. W przeciwnym razie należałoby przyjąć, że kryteria stanowiące element definicji danych osobowych są „łagodniejsze" niż kryteria przyjęcia, że osoba jest trudno identyfikowalna (art. 11 RODO). Jak się zdaje, w tym sensie problemy ze zidentyfikowaniem mogą dotyczyć zarówno przetwarzania zautomatyzowanego, jak i niezautomatyzowanego. Jednocześnie trzeba wyraźnie podkreślić, że sam fakt braku uporządkowania (w przypadku zautomatyzowanego przetwarzania) nie musi rodzić trudności w identyfikacji, jeżeli wykorzystywane są określone narzędzia informatyczne. Sumując, naszym zdaniem art. 11 ust. 2 RODO powinien znaleźć zastosowanie w sytuacji problemów z odszukaniem danych, tj. gdy dodatkowa informacja ma umożliwić identyfikację osoby (jak w przedstawionym powyżej przykładzie monitoringu CCTV, teoretycznie w celu odszukania konkretnej osoby na nagraniu można przeglądnąć dziesiątki/setki godzin nagrań, analizując dostępne dla administratora zasoby), jednak uwzględniając czas i koszt takiej analizy, trzeba przyjąć, że byłoby to nadmierne obciążenie administratora danych. Natomiast cały proces będzie możliwy do przeprowadzenia bez nadmiernych kosztów po uzyskaniu dodatkowych informacji umożliwiających identyfikację, np. o czasie, kiedy dana osoba mogła zostać nagrana.

Oczywiście, gdy dane osobowe są przetwarzane w sposób niezautomatyzowany, brak uporządkowania może w ogóle wyłączać zastosowanie RODO, które dotyczy w takiej sytuacji wyłącznie danych będących lub mogących być częścią zbioru danych.

! Co ważne, może być też tak, że podmiot danych dostarcza dodatkowe dane, a administrator danych w dalszym ciągu nie jest w stanie zidentyfikować danych (bo np. do identyfikacji potrzebne byłoby pozyskiwanie danych z innego źródła niż od podmiotu danych). Jak się zdaje, tej sytuacji dotyczy art. 12 ust. 2 zdanie drugie RODO (np. do

identyfikacji potrzebne są dane zebrane z innych niż podmiot danych źródeł). Zgodnie z tym przepisem w przypadkach, o których mowa w art. 11 ust. 2 RODO, administrator nie odmawia podjęcia działań na żądanie osoby, której dane dotyczą, pragnącej wykonać prawa przysługujące jej na mocy art. 15–22 RODO, chyba że wykaże, iż nie jest w stanie zidentyfikować osoby, której dane dotyczą. Inaczej trudno rozpoznać sens tego przepisu w relacji do art. 11 ust. 2 RODO.

Treścią przepisów RODO i definicją danych osobowych objęte są tylko te sytuacje, gdy pozyskanie tych dodatkowych danych – w celu identyfikacji – nie jest nadmiernie uciążliwie (niezależnie od tego, czy chodzi o pozyskanie danych spoza organizacji, czy „odszukanie" wewnątrz). W tym zakresie istotna jest zarówno perspektywa administratora danych, jak i innych podmiotów (np. czy gdyby dane zostały upublicznione, to czy osoba inna niż administrator danych byłaby w stanie powiązać je z konkretną osobą fizyczną).

Precyzuje to treść motywu 26 RODO, zgodnie z którym: „Aby stwierdzić, czy dana osoba fizyczna jest możliwa do zidentyfikowania, należy wziąć pod uwagę wszelkie rozsądnie prawdopodobne sposoby (w tym wyodrębnienie wpisów dotyczących tej samej osoby), w stosunku do których istnieje uzasadnione prawdopodobieństwo, iż zostaną wykorzystane przez administratora lub inną osobę w celu bezpośredniego lub pośredniego zidentyfikowania osoby fizycznej. Aby stwierdzić, czy dany sposób może być z uzasadnionym prawdopodobieństwem wykorzystany do zidentyfikowania danej osoby, należy wziąć pod uwagę wszelkie obiektywne czynniki, takie jak koszt i czas potrzebne do jej zidentyfikowania, oraz uwzględnić technologię dostępną w momencie przetwarzania danych, jak i postęp technologiczny. Zasady ochrony danych nie powinny więc mieć zastosowania do informacji anonimowych, czyli informacji, które nie wiążą się ze zidentyfikowaną lub możliwą do zidentyfikowania osobą fizyczną, ani do danych osobowych zanonimizowanych w taki sposób, że osób, których dane dotyczą, w ogóle nie można zidentyfikować lub już nie można zidentyfikować".

Uwzględniając powyższe, można wyróżnić następujące sytuacje:

1) zidentyfikowanie (czyli przypisanie określonych informacji do wyodrębnionej osoby fizycznej) osoby fizycznej jest możliwe bez dodatkowych informacji (biorąc pod uwagę kryteria z motywu 26 RODO)
   → taka sytuacja jest objęta RODO i nie dotyczy jej art. 11 ust. 2 RODO;
2) zidentyfikowanie osoby fizycznej nie jest możliwe bez dodatkowych informacji, a pozyskanie tych informacji nie jest nadmiernie uciążliwe (biorąc pod uwagę kryteria z motywu 26 RODO)
   → taka sytuacja jest objęta RODO z zastrzeżeniem, że nie aktualizują się niektóre obowiązki, chyba że osoba, której dane dotyczą, w celu wykonania praw dostarczy dodatkowych informacji pozwalających ją zidentyfikować (art. 11 ust. 2 RODO);
   Zestawiając treść art. 11 ust. 2 oraz art. 12 ust. 2 RODO, trzeba przyjąć, że rozporządzeniem może być objęte także przetwarzanie, gdy dostarczone przez podmiot danych dodatkowe informacje w dalszym ciągu nie pozwalają na zidentyfikowanie osoby fizycznej. Można to wytłumaczyć w ten sposób, że cały czas istnieje wtedy możliwość, iż dane pozyskane z innych źródeł przyczynią się do zidentyfikowania osoby fizycznej;
3) zidentyfikowanie osoby fizycznej nie jest możliwe bez dodatkowych informacji, a pozyskanie tych informacji, chociaż możliwe, jest nadmiernie uciążliwe (biorąc pod uwagę kryteria z motywu 26 RODO)
   → w tym przypadku nie dochodzi do przetwarzania danych osobowych (taka sytuacja nie jest co do zasady objęta regulacją RODO);
   Co istotne, w tym przypadku RODO może znaleźć zastosowanie w zakresie oceny ryzyka reidentyfikacji. Często niezwykle trudno przesądzić, czy istnieje możliwość identyfikacji. Należy zadać pytanie, czy wskazana konfiguracja może dotyczyć sytuacji, gdy te dodatkowe informacje mogą być dostarczone przez bezpośredniego zainteresowanego (podmiot danych). Nie można tego wykluczyć *a priori*;
4) zidentyfikowanie osoby fizycznej nie jest możliwe nawet po uzyskaniu dodatkowych informacji (biorąc pod uwagę kryteria z motywu 26 RODO)
   → taka sytuacja nie jest w ogóle objęta regulacją RODO.

Wskazaną dystynkcję dobrze obrazuje rozróżnienie 4 kategorii informacji, które zaproponował M. Hintze, przedstawione w poniższej tabeli.

	Dane dotyczące zidentyfikowanej osoby	**Łatwa identyfikacja**	**Art. 11 (deidentyfikacja)**	**Dane anonimowe**
Bezpośrednio połączone z danymi identyfikującymi	Tak	Nie	Nie	Nie
Znany jest sposób (*systematic way*) na (ponowną) identyfikację	Tak	Tak	Nie	Nie
Dane odnoszące się do konkretnej osoby	Tak	Tak	Tak	Nie

Źródło: M. Hintze, *Viewing the GDPR through a De-Identification Lens: A Tool for Compliance, Clarification, and Consistency*, „International Data Protection Law" 2018/1, s. 1.

Przy takim ujęciu art. 11 RODO znajdzie zastosowanie, gdy nie jest znany sposób na identyfikację. Mike Hintze zdaje się więc przyjmować podejście podobne do zaprezentowanego powyżej w niniejszej publikacji, ponieważ nie uzależnia możliwości zastosowania art. 11 RODO od tego, czy problem z identyfikacją ma charakter „wewnętrzny" (odszukanie danych), czy „zewnętrzny".

W kontekście powyższego należy zwrócić uwagę, że rodzaj źródła tych dodatkowych danych ma znaczenie w tym sensie, że w przypadku gdy osoba fizyczna w celu wykonania praw przysługujących jej na mocy art. 15–20 RODO dostarczy tych dodatkowych informacji pozwalających na jej zidentyfikowanie, administrator danych nie będzie mógł odmówić realizacji tych praw.

Co jednak istotne, art. 11 ust. 2 RODO nie wymaga, aby administrator danych aktywnie poszukiwał tych dodatkowych informacji tylko po to, aby móc zrealizować prawa wskazane w art. 15–20 RODO. Ciekawe, że art. 11 ust. 2 *in principio* RODO wymaga, aby administrator danych w miarę możliwości informował o braku możliwości realizacji praw wskazanych w art. 15–20 RODO w związku z brakiem

dodatkowych informacji umożliwiających identyfikację. Natomiast nie wymaga wprost, aby administrator danych informował o możliwości realizacji tych praw w sytuacji, gdy osoba, której dane dotyczą, dostarczy dodatkowych informacji pozwalających ją zidentyfikować. Otwarta jest kwestia, czy taki obowiązek można wyprowadzić z ogólnego wymogu „ułatwiania" realizacji praw podmiotów danych (art. 12 ust. 1 RODO).

Brak wskazanego obowiązku potwierdza treść art. 5 ust. 2 u.o.d.o., zgodnie z którym w przypadku, gdy wykonanie obowiązków, o których mowa w art. 15 ust. 1 i 3 RODO, wymaga niewspółmiernie dużego wysiłku związanego z wyszukaniem danych osobowych, administrator wykonujący zadanie publiczne wzywa osobę, której dane dotyczą, do udzielenia informacji pozwalających na wyszukanie tych danych. Przepis art. 64 k.p.a. stosuje się odpowiednio.

Przyjęcie odmiennej interpretacji art. 11 ust. 2 RODO (w zakresie braku obowiązku informowania o możliwości realizacji prawa po dostarczeniu dodatkowych informacji) oznaczałoby, że przedstawiony art. 5 ust. 2 u.o.d.o. jest częściowo zbędny (*superfluum*). Co istotne, wskazany powyżej przepis nie modyfikuje zakresu przedmiotowego RODO, czyli dotyczy danych przetwarzanych w sposób zautomatyzowany albo tradycyjny (niezautomatyzowany) tylko wówczas, gdy są one lub mają być częścią zbioru danych.

Należy jeszcze dodać, że możliwość przypisania określonych informacji do wyodrębnionej osoby fizycznej nie przesądza jeszcze, że administrator danych jest w stanie stwierdzić, iż jest to ta, a nie inna osoba. Kwestii tej dotyczy problem weryfikacji tożsamości (zob. art. 12 ust. 6 RODO). W przypadku weryfikacji tożsamości chodzi o ustalenie, czy dane osobowe, które możemy przypisać do wyodrębnionej osoby, dotyczą osoby, która deklaruje, że jest właśnie tą wyodrębnioną osobą (np. czy osoba pisząca z adresu jan.kowalski@___.pl i wnosząca o dostęp do swoich danych osobowych jest Janem Kowalskim zidentyfikowanym za pomocą numeru PESEL w ramach określonego procesu przetwarzania danych u administratora danych). Tutaj także – tj. w przypadku realizacji praw wskazanych w rozdziale III RODO – często wymagane będzie uzyskanie dodatkowych informacji, o czym mówi art. 12 ust. 6 RODO. W tym

przypadku może zachodzić sprzężenie zwrotne: jeżeli administrator danych nie będzie w stanie odpowiedzieć sobie na pytanie, jakich informacji potrzebuje, żeby zweryfikować tożsamość, oznaczać to będzie często, że nie jest w stanie w ogóle przypisać określonych informacji do konkretnej (wyodrębnionej) osoby fizycznej (czyli znajdzie zastosowanie regulacja art. 11 ust. 2 RODO). Może być też tak, że w ramach różnych procesów przetwarzania danych (ujawnionych w rejestrze czynności przetwarzania danych) przetwarzane są dane osobowe tej samej osoby, a jednocześnie administrator danych – ze względu na różne zakresy przetwarzanych danych i brak danych wspólnych – będzie musiał wprowadzić różne mechanizmy weryfikacji tożsamości.

Na zakończenie należy zwrócić uwagę, że RODO nie znajduje zastosowania do przetwarzania danych anonimowych, tj. tych, które nie dotyczą zidentyfikowanej lub możliwej do identyfikacji osoby fizycznej przy uwzględnieniu kryteriów z motywu 26 RODO. Jak wskazano powyżej, RODO znajdzie jednak zastosowanie do oceny ryzyka reidentyfikacji.

Na potrzebę oceny ryzyka w kontekście reidentyfikacji wskazywała już Grupa Robocza Art. 29 w opinii 6/2013[12], przedstawiając następujące rekomendacje co do szacowania ryzyka reidentyfikacji:

- szacowanie ryzyka reidentyfikacji powinno uwzględniać zarówno kategorie (anonimowych danych) poddane ponownemu wykorzystaniu, jak i inne dostępne informacje (np. dostępne poprzez wyszukiwarki internetowe)[13];

[12] Grupa Robocza Art. 29, *Opinia 6/2013 w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego (ISP)*, przyjęta 5.06.2013 r., WP 207 1021/00/PL, https://archiwum.giodo.gov.pl/pl/file/6380 (dostęp: 15.02.2022 r.).

[13] Grupa Robocza Art. 29, *Opinia nr 6/2013…*, pkt 6.4. Jest to zgodne z motywem 26 RODO, który wskazuje, że: „Aby stwierdzić, czy dana osoba fizyczna jest możliwa do zidentyfikowania, należy wziąć pod uwagę wszelkie rozsądnie prawdopodobne sposoby (w tym wyodrębnienie wpisów dotyczących tej samej osoby), w stosunku do których istnieje uzasadnione prawdopodobieństwo, iż zostaną wykorzystane przez administratora lub inną osobę w celu bezpośredniego lub pośredniego zidentyfikowania osoby fizycznej. Aby stwierdzić, czy dany sposób może być z uzasadnionym prawdopodobieństwem wykorzystany do zidentyfikowania danej osoby, należy wziąć pod uwagę wszelkie obiektywne czynniki, takie jak koszt i czas potrzebne do jej zidentyfikowania, oraz uwzględnić technologię dostępną w momencie przetwarzania danych, jak i postęp technologiczny".

- należy ocenić, czy dane, które mają zostać opublikowane, mogą być powiązane z innymi zestawami danych;
- należy uwzględnić, że ryzyko reidentyfikacji wzrasta, gdy określona grupa osób (np. rodzina, koledzy ze szkoły) dysponuje już dużą ilością informacji odnośnie do innych osób;

Jak wskazuje Grupa Robocza Art. 29, w tym przypadku liczy się nie tylko to, czy osoba posiadająca wcześniejszą wiedzę może zidentyfikować osobę, której dane dotyczą, ale czy nauczy się czegoś nowego na podstawie informacji uzyskanych w wyniku ponownej identyfikacji. Dwa poniższe przykłady, podane przez Grupę Roboczą Art. 29, ilustrują znaczenie tego rozróżnienia.

Przykład pierwszy: statystyki dotyczące odry. W jednym przypadku zanonimizowane dane statystyczne mogą ujawnić, że w mieście A w 2012 r. X liczba osób zachorowała na odrę. Nie podano dalszego podziału ani informacji. Lekarz, który wniósł wkład do statystyk, przekazując odpowiedniemu urzędowi zdrowia informacje o swoich pacjentach, przechowuje w swoim biurze pełniejsze zapisy dotyczące tych pacjentów, z zastrzeżeniem tajemnicy lekarskiej. Lekarz byłby w stanie łatwo ponownie zidentyfikować kilku pacjentów ze statystycznego zestawu danych. Podobnie matka, która wiedziała, że jej dziecko zachorowało na odrę w tym roku, może łatwo ponownie zidentyfikować swoje dziecko w zbiorze danych. Niemniej jednak ani matka, ani lekarz nie dowiedzieliby się niczego, czego wcześniej nie znali, z anonimowego zbioru danych, który został udostępniony publicznie.

Przykład drugi: nadużywanie środków odurzających i alkoholu, niegodziwe traktowanie w celach seksualnych a wyniki w nauce osiągane w szkole. Przykład ten można skontrastować 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.

Statystki pokazują między innymi, że w szkole A, do której chodzi ogółem 500 uczniów, w 2012 r. 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 żadni inni uczniowie w szkole A nie byli niegodziwie traktowani 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 alkoholiczką. 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.