

AGNIESZKA PISKORZ-RYŃ
Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie
ORCID 0000-0001-9788-0988

OTWIERANIE DANYCH BADAWCZYCH Z PERSPEKTYWY SZKOŁY WYŻSZEJ

Treść: Wstęp. – 1. Pojęcie danych badawczych. – 2. Zasady ponownego wykorzystywania danych badawczych w dyrektywie. – 3. Zasady ponownego wykorzystywania danych badawczych w ustawie o otwartych danych. – 4. Wyzwania dla uczelni wyższych. Próba podsumowania.

Wstęp

Prawo do ponownego wykorzystywania jest jednym z praw informacyjnych nowej generacji. Jego geneza związana jest z rozwojem nowych technologii oraz z ruchem społecznym, jakim jest otwarty rząd¹. To właśnie w jego ramach sformułowane zostały postulaty dzielenia się przez władzę publiczną danymi, tak by mogły być one wykorzystywane przez innych z pożytkiem dla społeczeństwa. Dzięki nowym technologiom łatwiej jest bowiem osiągnąć takie cele społeczne jak wzrost partycypacji społecznej, demokracji i współpracy. Postulat dzielenia się wiązać należy również z innym ruchem, jakim jest otwarta nauka². Ten zaś zakłada możliwości dostępności

¹ Na temat idei otwartego rządu i jego znaczenia dla prawa do ponownego wykorzystywania informacji sektora publicznego zob. A. PISKORZ-RYŃ, *Ponowne wykorzystanie informacji sektora publicznego. Zagadnienia administracyjnoprawne*, Warszawa 2018, s. 59-79.

² Zob. B. FECHER, S. FRIESIKE, *Open Science: One Term, Five Schools of Thought*, SSRN Electronic Journal Conference paper, April 2013, s. 44. Na ten temat również E. KULCZYŃSKI, *Otwarta nauka a komunikacja – perspektywa metateoretyczna*,

publikacji i danych badawczych dla innych naukowców. Dla tego artykułu właśnie te ostatnie mają kluczowe znaczenie.

O ponownym wykorzystywaniu danych badawczych możemy mówić dzięki rozszerzeniu zakresu ponownego wykorzystywania na poziomie europejskim³. W polskim porządku prawnym postanowienia dyrektywy implementuje ustawa z 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego (dalej: u.o.d.i.p.w.)⁴. To w tym właśnie akcie prawnym, choć w ograniczonym zakresie, zagwarantowano po raz pierwszy możliwość ponownego wykorzystywania danych badawczych.

W artykule tym przedstawiona zostanie regulacja prawna dotycząca ponownego wykorzystywania danych badawczych na poziomie unijnym i krajowym, tak by na tym tle wskazać jakie wyzwania stoją w tym zakresie przez szkołami wyższymi. Jak zostanie w tekście przedstawione ustalenie charakteru i zakresu wyzwań będzie o tyle trudne, że pewne problemy związane z praktyczną realizacją tego prawa wynikają z jednej strony z deficytów regulacji europejskiej oraz implementacji i transpozycji dyrektywy w prawie krajowym. Stąd w artykule tym zostaną one opisane by pokazać jakich zmian należy oczekiwać dla realizacji ponownego wykorzystywania danych badawczych od prawodawcy krajowego, a jakiego zakresu aktywności powinniśmy domagać się od szkół wyższych.

1. Pojęcie danych badawczych

Prawo do ponownego wykorzystywania informacji sektora publicznego przysługuje każdemu⁵. Zakres przedmiotowy prawa do ponownego wykorzystywania określa pojęcie informacji sektora publicznego

w: Komunikacja naukowa w humanistyce, Poznań 2017, <http://hdl.handle.net/10593/17589>, s. 67-85.

³ Zostało ono dokonane w dyrektywie Parlamentu Europejskiego i Rady (UE) 2019/1024 z 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego (dalej: „dyrektywa”) Dz. U. UE. L. z 2019 r. Nr 172, s. 56.

⁴ Dz. U. poz. 1641 z późn. zm.

⁵ Art. 5 ust. 1 u.o.d.i.p.w.

(dalej: i.s.p.). Zgodnie z definicją legalną jest to „każda treść lub jej część, niezależnie od sposobu utrwalenia, w szczególności w postaci papierowej, elektronicznej, dźwiękowej, wizualnej lub audiowizualnej, będącą w posiadaniu podmiotu zobowiązanego”⁶. Jedną z kategorii danych wyodrębnionych w ramach i.s.p. są dane badawcze. Są one definiowane na poziomie dyrektywy i w ustawie o otwartych danych.

Zgodnie z dyrektywą są to dokumenty w formie cyfrowej, inne niż publikacje naukowe, które są gromadzone lub produkowane w ramach działalności badawczo-naukowej i są wykorzystywane jako dowody w procesie badawczym bądź też są powszechnie akceptowane w środowisku naukowym jako konieczne do weryfikacji poprawności ustaleń i wyników badań⁷. W motywie 26 dyrektywy wskazano przykłady tych danych, do których zaliczono dane statystyczne, wyniki eksperymentów, pomiarów, obserwacji prowadzonych w terenie, wyniki ankiet, nagrania wywiadów i zdjęcia. Obejmują one także metadane, specyfikacje i inne obiekty cyfrowe. Przed rozpoczęciem analizy tego pojęcia podkreślenia wymaga jakościowa zmiana. Pojęcie to zostało zdefiniowane w przepisie prawa. Dotychczas jego definicje formułowali sami zainteresowani-badacze⁸. Definicje te były mniej precyzyjne, ale co do istoty nie odbiegały od tego, o czym mowa

⁶ Art. 2 pkt 8 u.o.d.i.p.w.

⁷ Art. 2 pkt 9 dyrektywy.

⁸ Zwrócić należy uwagę na pogląd wyrażony przez N. Dietricha i A. Wiebe, którzy stwierdzili, że termin „dane badawcze” nie wydaje się zbyt pomocny, gdyż nie ma ogólnie przyjętej definicji takich danych, albowiem „każdy z autorów i każde z badań naukowych przyjmuje własną definicję tego pojęcia” (N. DIETRICH, A. WIEBE, *Definicja danych badawczych*, w: *Otwarte, czyli bezpieczne. Analiza prawnych form ochrony danych badawczych i zalecenia w kwestii dostępu i korzystania*, red. L. Guibault, A. Wiebe, Kraków 2014, s. 16. Autorzy ci uznali, że termin „dane badawcze” nie zostanie ściśle zdefiniowany, lecz będzie obejmował każdy rodzaj danych stworzonych w toku badań naukowych, takich jak bazy surowych danych, tabele, grafiki, obrazki i wszystko inne (tamże, s. 16).

w dyrektywie⁹. Opierały się one na pojęciu danych surowych¹⁰. Pierwszy element definicji zawartej w dyrektywie dotyczy zapisu danych. Zgodnie z nim dane badawcze to dokumenty w postaci cyfrowej. W ten sposób prawodawca wiąże dane badawcze z zdefiniowanym w dyrektywie pojęciem dokumentu ograniczając je do dokumentów w postaci cyfrowej¹¹. Danymi badawczymi będą więc wszelkie treści lub ich części niezależnie od formy ich wyrażenia zapisane w postaci cyfrowej. Z tego zakresu prawodawca europejski wyłącza wprost publikacje naukowe. Dla uznania danych za dane badawcze, ważny jest nie tylko sposób ich zapisu, ale także proces ich powstania. Do tej kategorii zaliczamy każdą treść lub jej część, która spełnia łącznie dwa warunki. Jest gromadzona lub produkowana w ramach działalności badawczo-naukowej oraz jest wykorzystywana jako dowód w procesie badawczym bądź też jest powszechnie akceptowane w środowisku naukowym jako konieczne do weryfikacji poprawności ustaleń i wyników badań. W przypadku drugiego z warunków może wystąpić tylko jeden z przypadków w nim określonych. Zwrócić należy uwagę, że danymi badawczymi zgodnie z motywem 26 mogą być również metadane. W ustawie o otwartych danych wskazano, że metadane to ustrukturyzowane informacje opisujące, tłumaczące, lokalizujące i ułatwiające odnalezienie, wykorzystanie innych informacji

⁹ W literaturze przedmiotu wskazywano, że do tej kategorii danych należą dane obserwacyjne, eksperymentalne, symulacyjne, zależne (kompilowane), referencyjne (kanoniczne) (zob. B. FISCHER, A. PISKORZ-RYŃ, M. SAKOWSKA-BARYŁA, J. WYPORSKA-FRANKIEWICZ, *Ustawa o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego. Komentarz*, Warszawa 2022, art. 2).

¹⁰ Z ich trzon uznawano dane surowe, a więc nieprzetworzone, które mogą być uznawane za źródłowe dla jakichś metod, dane będące wynikiem przetworzenia danych źródłowych przez konkretne metody można by zaś wówczas określać jako przetworzone (zob. B. FISCHER, *Autorskoprawne konteksty ponownego wykorzystywania danych badawczych*, w: *Sto lat polskiego prawa handlowego. Księga jubileuszowa dedykowana Profesorowi Andrzejowi Kidybie*, t. 2, red. M. Dumkiewicz, K. Kopaczyńska-Pieczniak, J. Szczotka, Warszawa 2020, s. 561).

¹¹ Dokument oznacza dowolną treść niezależnie od jej nośnika (papier lub forma elektroniczna lub zapis dźwiękowy, wizualny bądź audiowizualny); lub dowolną część tej treści (art. 2 pkt 6 dyrektywy).

lub zarządzanie nimi¹². Niezależnie od sposobu zdefiniowania służą opisowi treści zapisanej w postaci cyfrowej. Metadanymi mogą być nimi opisane zarówno informacje, jak i dokumenty tekstowe, zdjęcia, nagrania, rysunki oraz mapy. Ich celem jest zapewnienie możliwości wyszukania informacji lub innych obiektów zapisanych w formie cyfrowej, a także ich kontrolę, zrozumienie i długotrwałe przechowanie oraz zarządzanie. Pozwalają one również rozpoznawanie i czytanie danych przez maszyny, co jest bardzo istotne dla ponownego wykorzystywania i.s.p. W związku z tym przyjąć należy, że metadanymi będą informacje opisujące wszelkie treści lub ich części zapisane w formie cyfrowej służące ich wyszukiwaniu, kontroli, zrozumieniu i długotrwałemu przechowaniu oraz zarządzaniu. Dzięki nim możliwe jest również zautomatyzowane łączenie danych¹³. Ustawodawca krajowy pojęcie danych badawczych definiuje zgodnie z postanowieniami dyrektywy dostosowując ją do warunków krajowych poprzez odwołanie do art. 4 Prawo o szkolnictwie wyższym i nauce¹⁴. Przepis ten określa zakres pojęcia działalność naukowa¹⁵. W definicji legalnej

¹² Art. 2 pkt 10 ustawy z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego (Dz. U. poz. 1641 z późn. zm.), zwanej dalej u.o.d.

¹³ A. PISKORZ-RYŃ, *Ponowne wykorzystywanie danych badawczych*, w: *Prawo sztucznej inteligencji i nowych technologii*, red. B. Fischer, A. Pązik, M. Świerczyński, Warszawa 2024, s. 147-148.

¹⁴ Ustawa z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (t.j. Dz. U. z 2023 r. poz. 742 z późn. zm.).

¹⁵ Zgodnie z tym przepisem działalność naukowa obejmuje badania naukowe, prace rozwojowe oraz twórczość artystyczną (ust. 1). Badania naukowe są działalnością obejmującą: badania podstawowe rozumiane jako prace empiryczne lub teoretyczne mające przede wszystkim na celu zdobywanie nowej wiedzy o podstawach zjawisk i obserwowalnych faktów bez nastawienia na bezpośrednie zastosowanie komercyjne oraz badania aplikacyjne rozumiane jako prace mające na celu zdobycie nowej wiedzy oraz umiejętności, nastawione na opracowywanie nowych produktów, procesów lub usług lub wprowadzanie do nich znaczących ulepszeń (ust. 2). Prace rozwojowe są działalnością obejmującą nabywanie, łączenie, kształtowanie i wykorzystywanie dostępnej aktualnie wiedzy i umiejętności, w tym w zakresie narzędzi informatycznych lub oprogramowania, do planowania produkcji oraz projektowania i tworzenia zmienionych, ulepszonych lub nowych

danych badawczych wskazano, że dane badawcze są to informacje sektora publicznego utrwalone w postaci elektronicznej, inne niż publikacje naukowe, które zostały wytworzone lub zgromadzone w ramach działalności naukowej w rozumieniu ww. przepisu i są wykorzystywane jako dowody w procesie badawczym lub służą do weryfikacji poprawności ustaleń i wyników badań¹⁶.

2. Zasady ponownego wykorzystywania danych badawczych w dyrektywie

Dyrektywa problematykę danych badawczych podejmuje w art. 10. W przepisie tym wskazuje, że państwa członkowskie wspierają dostępność danych badawczych poprzez przyjęcie polityki krajowej oraz podejmowanie odpowiednich działań mających na celu zapewnienie powszechnej dostępności danych badawczych finansowanych ze środków publicznych („polityka otwartego dostępu”), zgodnie z zasadą „otwartości domyślnej” i zgodnie z zasadami FAIR. W tym kontekście należycie uwzględnia się kwestie związane z prawami własności intelektualnej, ochroną danych osobowych i poufnością, bezpieczeństwem i uzasadnionymi interesami handlowymi zgodnie z zasadą „otwarty jak to najbardziej możliwe, zamknięty jak to konieczne”. Wspomnianą politykę otwartego dostępu kieruje się do organizacji prowadzących badania naukowe i organizacji finansujących badania naukowe¹⁷. Przepis ten wskazuje granice ponownego wykorzystywania danych badawczych. Nakłada on na państwa członkowskie obowiązek wspierania dostępności danych badawczych poprzez przyjęcie polityki krajowej oraz podjęcie odpowiednich działań mających na celu zapewnienie powszechnej dostępności danych badawczych finansowanych ze środków publicznych. W ten sposób wskazano zakres

produktów, procesów lub usług, z wyłączeniem działalności obejmującej rutynowe i okresowe zmiany wprowadzane do nich, nawet jeżeli takie zmiany mają charakter ulepszeń (ust. 3). Twórczość artystyczna jest działalnością obejmującą działania twórcze w sztuce, której efektem jest stanowiące wkład w rozwój kultury dzieło artystyczne materialne lub niematerialne, w tym także artystyczne wykonanie (ust. 4).

¹⁶ Art. 2 pkt 2 u.o.d.i.p.w.

¹⁷ Art. 10 ust. 1 dyrektywy.

ponownego wykorzystywania danych badawczych i ograniczono go do danych finansowanych ze środków publicznych. Bez precyzowania, co dyrektywa pod pojęciem tym rozumie. W przepisie tym prawodawca europejski nakłada na ustawodawcę krajowego obowiązek podjęcia określonego zachowania. Jest nim przyjęcie polityki krajowej, ale również podjęcie innych odpowiednich działań wskazując przy tym rezultat, który przez ustawodawcę krajowego ma być wypracowany w procesie implementacji. Jest nim osiągnięcie określonego stanu rzeczy – zapewnienia powszechnej dostępności danych badawczych finansowanych ze środków publicznych („polityka otwartego dostępu”). Nie jest to cel generalny dyrektywy, ale jej cel cząstkowy wynikający z cytowanej normy. Dyrektywa jako akt dwustopniowy pozostawia państwom członkowskim margines wyboru form i środków dla realizacji wskazanego w niej celu¹⁸. W tym przypadku zakres tej swobody jest szczególnie duży. Dyrektywa wskazuje bowiem, że państwa członkowskie powinny podjąć inne odpowiednie działania. Wybór w tym zakresie pozostawiono ustawodawcy krajowemu. Bezpośrednimi adresatami dyrektyw są państwa członkowskie, a sama dyrektywa wiąże jedynie co do „rezultatu, który ma być osiągnięty”, pozostawiając organom krajowym „swobodę wyboru form i środków” służących realizacji celów dyrektywy¹⁹.

Jednym z mechanizmów wspierania danych badawczych jest przyjęcie polityki krajowej. Dyrektywa tego pojęcia nie precyzuje. Wydaje się, że powinno być ono rozumiane w sposób zgodny z ideą triady nauk administracyjnych²⁰, odwołując się do koncepcji nauki polityki administracyjnej. Jej „przedmiotem jest kwestia optymalizacji skuteczności działalności publicznej, a więc zewnętrzna sprawność działalności administracyjnej państwa. Wartościowanie zjawisk społecznych (w tym administracyjnych) następuje tu zatem

¹⁸ B. KURCZ, *Dyrektywy Wspólnoty Europejskiej i ich implementacja do prawa krajowego*, Kraków 2004, rozdz. 1.1. pkt 3-5.

¹⁹ M. KAMIŃSKI, *Bezpośrednie i pośrednie stosowanie dyrektyw unijnych przez polskie sądy administracyjne*, PS 1 (2011), s. 24.

²⁰ F. LONGCHAMPS, *Założenia nauki administracji*, Wrocław 1991, s. 35-39.

nie ze względu na samą administrację, lecz ze względu na skutki, jakie jej działalność rodzi w otoczeniu²¹. W tym kontekście politykę rozumieć należy jako działanie oparte na wiedzy, polegające na opisie stanu spraw publicznych w wyodrębnionej dziedzinie, ich analizie oraz ocenie według racjonalnych i wewnętrznie spójnych kryteriów, a następnie sformułowanie programu, który wskazuje cele, sposoby i środki do ich osiągnięcia, pożądane z uwagi na dokonaną ocenę spraw publicznych²².

Poza wskazaniem rezultatu, jakim jest zapewnienie powszechnej dostępności danych badawczych finansowanych ze środków publicznych, prawodawca europejski określa podstawy na jakich powszechna dostępność ma się opierać. Są to zasady „otwartości domyślnej” i zasady FAIR. Otwartość domyślna (*open by default*) nie jest zdefiniowana na potrzeby dyrektywy. Wydaje się uzasadnione odwołanie w tym zakresie do Międzynarodowej Karty Otwartych Danych²³, która stanowi zbiór zasad i dobrych praktyk w zakresie otwarcia danych publicznych. Przez „otwartość domyślną” na jej gruncie rozumieć należy działanie władzy publicznej w stosunku do obywateli i innych podmiotów, zakładające domniemanie dostępności danych. Nie oznacza to, jednakże, iż wszystkie dane są dostępne. Może bowiem ograniczyć ich dostępność w oparciu o uzasadnione przyczyny. Sama dyrektywa wskazuje na ich katalog. Dane powinny być o ile to możliwe otwarte, ale zamknięte w tym zakresie w jakim to konieczne. Karta odnosi się również do budowania otwartości domyślnej wskazując, że następuje ono nie tylko poprzez decyzje polityczne, rozwiązania prawne, ale również za pomocą programów szkoleniowych i uświadamiających, narzędzi, wytycznych

²¹ H. IZDEBSKI, M. KULESZA, *Administracja publiczna, zagadnienia ogólne*, Warszawa 2004, s. 12.

²² J. JEŻEWSKI, *Polityka administracyjna – przedmiot i metoda*, w: *Polityka administracyjna*, IV Międzynarodowa Konferencja Naukowa Stryków 7-9 września 2008 r., red. J. Łukasiewicz, Rzeszów 2008, s. 299.

²³ International Open Data Charter, Wrzesień 2015 r., https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf [dostęp 7.03.2024].

oraz strategii komunikacyjnych²⁴. Państwa członkowskie w ramach podejmowania innych działań powinny dążyć do stworzenia kultury otwartości i zachęcać do proaktywnych działań związanych z sposobem projektowania i prowadzenia badań. Powstanie zasad FAIR wiązać należy z ruchem otwartej nauki²⁵. Powstały one jako oddolna inicjatywa środowisk akademickich, biznesu, agencji finansujących badania i wydawnictw naukowych jako wytyczne dla poprawy infrastruktury wspierającej ponowne wykorzystywanie danych naukowych²⁶. Dopiero później zostały uznane przez UE²⁷. W ich nazwie wykorzystano akronim ułatwiający opisywanie propagowanej koncepcji, który sam w sobie niesie pozytywne przesłanie. Przymiotnik FAIR oznacza uczciwy, sprawiedliwy, ale też równy i niedyskryminacyjny²⁸. Akronim powstał od słów *Findable* (możliwe do odnalezienia), *Accessible* (dostępne), *Interoperable* (interoperacyjne) oraz *Reusable* (możliwe do ponownego użycia). Zgodnie z zasadami FAIR wszystkie obiekty badawcze powinny być możliwe do

²⁴ Zob. zasada 1 pkt 13-17 Międzynarodowej Karty Otwartych Danych oraz podsumowanie zasady pierwszej (<https://opendatacharter.net/principles/#open-by-default>).

²⁵ Dokonany w artykule tym opis zasad FAIR opiera się na informacjach zawartych w tej publikacji oraz na stronie Go FAIR Inciative- Zasady FAIR <https://www.go-fair.org/fair-principles>.

²⁶ M.D. WILKINSON i in., The FAIR Guiding Principles for scientific data management and stewardship, *Sci. Data* 3:160018 doi 10.1038/sdata.2016.18 (2016), <https://www.nature.com/articles/sdata201618#citeas>.

²⁷ Na temat podmiotów propagujących zasady FAIR i prowadzące badania w tym zakresie zob. *Cost-Benefit analysis for FAIR research data – Cost of not having FAIR research data*, European Commission, marzec 2018, doi 10.2777/02999, s. 6. Na temat zasad FAIR w działaniu UE zob. Final Report and Action Plan from the European Commission Expert Group on FAIR Data European Commission Expert Group on FAIR Data, 2018 November, doi 10.2777/1524. Unia przyjęła również zasady FAIR w ramach program Horizont 2020, *Guidelines on Data Management in Horizon 2020*, s. 6; http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf; *Executive Board of the European Open Science Cloud (EOSC) – Strategic Implementation Plan*, 2019, s. 10-14.

²⁸ Na temat tego, czy zasady FAIR są „fair”, zob. B. MONS i in., *Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud*, s. 5; *Information Services and Use* 37 (2017), s. 49-5; doi 10.3233/ISU-170824.

znalezienia, dostępne, interoperacyjne i wielokrotnego użytku (FAIR) zarówno dla maszyn, jak i dla ludzi, co ma szczególne znaczenie wraz z rozwojem sztucznej inteligencji. Maszyny stanowią bowiem istotne wsparcie dla człowieka w związku z ilością gromadzonych danych badawczych. Zgodnie z zasadami FAIR pierwszym krokiem do wykorzystania danych jest ich odnalezienie. Powinny być więc możliwe do odnalezienia, zarówno dla ludzi jak i komputerów. Dotyczy to zarówno danych jak i opisujących je metadanych. Metadane nadające się do odczytu maszynowego są niezbędne do automatycznego wykrywania zbiorów danych i usług, a zatem jest to konieczny element procesu publikacji zgodnie z zasadami FAIR²⁹. Dla możliwości odnalezienia (meta)dane³⁰ muszą mieć po pierwsze, przypisany globalnie niepowtarzalny i trwały identyfikator; po drugie, być opisane za pomocą rozbudowanych metadanych, po trzecie metadane muszą zawierać identyfikator danych, które opisują; po czwarte (meta)dane muszą być rejestrowane lub indeksowane w zasobach możliwych do przeszukiwania. Dane zgodnie z zasadą FAIR winny być dostępne, przez co rozumieć należy, że użytkownik po ich znalezieniu musi wiedzieć, w jaki sposób można uzyskać do nich dostęp, a jeżeli to konieczne uwierzytelnienie i autoryzację. Dla realizacji wskazanej zasady (meta)dane muszą być pozyskiwane za pomocą identyfikatora przy użyciu standardowego protokołu komunikacyjnego. Musi być on otwarty, bezpłatny i uniwersalnie wykonalny. Musi pozwalać w razie potrzeby na przeprowadzenie procedury uwierzytelniania i autoryzacji. Metadane zaś muszą być dostępne nawet wtedy, gdy opisywane przez nie dane nie mogą być już pozyskane. Kolejną zasadą FAIR jest interoperacyjność, przez co rozumieć należy, że dane zazwyczaj powinny być zintegrowane z innymi danymi. Muszą współdziałać

²⁹ Dokonując opisu, zdecydowano się wskazać poszczególne zasady FAIR odnoszące się do czterech filarów FAIR. Każda z tych zasad jest dalej opisywana. W tym miejscu ze względu na ograniczone ramy tego opracowania zrezygnowano z uwzględnienia tego opisu.

³⁰ Przyjęty sposób zapisu oznacza, iż wskazane treści odnoszą się zarówno do danych, jak i metadanych.

z aplikacjami lub poprogramowaniami do pracy grupowej służącymi do celów analizy, przechowywania oraz przetwarzania danych. (Meta)dane powinny używać po pierwsze, formalnego, dostępnego, współdzielonego i szeroko stosowanego języka do przedstawiania wiedzy; po drugie, wykorzystywać słowniki zgodne z zasadami FAIR; po trzecie obejmować kwalifikowane odniesienia do innych (meta) danych. Ostatecznym celem FAIR jest optymalizacja ponownego wykorzystania danych. Aby to osiągnąć, metadane i dane powinny być dobrze opisane za pomocą mnogości dokładnych i istotnych atrybutów, tak aby mogły być replikowane i/lub łączone w różnych ustawieniach. Zasada ta koncentruje się na zdolności użytkownika (człowieka lub maszyny) na decydowaniu o przydatności danych w określonych celach. Z tego względu wydawca danych powinien dostarczyć nie tylko metadane pozwalające wykrycie danych, ale opisujące kontekst w jakim zostały wygenerowane może to obejmować np. protokoły eksperymentalne, producenta i markę maszyny lub czujnika, który utworzył dane itp. W związku z tym meta(dane) powinny być udostępniane z jasną i dostępną licencją na użytkowanie danych, jak również powinny mieć udokumentowane pochodzenie. Inne osoby korzystające z danych powinny wiedzieć skąd pochodzą, kogo zacytować i/lub w jaki sposób autor chce zostać uznany. Z tego względu należy załączyć opis przepływu pracy, który doprowadził do powstania danych tak by było wiadomo kto je wygenerował lub zebrał, ale również w jaki sposób zostały one przetworzone, czy były one już wcześniej publikowane? Czy zawierają dane od kogoś innego, które mogły zostać przekształcone lub uzupełnione? Idealnie byłoby, gdyby ten przepływ pracy był opisany w formacie nadającym się do odczytu maszynowego. Ponadto (meta)dane spełniają wspólne normy dotyczące danej dziedziny. Łatwiej jest ponownie wykorzystywać dane, jeśli są one ustrukturyzowane w znormalizowany sposób. Istnieją wspólne standardy dotyczące archiwizacji i udostępniania danych, przykładem może być MIAME, czy MIAPLE.

Uzupełnieniem określenia zasad udostępniania danych badawczych do ponownego wykorzystywania są postanowienia art. 10 ust. 2 dyrektywy. Zgodnie z nimi bez uszczerbku dla art. 1 ust. 2 lit. c) dane

badawcze udostępnia się do ponownego wykorzystywania w celach komercyjnych lub niekomercyjnych zgodnie z rozdziałami III i IV, w zakresie, w jakim dane te są finansowane ze środków publicznych i zostały już publicznie udostępnione przez naukowców, organizacje prowadzące badania naukowe lub organizacje finansujące badania naukowe za pośrednictwem repozytorium instytucjonalnego lub tematycznego. W tym kontekście uwzględnia się uzasadnione interesy handlowe, działania związane z transferem wiedzy oraz istniejące już prawa własności intelektualnej. Podkreślenia wymaga to, że dane badawcze mogą być wykorzystywane w celach komercyjnych i niekomercyjnych. Jednocześnie powtórzono, że warunkiem jest finansowanie ze środków publicznych. Przepis ten jednak ogranicza ponowne wykorzystywanie danych badawczych formułując kolejny warunek. Ponowne wykorzystywanie odnosi się wyłącznie do danych badawczych publicznie udostępnione przez naukowców, organizacje prowadzące badania naukowe lub organizacje finansujące badania naukowe za pośrednictwem repozytorium instytucjonalnego lub tematycznego³¹. W związku z powyższym dane badawcze podlegają ponownemu wykorzystywaniu w celach komercyjnych i niekomercyjnych, jeżeli po pierwsze nie narusza to prawa własności należących do osób trzecich (art. 10 ust. 2 w zw. z art. 1 ust. 2 lit. c), po drugie dane te są finansowane ze środków publicznych, po trzecie, zostały

³¹ Tłumaczy to motyw 28 zd. 11–13, w którym wskazano, żeby „uniknąć obciążenia administracyjnego, wynikające z niniejszej dyrektywy obowiązki powinny mieć zastosowanie wyłącznie do tych danych badawczych, które zostały już publicznie udostępnione przez naukowców, organizacje prowadzące badania naukowe lub organizacje finansujące badania naukowe za pośrednictwem repozytorium instytucjonalnego lub tematycznego, i nie powinny pociągać za sobą dodatkowych kosztów związanych z pobieraniem zbiorów danych ani wymagać dodatkowej konserwacji danych. Państwa członkowskie mogą rozszerzyć stosowanie niniejszej dyrektywy na dane badawcze publicznie udostępnione za pośrednictwem innych infrastruktur danych niż repozytoria, za pośrednictwem publikacji w otwartym dostępie, w postaci pliku dołączonego do artykułu, artykułu udostępniającego nieprzetworzone dane (*data paper*) lub publikacji w czasopiśmie naukowym udostępniającym wysokiej jakości zbiory danych (*data journal*). Dokumenty inne niż dane badawcze powinny nadal być wyłączone z zakresu niniejszej dyrektywy”.

udostępnione przez naukowców, organizacje prowadzące badania naukowe lub organizacje finansujące badania naukowe za pośrednictwem repozytorium instytucjonalnego lub tematycznego.

3. Zasady ponownego wykorzystywania danych badawczych w ustawie o otwartych danych

Kwestie danych badawczych w ustawie reguluje art. 22 ustawy o otwartych danych. Zgodnie z tym przepisem dane badawcze, będące w posiadaniu podmiotów, o których mowa w art. 7 ust. 1 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce, podlegają bezpłatnie ponownemu wykorzystywaniu, jeżeli zostały wytworzone lub zgromadzone w ramach działalności naukowej finansowanej ze środków publicznych oraz są już publicznie udostępniane w systemie teleinformatycznym podmiotu zobowiązanego, w szczególności w repozytorium instytucjonalnym lub tematycznym. Podmiot zobowiązany, udostępniając te dane, wraz z ich udostępnieniem informuje o braku warunków ponownego wykorzystywania albo określa te warunki. Postanowienia te uzupełnia art. 23 ustawy. Nakłada on na Ministra właściwego do spraw szkolnictwa wyższego i nauki obowiązek opracowania Polityki otwartego dostępu do danych badawczych finansowanych ze środków publicznych, uwzględniając w szczególności: 1) aktualny stan udostępniania danych badawczych oraz standardy i dobre praktyki w zakresie udostępniania danych badawczych; 2) obowiązujący stan prawny; 3) kierunki rozwoju w zakresie otwartego dostępu do danych badawczych i zarządzania wynikami badań naukowych; 4) Program otwierania danych, o którym mowa w art. 28³². Politykę otwartego dostępu do danych badawczych finansowanych ze środków publicznych kieruje się do podmiotów, o których mowa w art. 7 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce. Rozwiązania przyjęte w ustawie o otwartych danych są zgodne z dyrektywą. Brak jednak w tym akcie prawnym rozwiązań o charakterze zasadniczym. Powoduje to, że implementację dyrektywy do prawa krajowego należy ocenić negatywnie. Brak zaś innych działań ma wpływ na transpozycję dyrektywy. Najpierw sformułować

³² Art. 23 ust. 1 u.o.d.

należy uwagi dotyczące implementacji. Po pierwsze zgodnie z postanowieniami dyrektywy polityka krajowa powinna opierać się o zasadę otwartości domyślnej i zasady FAIR. Ustawodawca kwestię tę pomija w treści art. 23 u.o.d.p.w. Dla realizacji rezultatu określonego w dyrektywie, polityka krajowa powinna być związana z takim proaktywnym działaniem, który wiąże badacza już na etapie projektowania badań tak aby dane były tworzone jako FAIR. Zasady te poprzez ich uwzględnienie w przepisach prawa mogłyby kształtować obowiązek badacza, ustanowiony w celu realizacji dobra wspólnego, tak by planować badania tak by dane powstałe w ich wyniku były wielokrotnego użytku służąc innym badaczom, rozwojowi nauki i społeczeństwu. Kolejnym krokiem w polityce otwartego dostępu powinno być zobowiązanie podmiotów finansujących badania ze środków publicznych, aby uzależniały ich finansowanie od stosowania w procesie badawczym zasad FAIR. Chodzi o to by tworzone dane odpowiadały określonym standardom będąc użytecznie nie tylko dla naukowców, którzy je tworzą na potrzeby swoich badań, ale również dla innych badaczy. Ci bowiem mogą ich używać w innym celu niż zostały pierwotnie wytworzone lub mogą na ich podstawie prowadzić dalsze badania. Jest tak ponieważ dane badawcze są danymi wielokrotnego użytku. Jednocześnie zauważyć należy, że zakres danych podlegających ponownemu wykorzystywaniu ustalony w art. 10 dyrektywy i art. 22 ustawy nie jest tożsamy.

Polski ustawodawca wskazał, że ponownemu wykorzystywaniu podlegają dane badawcze będące w posiadaniu podmiotów, o których mowa w art. 7 ust. 1 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce. W ten sposób wyłączył spod stosowania ustawy części badań finansowanych ze środków publicznych na podstawie ustawy z 30 kwietnia 2010 r. o Narodowym Centrum Badań i Rozwoju (t.j. Dz. U. z 2024 r. poz. 1170). Podmioty, o których mowa w art. 37 pkt 4-10 tej ustawy mogące startować w konkursach na wykonanie projektów nie będąc podmiotami, o których mowa w art. 7 ust. 1 prawo o szkolnictwie wyższym. Podobnie jest, jak chodzi o podmioty, o których mowa w art. 27 ust. 1 pkt 2-9 ustawy z 30 kwietnia 2010 r. o Narodowym Centrum Nauki (t.j. Dz. U. z 2023 r. poz. 153) oraz podmioty, o których mowa w art. 17 ust. 1 pkt 2-5 ustawy

z 21 lutego 2019 r. o Agencji Badań Medycznych (t.j. Dz. U. z 2023 r. poz. 2064). Z charakteru tych trzech wymienionych podmiotów wynika, że finansowane przez nie badania mają charakter działalności naukowej, o której mowa w art. 4 ustawy z dnia 20 lipca 2018 r. – prawo o szkolnictwie wyższym i nauce. Co za tym idzie więc dane badawcze przez nie wytworzone powinny podlegać ponownemu wykorzystywaniu.

Polska nie dokonała również prawidłowej transpozycji dyrektywy o otwartych danych w odniesieniu do danych badawczych. Przed wszystkim nie została do chwili obecnej opracowana polityka krajowa, o której mowa w art. 23 ustawy o otwartych danych. Za taki nie może być uznany program otwierania danych na lata 2021-2027³³. Odwołuje się on do rekomendacji zawartych w „Kierunkach rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce”. Należy do nich: opracowanie i przyjęcie polityk w zakresie otwartego dostępu przez podmioty tworzące system szkolnictwa wyższego i nauki, o których mowa w art. 7 ust. 1 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce, b) wyznaczenie przez kierowników jednostek naukowych i uczelni pełnomocników do spraw otwartości danych, c) przechodzenie czasopism naukowych do modeli otwartych, d) udostępnienie rozpraw doktorskich w otwartych repozytoriach, e) monitorowanie i raportowanie do MNiSW (od 1 stycznia 2021 r. do MEiN) postępów we wdrażaniu OD, w tym systematyczne analizowanie liczby publikacji powstających w danej jednostce naukowej lub uczelni w celu określenia proporcji publikacji w OD w stosunku do wszystkich publikacji, f) organizowanie szkoleń w zakresie OD dla pracowników naukowych i doktorantów danej jednostki naukowej lub uczelni, g) uwzględnienie doświadczeń i potencjału bibliotek naukowych, które często koordynują proces redakcji i deponowania publikacji naukowych w repozytoriach³⁴.

³³ Uchwała nr 28 Rady Ministrów z dnia 18 lutego 2021 r. w sprawie Programu otwierania danych na lata 2021-2027 (M. P. z 2021 r. poz. 290). Treść programu <https://sipdata-1lex-1pl-1fy8aq6g108f1.han.uksw.edu.pl/dane/monitory/2021/2468546/1.pdf>

³⁴ Zob. Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce, przyjęty w 2015 r. file:///Users/Agnieszka/Downloads/20180413_Kierunki_rozwoju_OD_wersja_ostateczna%20(2).pdf [dostęp 9.08.2023], s. 72.

W programie pojawia się pojęcie danych badawczych. Są one zdefiniowane w sposób zgodny z dyrektywą, ale jego treść w tym przede wszystkim pkt 5.4. dotyczący tej problematyki odnosi się właściwie wyłącznie do publikowania w otwartym dostępie. Oznacza to więc, że Polska nie przyjęła w tym zakresie polityki krajowej, ani nie podjęła innych odpowiednich działań.

4. Wyzwania dla uczelni wyższych. Próba podsumowania

Jak wskazano w tekście, zapewnienie powszechnej dostępności danych badawczych na poziomie wyższej uczelni jest celem o charakterze długoterminowym. O jego praktycznej realizacji na chwilę obecną nie możemy jeszcze mówić. Powodów tego stanu rzeczy można dopatrywać się z jednej strony w bardzo zachowawczym podejściu do otwierania nowego zasobu przez UE. Jest to jednak dla niej zachowanie typowe. Zapewne przy kolejnej zmianie dyrektywy będziemy obserwowali wzmocnienie obowiązków w tym zakresie, tak by w konsekwencji w dalszej perspektywie zostało zagwarantowane prawo do ponownego wykorzystywania danych badawczych finansowanych ze środków publicznych. Przede wszystkim jednak, deficyty jak chodzi o ponowne wykorzystywanie danych badawczych, wynikają z nieprawidłowej implementacji i transpozycji dyrektywy do prawa krajowego, o czym już była mowa.

Sytuacja szkoły wyższej w opisanej sytuacji nie jest godna podziwienia. W związku z brakiem polityki krajowej, ale również podejmowania działań na rzecz kultury otwartości otwieranie danych badawczych nie jest zadaniem łatwym. Trudno oczekiwać by podmioty, o których mowa w art. 7 ustawy Prawo o szkolnictwie wyższym oddolnie podjęły działania w tym zakresie zastępując podmioty zobowiązane. Warto jednak w swojej działalności dostrzegać ten problem, nawet jeżeli będzie to polegało wyłącznie na tym by propagować otwieranie danych wśród samych naukowców poprzez rozszerzanie ich wiedzy i budowanie kultury otwartości w tym zakresie. Warto również propagować wiedzę jak chodzi o zasady FAIR. Dane badawcze to bardzo rozbudowany zasób o zróżnicowanym charakterze. Naturalnym procesem jest więc dostosowywanie zasad FAIR do potrzeb prowadzonych badań i stałych zmian technologicznych. Promować należy wszelkie oddolne działania w tym zakresie.

Opening up research data from the perspective of higher education

The article addresses the issue of reusing research data as part of the right to public sector information in the context of the development of the ideas of open government and open science. The aim of the study is to analyze the legal regulations concerning the reuse of research data at the level of European Union law and Polish law. The study reconstructs the concept of 'research data' and discusses the principles of its disclosure, including the principle of 'default openness' and the FAIR standard. It also presents the conditions, scope and limitations of the re-use of research data, in particular with regard to public funding and the sharing of data through institutional and thematic repositories. The conclusion identifies the main challenges facing higher education institutions in implementing this issue.

Słowa kluczowe: dane badawcze; otwarte dane; ponowne wykorzystywanie; dyrektywa 2019/1024; FAIR; otwarta nauka

Keywords: research data; open data; reuse; Directive 2019/1024; FAIR; open science

Bibliografia

- DIETRICH N., WIEBE A., *Definicja danych badawczych*, w: *Otwarte, czyli bezpieczne. Analiza prawnych form ochrony danych badawczych i zalecenia w kwestii dostępu i korzystania*, red. Guibault L., Wiebe A., Kraków 2014, s. 78-88.
- FECHER B., FRIESIKE S., *Open Science: One Term, Five Schools of Thought*, SSRN Electronic Journal Conference paper, April 2013.
- FISCHER B., *Autorskoprawne konteksty ponownego wykorzystywania danych badawczych*, w: *Sto lat polskiego prawa handlowego. Księga jubileuszowa dedykowana Profesorowi Andrzejowi Kidybie*, t. 2, red. Dumkiewicz M., Kopaczyńska-Pieczniak K., Szczotka J., Warszawa 2020, s. 17-32.
- FISCHER B., PISKORZ-RYŃ A., SAKOWSKA-BARYŁA M., WYPORSKA-FRANKIEWICZ J., *Ustawa o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego. Komentarz*, Warszawa 2022.
- HODSON S. i in., *Turning FAIR into reality, Final Report and Action Plan from the European Commission Expert Group on FAIR Data*, 2018, doi 10.2777/1524.

- IZDEBSKI H., KULESZA M., *Administracja publiczna, zagadnienia ogólne*, Warszawa 2004.
- LONGCHAMPS F., *Założenia nauki administracji*, Wrocław 1991.
- KULCZYŃSKI E., *Otwarta nauka a komunikacja – perspektywa metateoretyczna*, w: *Komunikacja naukowa w humanistyce*, Poznań 2017, s. 34-49.
- KURCZ B., *Dyrektywy Wspólnoty Europejskiej i ich implementacja do prawa krajowego*, Kraków 2004.
- MONS B. i in., *Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud*, s. 5, *Information Services and Use* 37 (2017), doi 10.3233/ISU-170824.
- PISKORZ-RYŃ A., *Ponowne wykorzystywanie danych badawczych*, w: *Prawo sztucznej inteligencji i nowych technologii*, red. Fischer B., Pązik A., Świerczyński M., Warszawa 2024, s. 17-22.
- PISKORZ-RYŃ A., *Ponowne wykorzystanie informacji sektora publicznego. Zagadnienia administracyjnoprawne*, Warszawa 2018.
- WILKINSON M.D., i in., *The FAIR Guiding Principles for scientific data management and stewardship*. *Sci. Data* 3 (2016), doi 10.1038/sdata.2016.18.

NOTA O AUTORZE

dr hab. Agnieszka Piskorz-Ryń, prof. uczelni, pracownik naukowo-dydaktyczny nauk prawnych na Wydziale Prawa i Administracji Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie.

Artykuł stanowi opracowanie własne autora/autorów. Treść uzupełniona o pracę badawczą została wygłoszona podczas międzynarodowej konferencji naukowej realizowanej w maju 2024 roku w ramach programu „Nauka dla Społeczeństwa” nr umowy NdS/550272/2022/2022 pod tytułem „Rola systemu szkolnictwa wyższego i nauki. Historia – terażniejszość – ku przyszłości” realizowanego przez Fundację „Pro Polonia” we współpracy z Wydziałem Prawa Kanonicznego oraz Wydziałem Prawa i Administracji Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie.

