Uwolnić potencjał danych.

Dane jako zasób wspólny.

instrat

Blanka Wawrzyniak
Marta Musidłowska

Warszawa, maj 2022

Fundusze Europejskie
Wiedza Edukacja Rozwój

Rzeczpospolita Polska

Unia Europejska
Europejski Fundusz Społeczny

Uwolnić potencjał danych

Dane jako zasób wspólny

Instrat Raport 05/2022
Blanka Wawrzyniak
Marta Musidłowska

Rekomendujemy cytowanie:
Wawrzyniak, B., Musidłowska, M. (2022) *Uwolnić potencjał danych. Dane jako zasób wspólny.* Instrat Raport 05/2022

Autorzy:
Blanka Wawrzyniak
Marta Musidłowska

Kontakt:
Blanka Wawrzyniak, Liderka programu badawczego Gospodarka Cyfrowa
blanka.wawrzyniak@instrat.pl

Projekt okładki: Anna Olczak
Ilustracja na okładce: Anna Olczak
Skład: Anna Olczak

Wsparcie merytoryczne: Jan J. Zygmuntowski

Warszawa, maj 2022

Instrat - Fundacja Inicjatyw Strategicznych
ul. Oleandrów 7/16
00-629 Warszawa
www.instrat.pl

Spis treści

Słownik pojęć

Dane dotyczące zdrowia	zgodnie z wytycznymi Europejskiej Rady Ochrony Danych (EROD), do danych tych należą: ● informacje zebrane przez świadczeniodawcę opieki zdrowotnej w dokumentacji medycznej pacjenta; ● informacje, które stały się danymi dotyczącymi zdrowia w wyniku ich odniesienia do innych danych, co ujawniło stan zdrowia lub zagrożenia dla zdrowia; ● informacje przekazane w ankietach „samokontroli" przez osoby, których dane dotyczą, w ramach udzielanych odpowiedzi na pytania dotyczące ich stanu zdrowia (opisy objawów); ● informacje, które stały się danymi dotyczącymi zdrowia ze względu na sposób ich wykorzystania w określonym kontekście (np. informacje dotyczące niedawnej podróży lub obecności w regionie dotkniętym COVID-19).
Dług innowacyjny	nieuzasadnione koszty, które ponosi gospodarka kraju lub blok gospodarczy z powodu braku odpowiednich inwestycji we własne innowacje (Borowik et al, 2018). W odniesieniu do współdzielenia danych, dług innowacyjny oznacza **utratę** przez polskich przedsiębiorców **potencjalnych zysków bądź ponoszenie dodatkowych kosztów wynikających z braku dostępu do ich narzędzi sztucznej inteligencji i innowacyjnych rozwiązań.** Charakterystyczną metodą zmniejszania długu innowacyjnego są "żabie skoki" (leapfrogging), czyli nagła modernizacja branży zapóźnionej przez inwestycje bezpośrednio w najlepsze technologie, z pominięciem przejściowych i poprzedniej generacji.
DICOM	standard określający f**ormat i sposób transmisji danych obrazowych między urządzeniami obrazującymi** (aparaty TK, MRT, cyfrowe angiografy czy cyfrowe aparaty rtg) a jednostkami służącymi do ich analizy i wtórnego przetwarzania (diagnostyczne stanowiska opisowe), czy też systemami archiwizacji (infoRadiologia).
ENISA	Agencja Unii Europejskiej ds. Cyberbezpieczeństwa.
HL7	standard cyfrowej wymiany informacji w środowiskach medycznych. **Protokoły opisane w tym standardzie dotyczą warstwy aplikacyjnej (siódmej) modelu OSI**. To protokół komunikacyjny służący do wymiany danych medycznych, który definiuje komunikaty poziomu aplikacji używane przez

	kilka głównych systemów szpitalnych. Główne funkcje systemu obejmują komunikaty dotyczące: dostępu do danych, pobierania danych, przesyłania danych, sterowania, pobierania wyników i obserwacji klinicznych.
Interoperacyjność silna	zdolność systemu lub produktu do pełnej współpracy z innymi systemami lub produktami o charakterze międzysektorowym i powszechnym, w tym np. przekazywania danych z sektora prywatnego do publicznego i odwrotnie
Interoperacyjność słaba	zdolność systemu lub produktu do współpracy z innymi systemami lub produktami jedynie w obrębie jednego sektora, najczęściej publicznego.
Kapsuły danych	prywatne silosy danych służące indywidualnym osobom do gromadzenia danych, które ich dotyczą lub danych wygenerowanych z użytkowanych przez nie urządzeń.
Pośrednicy danych	zaufana trzecia strona współdzielenia danych, pełniąca rolę mediatora między tymi, którzy chcą udostępnić swoje dane a tymi, którzy chcą je wykorzystać; w odpowiedni sposób zarządza danymi oraz udziela wsparcia użytkownikom w dokonywaniu świadomych wyborów z zakresie wyrażania zgody na przetwarzanie ich danych (Janssen, H., Singh, J. 2022).
RODO	Rozporządzenie 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 swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych), OJ L 119, 4.5.2016, p. 1–88.
Składnica danych	przestrzeń dla danych przemysłowych, rolniczych, biznesowych, której celem jest umożliwianie działań z zakresu analizy biznesowej (Business Intelligence, BI), w szczególności analityki. Model oparty na współpracy między zainteresowanymi podmiotami; może występować w formie platformy zarządzanej przez niezależnego pośrednika (Zaufany Pośrednik) bądź przez wszystkich interesatiuszy budujących wspólną strategię w zakresie wymiany danych (Data Pooling).
Suwerenność cyfrowa	zdolność państw, organizacji międzynarodowych i każdego użytkownika i użytkowniczki z osobna do egzekwowania swoich praw oraz wpływania na platformy cyfrowe i firmy technologiczne zgodnie z własnymi potrzebami społecznymi i rozwojowymi" (Wawrzyniak, Zygmuntowski i Lamański, 2020).
Sztuczna inteligencja	zdolność systemu do prawidłowego interpretowania danych pochodzących z zewnętrznych źródeł, nauki na ich podstawie

	oraz wykorzystywania tej wiedzy, aby wykonywać określone zadania i osiągać cele poprzez elastyczne dostosowanie (Kaplan i Haenlein, 2019)
Ślad cyfrowy	zapis każdego zachowania użytkownika w Internecie. Składa się z dwóch części - danych o użytkowniku ustalanych na podstawie adresu IP, dzięki któremu można ustalić m.in. położenie geograficzne urządzenia, oraz danych o ruchu w cyberprzestrzeni poprzez zalogowanie się użytkownika do sieci, dzięki którym cały ruch wykonany na komputerze zostanie zapisany na serwerze dostawcy usług telekomunikacyjnych (Nowoczesna Polska, 2013).
Wspólnice danych	zaufana instytucja zarządzająca danymi zgodnie z interesem publicznym, przede wszystkim dla danych osobowych o znaczeniu publicznym (np. dane zdrowotne), wytwarzanych w systemie usług publicznych, danych społecznościowych, administracyjnych czy samorządowych. Przy zapewnieniu odpowiednich zabezpieczeń prywatności, instytucja ta przyznaje różnym podmiotom (zarówno publicznym, jak i prywatnym) możliwość dostępu do zgromadzonych danych na określonych zasadach i za odpowiednią opłatą.

1. Wprowadzenie

Wraz z pojawianiem się nowych rozwiązań technologicznych i coraz powszechniejszym dostępem do sieci, stopniowemu przeobrażeniu ulegał nie tylko globalny rynek gospodarczy, ale przede wszystkim cyfrowa świadomość społeczna. Pierwsi użytkownicy Internetu wierzyli, że dostęp do darmowych i otwartych treści - przy zachowaniu anonimowości - może ułatwić proces demokratyzacji wiedzy. (Mayer-Schönberger, Ramge, 2022). Pomimo iluzorycznej nieodpłatności usług sieciowych, w rzeczywistości osoby z nich korzystające same przyczyniały się do powiększania zasobów informacji w Internecie, pozostawiając po sobie i swojej działalności bezcenny, cyfrowy ślad (Nowoczesna Polska, 2013).

Prawdziwą wartość ekonomiczną danych ujawniły dopiero światowej skali nadużycia największych spółek branży technologicznej. Firmy te przez wiele lat unikały opodatkowania ich działalności i przestrzegania jakichkolwiek, wewnątrzkrajowych regulacji twierdząc, że nie da się jednoznacznie ustalić, gdzie znajduje się faktyczna alokacja ich kapitału. W konsekwencji, 80% bogactwa korporacyjnego znalazło się w rękach zaledwie 10% światowych przedsiębiorstw (Foroohar, 2019).

Jednocześnie cele, dla których firmy te wykorzystywały zbierane zasoby danych, znacznie wykraczały poza zwyczajowe podstawy przetwarzania informacji związane m.in. z poprawą wyświetlanych sugestiii czy wyników wyszukiwania. Dochodziło bowiem do sytuacji, w których dane służyły do tendencyjnego profilowania użytkowników, wzmacniania krzywdzących stereotypów i nierówności czy radykalizowania niektórych grup społecznych,

świadomie destabilizując porządek publiczny i demokratyczne rządy (Zygmuntowski, 2020a).

Efektem tego stanu rzeczy była zmiana paradygmatu myślenia w stronę zwiększenia ochrony prywatności w sieci, a także własnościowego traktowania danych osobowych. Publiczne ujawnienie skandalu Cambridge Analytica zbiegło się z wejściem w życie RODO, które do dziś stanowi istotny punkt wyjścia do dyskusji o możliwości "swobodnego przepływu danych osobowych" (Rozporządzenie o Ochronie Danych Osobowych, 2016). Głównym celem RODO miało być odzyskanie kontroli jednostki nad danymi, które jej dotyczą poprzez zobowiązanie platform do ujawnienia informacji na temat sposobu przetwarzania danych i zapewnienie, aby informacje te były dostępne w przejrzysty i zrozumiały sposób. Wymaganie uzyskania zgody na przetwarzanie danych osobowych od użytkowników platform przyczyniło się do wzrostu świadomości w zakresie posiadanych praw cyfrowych i pogłębienia niechęci wobec największych spółek technologicznych. Jak wskazuje bowiem raport Polskiego Instytutu Ekonomicznego, przeciętni użytkownicy oczekują pieniężnej rekompensaty w zamian za szeroko zakrojony dostęp do danych i wyświetlanie spersonalizowanych reklam przez platformy cyfrowe (Polski Instytut Ekonomiczny, 2020).

Przełomowym krokiem w kierunku zmiany myślenia o danych jako o własności indywidualnej jednostki było złożenie skargi przez Maximiliana Schremsa do irlandzkiego organu ochrony danych osobowych, dotyczącej zasad umożliwiających przekazywanie danych osobowych z UE do Stanów Zjednoczonych. W jej wyniku, TSUE unieważnił Tarczę Prywatności orzekając jednocześnie, że dalszy transfer danych na podstawie niniejszej decyzji jest zabroniony. Pomimo działania w imieniu własnym, naświetlony przez Maxa Schremsa problem dotyczył w rzeczywistości europejskich danych rozumianych jako pewne dobro wspólne, zasługujące na co najmniej tak wysublimowaną ochronę, jaką zapewnia RODO.

Przyznawanie danym wyłącznie przymiotu własności wydaje się z natury rzeczy hamować korzystne społecznie wykorzystanie danych. To właśnie takie podejście sprawia, że firmy nie dostrzegają wielu możliwości, w których gromadzone i przechowywane przez nie dane tworzą również cenną wartość publiczną (Swant, 2019). Pojawienie się i dynamiczny rozwój narzędzi technologicznych opartych na systemach sztucznej inteligencji w sposób szczególny uwydatniło konieczność odejścia od towarowego podejścia do danych i ich ponownego wykorzystania w celu zapewnienia ogólnego społecznego wzrostu i wytworzenia Wspólnej Wartości Społecznej (Shared Social Value).

Chociaż więc niektórzy określają dane jako "nową ropę" (np. The Economist), w rzeczywistości posiadają one fundamentalnie odmienne cechy. Przede wszystkim, cyfrowe dane nie są konkurencyjne w konsumpcji, ponieważ ich zasobów nie da się wyczerpać poprzez wielokrotną eksploatację (Zygmuntowski, 2020a). Nie pochodzą one z samoistnie występujących złóż, lecz są wynikiem aktywności człowieka (dane osobowe) bądź urządzeń obsługiwanych przez człowieka (dane nieosobowe). Odpowiedzialne zarządzanie danymi musi więc uwzględniać prawa ludzi, dając im podmiotowość w zakresie ochrony prywatności czy udzielania dostępu do danych, ale też brać pod uwagę możliwość wielokrotnego wykorzystywania informacji w rozmaitych celach.

Z potencjału danych powinni zatem korzystać nie tylko giganci technologiczni, ale także (albo przede wszystkim) organy administracji, dostawcy usług publicznych, ośrodki naukowo-badawcze, organizacje pozarządowe i nietechnologiczne MŚP. Ignorując bowiem

potęgę informacji cyfrowej, podmioty te pozbawiłyby się dostępu do nowych, inteligentnych rozwiązań, a tym samym groziłoby im zacofanie technologiczne i wynikające z niego straty. Najbardziej perspektywiczną strategią dla nowoczesnych gospodarek jest zarządzanie **danymi jako wspólnym zasobem**, o charakterystyce infrastruktury, a nie zwykłego towaru. Maksymalizacja potencjału danych to szansa rozwojowa dla Polski, by wytworzyć własne wysokotechnologiczne rozwiązania, zmodernizować usługi publiczne i podnieść jakość podejmowania decyzji we wszystkich sektorach. Aby to osiągnąć, musimy uwolnić potencjał danych poprzez ich efektywne współdzielenie w ramach zaufanych przestrzeni, instytucji i technologii powołanych do tego celu. Chociaż instytucje służące do dzielenia się danymi to całkowicie nowa koncepcja, ich powstanie może być porównane do niegdyś innowacyjnego pomysłu zakładania banków czy spółek akcyjnych, bez których funkcjonowanie współczesnej gospodarki jest trudne do wyobrażenia.

2. Korzyści płynące ze współdzielenia danych

Ze względu na duże znaczenie ochrony prywatności w dyskusji o danych, do tej pory przypisywało się im wartość w ujęciu jednostkowym, przez pryzmat odpowiedniego egzekwowania praw osób, których dane dotyczą. W związku z tym, zdaniem niektórych można mówić o “prywatnej własności danych”, pozwalając na ukształtowanie się rynku ich komercyjnej wymiany. Takie podejście doprowadzi jednak do zamykania danych w prywatnych silosach i utraty potencjału płynącego z agregacji i dalszego wykorzystywania. Współdzielenie danych, aby zmaksymalizować korzyści, musi jednak uniknąć pułapki utowarowienia danych, ale również pułapki indywidualnej prywatności. Zarządzanie danymi prezentuje się raczej jako trylemat, który musi być rozstrzygany w zaufanej, bezpiecznej przestrzeni współpracy - zarówno w kontekście infrastruktury, jak i procedur instytucji.

Wzajemne zaufanie uczestników procesu współdzielenia danych odgrywa zdecydowanie kluczową rolę. Współpraca na szeroką skalę, zarówno międzysektorowa jak również wewnątrz danej branży, pozwala na ukształtowanie świadomego procesu decyzyjnego w zakresie dalszego, bardziej zrównoważonego rozwoju. Jedynie uczciwa i transparentna współpraca interesariuszy jest w stanie wyeliminować nadmierną dominację cyfrową opartą na ograniczonym dostępie do informacji. Koncentracja władzy informacyjnej jest bowiem korzystna dla nielicznych, ponieważ hamuje innowacje i utrudnia dostęp do korzyści dla społeczeństwa, a co za tym idzie - każdego z nas (Mayer-Schönberger, Ramge, 2022).

Poniżej przedstawiamy przykładowe korzyści płynące z ponownego wykorzystania danych przez poszczególne sektory:

ADMINISTRACJA PAŃSTWOWA I SAMORZĄDY:

- redukcja kosztów usług publicznych;
- wdrażanie rozwiązań typu smart city (np. optymalizowanie konsumpcji energii poprzez inteligentne taryfy miejskie; efektywne zarządzanie transportem miejskim);
- zintegrowana wiedza nt. rynku nieruchomości i potrzeb mieszkaniowych;
- prognozowanie zużycia infrastruktury i koniecznych inwestycji;
- rozwój lokalny i tworzenie nowych firm, produktów i usług w oparciu o znane nawyki mieszkańców;
- modernizacja usług publicznych z użyciem technologii opartych o dane;
- poprawa jakości opieki zdrowotnej, urządzeń medycznych i lepsze wykrywanie chorób (np. rozwój AI w medycynie);

PRZEDSIĘBIORCY I BIZNES:

- zwiększenie dostępności danych handlowych przydatnych do prognozowania trendów;
- identyfikacja przewag konkurencyjnych;
- opracowywanie nowych produktów i usług, w tym wysokotechnologicznych;
- budowanie strategii cenowych i analiza rynku;
- optymalizacja procesu obsługi klienta;
- obniżenie kosztów prowadzenia firmy poprzez optymalizację logistyki;

PRZEMYSŁ I ROLNICTWO:

- wdrażanie innowacji w przemyśle (np. wynajdywanie nowych sposobów wytwarzania towarów; poprawa mechanizmów maszyn);
- obniżanie kosztów produkcji przemysłowej wskutek optymalizacji;
- poprawa wydajności energetycznej (przewidywanie zapotrzebowania na prąd i tworzenia bilansujących się źródeł; prognozowanie strat energii w sieciach; zoptymalizowane planowanie inwestycji w energetyce);
- podnoszenie potencjału technologicznego upraw żywności w Polsce;
- poprawa efektywności wykorzystania zasobów, wydajności i zrównoważenia ekologicznego;
- zapewnienie społecznościom wiejskim lepszych warunków życia
- poprawę relacji między konsumentem a różnymi uczestnikami łańcucha wartości;

NAUKI, NGO, SPOŁECZEŃSTWO OBYWATELSKIE:

- polepszenie jakości debaty publicznej i procesów podejmowania decyzji dla wspólnego dobra;

- zwiększanie partycypacji społecznej (np. tworzenie narzędzi IT służących do angażowania społeczeństwa w procesy zachodzące na szczeblu lokalnym);
- zwiększenie kontroli społecznej nad danymi;
- nadzór nad jakością i rzetelnością administracji publicznej, biznesu i innych współdzielących dane;
- medyczne zastosowanie “wearables” (np. zegarków ostrzegających przed napadem padaczkowym;
- prowadzenie badań naukowych, zarówno przez naukowców jak i naukę obywatelską.

3. Panorama praktyk międzynarodowych

3.1 Działania na szczeblu wspólnotowym

W kontekście znaczenia danych, do niedawna regulacje unijne dotyczyły jedynie indywidualnej ochrony praw podmiotu, którego cyfrowe informacje dotyczą. Ale mimo że jedną z naczelnych zasad RODO pozostaje zasada ograniczenia wykorzystania danych do jednego konkretnego celu, rozporządzenie oferuje od niej odstępstwa w określonych przypadkach. Wyraźnie dopuszcza bowiem możliwość "dalszego przetwarzania do celów archiwizacji w interesie publicznym, do celów badań naukowych lub historycznych lub do celów statystycznych” (Alemanno, 2018). Zarówno ta zasada jak i dotychczas niedostatecznie wykorzystywane prawo do przenoszalności danych (*data portability*, art. 20 RODO), stanowią podwaliny ekosystemu współdzielenia danych.

Ta zmiana paradygmatu wynika z dostrzeżenia przez regulatora unijnego potencjału danych oraz szeregu możliwości, jakie daje ich ponowne użycie. W strategii z 19 lutego 2020 r. Komisja Europejska wskazała, iż jej celem jest stworzenie wspólnej europejskiej przestrzeni danych w ramach jednolitego rynku, na którym mogłyby być one wykorzystywane – w zgodzie z obowiązującymi przepisami oraz bez względu na ich fizyczne miejsce przechowywania w Unii (Europejska strategia w zakresie danych). W związku z tym, według ustalonej strategii Unia dąży do otwierania danych publicznych, zachęcania prywatnych podmiotów do dzielenia się danymi, zwiększania dostępności sprawdzonych usług chmurowych. Zarówno działania miękkie (podnoszenie świadomości czy kreowanie zachęt), jak i tworzenie odpowiednich ram regulacyjnych (Data Governance Act; Data Act) mają sprzyjać realizacji zasadniczego, być może najbardziej ambitnego, celu unijnej polityki cyfryzacyjnej (Komisja Europejska, 2020a).

Pierwszym aktem z zakresu zestawu środków zapowiedzianych we wspomnianej strategii jest **Akt w sprawie zarządzania danymi** (Data Governance Act), który stanowi uzupełnienie dyrektywy w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego. Celem Aktu jest poszerzenie skali tego zjawiska i poprawa warunków udostępniania danych. Jak podaje Komisja we wniosku, najważniejsze postanowienia rozporządzenia obejmują cztery obszary:

- udostępnianie danych sektora publicznego do ponownego wykorzystywania w sytuacjach, w których dane te są objęte prawami innych osób;
- udostępnianie danych między przedsiębiorstwami w zamian za wynagrodzenie w dowolnej postaci;

- umożliwianie wykorzystywania danych osobowych z pomocą „pośrednika w udostępnianiu danych osobowych", który ma pomagać osobom fizycznym w wykonywaniu ich praw wynikających zogólnego rozporządzenia oochronie danych (RODO);
- umożliwianie wykorzystywania danych z pobudek altruistycznych.

Projektowane rozporządzenie szczególny nacisk kładzie także na zwiększenie zaufania do instytucji pośredników danych (*data intermediaries*). Podmioty te, określane w rozporządzeniu jako dostawcy usług udostępniania danych, mają za zadanie zapewnić bezpieczne i godne zaufania środowisko dla udostępniania danych przez ich posiadaczy. Możliwość ta dotyczy zarówno osób prawnych, jak i fizycznych. Usługi przeznaczone dla osób prawnych będą polegać na tworzeniu dedykowanych przestrzeni, w której podmioty będą mogły nie tylko umieszczać swoje dane, ale również korzystać z tych, które w niej się znalazły. Dla osób fizycznych regulator unijny przewidział natomiast ułatwienia w zakresie udostępniania danych ich ewentualnym, przyszłym użytkownikom. W tym celu stworzone zostaną odpowiednie aplikacje umożliwiające dostęp do przestrzeni/portfeli danych, za pośrednictwem których udziela się dostępu do danych podmiotom chcącym z nich korzystać. Środki te mają ułatwić osobom fizycznym korzystanie z ich uprawnień wynikających z RODO, poprzez świadome decydowanie o tym, komu udostępnić swoje dane (Małobęcka-Szwast, 2021). Więcej o instytucji pośredników danych jako jednej z metod ich współdzielenia, korzyści płynących z tego modelu jak i zidentyfikowanych barier, można przeczytać w kolejnej części raportu.

Innym projektowanym aktem regulującym zasady udostępniania danych jest **Akt w sprawie danych (Data Act)**. Dotyczy on m.in. danych generowanych przez użytkowników przy pomocy urządzeń IoT (Internet of Things) i możliwości ich ekstrahowania w celu dalszego wykorzystania lub przekazania innym podmiotom. Projekt uwzględnia również dostęp sektora publicznego do danych podmiotów prywatnych w wymagających nadzwyczajnej interwencji przypadkach (np. w sytuacjach klęski żywiołowej). Dostawcy usług przetwarzania danych oraz operatorzy przestrzeni, w których dane się znajdą, zostaną również zobowiązani do przestrzegania reguł dotyczących interoperacyjności formatów rejestrowania danych (Małobęcka-Szwast, 2021).

Stosunkowo nowym przedsięwzięciem o zasięgu sektorowym jest projekt Komisji Europejskiej dotyczący stworzenia **Europejskiej przestrzeni danych dotyczących zdrowia** (Komisja Europejska, 2022). Jego głównym założeniem jest zachęcanie do korzystania z danych dotyczących zdrowia do celów badawczych, kształtowania polityki i stanowienia prawa, w oparciu o zaufane ramy zarządzania i zasady ochrony danych. Projekt zwraca uwagę również na kwestie bezpieczeństwa i odpowiedzialności w kontekście korzystania z AI w dziedzinie zdrowia, promując jednocześnie proaktywną postawę obywateli w zakresie kontroli nad ich danymi zdrowotnymi. W jego najnowszej wersji pojawia się również instytucja "data access body", czyli organu na szczeblu krajowym udzielającego pozwoleń na dostęp do danych.

Na początku marca 2022 roku wypłynął natomiast projekt Komisji Europejskiej dotyczący nowych **ram zarządzania danymi dotyczącymi zdrowia**. Ma on na celu wprowadzenie infrastruktury i wymogów interoperacyjności obejmujących cały obszar Unii Europejskiej (Bertuzzi & Fortuna, 2022). Projekt jest spójny z wcześniej wymienionymi Data Governance Act oraz Data Act, które także pozostają na etapie projektowania. Zgodnie z projektem, to pacjenci mają być głównymi podmiotami decydującymi o ograniczaniu dostępu do swoich danych, czy ich bezpłatnym udostępnianiu. Inaczej niż dotychczas, zbierane dane będą pochodzić z różnych źródeł: dokumentacji zdrowotnej, rejestrów publicznych, danych o

charakterze społecznym, administracyjnym, genetycznym i genomicznym, badań klinicznych, kwestionariuszy badawczych oraz danych biomedycznych (np. biobanki). Dane te będą mogły być wykorzystywane m.in. w działalności organów publicznych w wykonywaniu ich zadań, do celów badawczo-rozwojowych i naukowych, czy tworzenia nowych rozwiązań dla interesu publicznego. Ponadto, zgodnie z projektem osobom fizycznym przysługiwać będzie bezpłatny dostęp do minimalnego zestawu "podstawowych" danych dotyczących zdrowia, w tym szczepień, elektronicznych recept, zdjęć, wyników badań laboratoryjnych, raportów z wypisów i innych. Niezwykle ważna jest również kwestia zapewnienia odpowiedniego bezpieczeństwa systemów elektronicznych kart zdrowia (EHR) - będą one musiały spełniać ściśle określone wymagania techniczne, w tym związane z ich interoperacyjnością.

3.2 Działania na szczeblach krajowych

Ze względu na skomplikowany proces realizowania polityk unijnych, niektóre europejskie państwa już wcześniej zaczęły podejmować kroki mające na celu uwspólnianie danych oraz budowanie dedykowanych im otwartych przestrzeni. Ekosystemy współdzielenia danych funkcjonują już bądź są wdrażane w takich krajach jak Estonia, Wielka Brytania, Niemcy, Finlandia czy Francja.

Poniżej przedstawiamy niektóre rozwiązania:
Estonia
XRoad:

- Dostęp do danych możliwy tylko za pomocą kart identyfikacyjnych służących do uwierzytelniania i podpisów cyfrowych;
- Rozwiązanie dla sektora publicznego - podmioty spoza XRoad nie mają dostępu do zgromadzonych tam danych;
- Obowiązek pracowników służby zdrowia dotyczący przesyłania danych do systemu informacji zdrowotnej (HIS), dostępnej wyłącznie dla licencjonowanych pracowników;
- Możliwość korzystania z większości ogólnoeuropejskich danych znajdujących się w głównych bazach w Estonii;
- Dostęp do danych możliwy tylko za pomocą kart identyfikacyjnych służących do uwierzytelniania i podpisów cyfrowych;
- Wysokie zabezpieczenia np. nakładki na dane poszczególnych pacjentów, które ujawniają tylko niezbędne informacje;

Wielka Brytania
Open Data Institute:

- Pozarządowy instytut badawczy współpracujący z podmiotami z różnych sektorów na rzecz tworzenia pilotaży w zakresie bezpiecznych i etycznych ekosystemów danych;
- Współpraca zarówno z podmiotami prywatnymi, jak i publicznymi;
- Cel: wspieranie podmiotów z różnych sektorów w budowaniu otwartych, godnych zaufania ekosystemów danych w ich organizacjach;
- Przeprowadzenie badania wspólnie z brytyjskim Biurem ds. Sztucznej Inteligencji i Innowacji służącego poddaniu ocenie potencjału płynącego z wykorzystania jednego z modeli współdzielenia danych (*data trusts)* na podstawie 3 programów pilotażowych: dotyczących danych miejskich, danych żywnościowych i o międzynarodowym nielegalnym handlu dziką fauną i florą;

Niemcy

Daten-Treuhänder:

- Planowany pośrednik dla danych przemysłu motoryzacyjnego oraz ubezpieczycieli;
- Zbieranie danych dotyczących wypadków motoryzacyjnych i napraw środków transportu motoryzacyjnego;
- Dane udostępniane za zgodą właścicieli samochodów;
- Dostęp do danych dla sektora publicznego, ubezpieczycieli, związków inspekcji technicznej i centrum serwisowych;
- Dane udostępniane do pracy nad zwiększeniem bezpieczeństwa na drogach oraz usprawnianiem budowy i naprawy maszyn;

Finlandia

Findata:

- Fiński Urząd ds. Pozwoleń na Udostępnianie Danych Społecznych i Zdrowotnych;
- Publiczna instytucja zarządzająca dostępem do różnych zbiorów danyc dotyczących zdrowia i spraw społecznych;
- Przechowywanie danych zarówno od prywatnych, jak i publicznych dostawców usług medycznych;
- Udzielanie dostępu do danych;
- System odpłatnego udostępniania danych;
- Różnicowanie kwoty za możliwość korzystania z danych w zależności od ilości danych i celu ich wykorzystania;
- System opt-out dla obywateli (od rozpoczęcia działalności Findata zaledwie 200 osób zgłosiło taką chęć);
- Dostęp udzielany różnym podmiotom dla prowadzenia badań;
- Rezultat badań, do których potrzebne były dane Findata powinien być udostępniony publicznie;
- Wprowadzenie funkcji kontrolerów danych (osób czuwających nad ich kompletnością i odpowiednim wprowadzaniem);

Francja

Health Data Hub:

- Centralny punkt dostępu do danych;
- Zbieranie różnego rodzaju danych dotyczących zdrowia, m.in. związanych z refundacją ubezpieczenia zdrowotnego, niezależnie od podmiotu i rodzaju usługi medycznej;
- Zabezpieczanie danych poprzez ich pseudonimizację;
- Wykorzystywanie danych jedynie dla celów związanych z interesem publicznym;
- Konieczna uprzednia zgoda Krajowej Komisji Ochrony Danych i Wolności (CNIL);

Polska:

Paradoksalnie, pomimo prężnej cyfryzacji usług publicznych w ostatnich latach (m.in. mObywatel, Internetowe Konto Pacjenta), Polska plasuje się na końcu rankingu Digital Economy and Society Index przygotowanego przez Komisję Europejską w 2021 roku (Komisja Europejska, 2021). Wynika to z wielu różnych czynników, które miały wpływ na ostateczne wyniki zestawienia. Pomimo podwyższonego wskaźnika dostępności do usług publicznych za pośrednictwem Internetu dla obywateli, poziom umiejętności cyfrowych Polaków wypada poniżej ogólnoeuropejskiej średniej (44 % posiada umiejętności na poziomie podstawowym, podczas gdy średnia wynosi 56 %). Mimo to, według raportu Open Data Maturity z 2021 roku, wskazującego stopień postępu w dziedzinie otwierania danych publicznych w poszczególnych państwach w Europie, Polska plasuje się na 4. miejscu,

osiągając poziom dojrzałości danych oszacowany na 95 % (Van Hesteren et al 2021), zaraz za państwami takimi jak Francja, Irlandia i Hiszpania.

Sukces polskich projektów cyfryzacji administracji publicznej i ochrony zdrowia nie powinien jednak przysłonić ich niedoskonałości i barier, z którymi należy się zmierzyć. Internetowe Konto Pacjenta i związana z nim Elektroniczna Dokumentacja Medyczna w dalszym ciągu zapewniają jedynie interoperacyjność słabą. Obowiązek stosowania jednolitych standardów został nałożony bowiem jedynie na przedsięwzięcia związane z utworzeniem Elektronicznej Platformy Gromadzenia, Analizowania i Udostępniania zasobów cyfrowych o Zdarzeniach Medycznych, czy platformy udostępniania rejestrów medycznych prywatnym placówkom ochrony zdrowia (Ministerstwo Zdrowia, 2018). Projekty te nie dotyczą wymiany danych o pacjentach między sektorem publicznym i prywatnym. Na tle innych państw Unii Europejskiej, pod względem interoperacyjności danych medycznych zawartych w Elektronicznej Dokumentacji Medycznej pacjenta, Polska razem z Rumunią znajdują się na ostatnim miejscu (Empirica, 2022).

Sukces programu Otwarte Dane daje nadzieję na powodzenia programu współdzielenia danych - zarówno osobowych, jak i nieosobowych, pochodzących od sektora publicznego, jak również przedsiębiorstw czy zwykłych obywateli. Polska ma wciąż jeszcze szansę znaleźć się w gronie państw promujących egalitarny, uspołeczniony charakter baz danych oraz zapewniających ich dostępność dla różnych aktorów społecznych i gospodarczych. Aby dokonać zmian w dotychczasowym modelu zarządzania cyfrowymi informacjami, konieczne jest zidentyfikowanie występujących barier, określenie priorytetów oraz opracowanie wytycznych i standardów współdzielenia danych w zaufanych przestrzeniach.

4. Modele dzielenia się danymi

Współdzielenie danych to umożliwianie dostępu do cennych zasobów informacji i wspieranie bardziej efektywnego i świadomego rozwoju w różnych dziedzinach. Aby jednak uniknąć ich powtórnego zamknięcia przez wybrane podmioty lub sprzecznego z interesem publicznym wykorzystania zasobów, konieczne jest stworzenie reguł udostępniania.

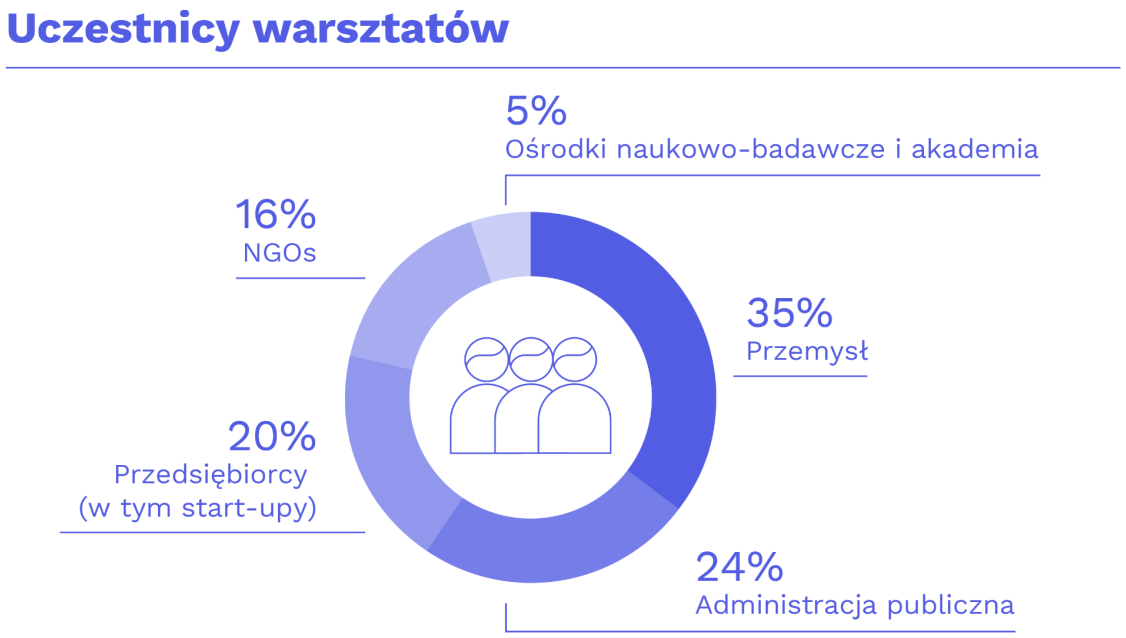
W niniejszej publikacji przedstawione zostały podstawowe modele zarządzania danymi zidentyfikowane podczas warsztatów “Uwolnić potencjał danych”. Cały cykl prac obejmował cztery bloki warsztatowe, trwające łącznie 40 godzin, a każdy z modułów zawierał zarówno część teoretyczną, jak i badawczą (dyskusja; *problem-solving*). Pierwsze trzy spotkania dotyczyły kolejno: pośredników danych osobowych; wirtualnych składnic dla danych nieosobowych; wspólnic dla danych osobowych. Czwarty blok poświęcony był analizie sposobów zarządzania danymi w zależności od stopnia ich wrażliwości.

		Kontrola interesariuszy / sposoby zarządzania	
		INDYWIDUALNA / OPARTE NA UPRAWNIENIACH	**INSTYTUCJONALNA / OPARTE NA ZAUFANIU**

Alokacja wartości/ cele zarządzania	**PRYWATNA/ ZYSK**	Kapsuły danych	Wirtualne składnice danych
	PUBLICZNA/ DOBRO WSPÓLNE	Pośrednicy danych	Wspólnice danych

Źródło: *Zygmuntowski, J. J., Zoboli, L., Nemitz, P. F. (2021). Embedding European values in data governance: a case for public data commons. Internet Policy Review, 10(3). https://doi.org/10.14763/2021.3.1572*

W projekcie wzięli udział przedstawiciele administracji państwowej, organizacji pozarządowych, przemysłu, biznesu, a także reprezentanci ośrodków naukowo-badawczych i środowisk akademickich, w tym pracujący w wiodących instytucjach współdzielenia danych z zagranicy.



Podczas warsztatów oraz późniejszej analizy zastosowaliśmy poniższe metody badawcze:

- **Analiza SLEPT** (*Socio-cultural, Legal, Economic, Political, Technological*) - metoda generalnej segmentacji makro-otoczenia i identyfikacji trendów, barier oraz innych zmiennych mająca na celu analizę rynku w wielu kontekstach (nie tylko pozycji konkurencyjnej). Stosowana jako wprowadzenie do studiów wykonalności. Najczęściej stosowana w wariancie PEST, tutaj z uwzględnieniem osobno czynników prawnych od politycznych (na poziomie strategii rządowych Polski i Unii Europejskiej).

- **Kluczowe czynniki sukcesu** (*Critical Success Factors*) - elementy działania firmy/organizacji niezbędne dla odniesienia przez nią sukcesu, na podstawie których często ustala się poziomu sukcesu i cele np. KPI.
- **Metoda delficka** - metoda heurystyczna (formująca myślenie kreatywne) polegająca na wykorzystywaniu wiedzy, doświadczenia i opinii ekspertów w seriach pytań badawczych, gdzie w kolejnych rundach wyniki poprzedniej traktowane są jako dane wejściowe. Dzięki temu następuje pętla informacji zwrotnej korygująca odchylenia od głównej prognozy.

W Raporcie wyodrębniamy korzyści płynące z współdzielenia danych, a także bariery wynikające z analizy SLEPT, które potencjalnie hamować mogą proces *data-sharing* w danym modelu. Zaproponowane zostały także konkretne rozwiązania oraz czynniki, które według badanych są kluczowe dla stworzenia oraz prawidłowego funkcjonowania bezpiecznych przestrzeni wymiany danych.

Publikacja powstała na kanwie materiału zgromadzonego podczas warsztatów, dlatego znajdziemy w niej części poświęcone każdej z przestrzeni danych omówionej podczas cyklu. Jednak pomimo kompleksowego podejścia do kwestii współdzielenia danych, uwaga autorów skupiła się przede wszystkim na **wspólnicy dla sektora rolniczego** oraz **wspólnicy danych zdrowotnych**. Z warsztatów wyniknęło bowiem, że to właśnie te przestrzenie mają największą szansę powstać w polskiej przestrzeni publicznej,przyczyniając się do poprawy jakości życia wszystkich obywateli oraz zmniejszania długu innowacyjnego.

4.1 Pośrednicy danych osobowych

Pośrednicy danych osobowych stanowią obiecującą koncepcję umożliwiającą wykorzystanie danych przy jednoczesnym poszanowania zasad dotyczących prywatności. Instytucje te mogą być tworzone dla celów, takich jak chociażby skuteczniejsze wdrażanie ochrony danych osobowych bądź efektywniejsze zachęcanie do udostępniania danych w całym łańcuchu wartości. Dzięki podejmowaniu działań w zgodzie z interesem internautów oraz stawianiu ich potrzeb na pierwszym miejscu, pośrednicy mają szansę stać się modelem alternatywnym dla narzucanego przez największe platformy cyfrowe. O ile bowiem gigantom cyfrowym zarzuca się gromadzenie ogromnej ilości danych wykorzystywanych przede wszystkim do celów komercyjnych, pośrednicy danych działaliby przede wszystkim z myślą o dobru swoich użytkowników.

Głównym założeniem tego modelu jest zapewnianie bezpiecznego i godnego zaufania środowiska, w którym osoby prawne lub osoby fizyczne (posiadacze danych) mogą udostępniać swoje dane. Do zadań pośredników danych osobowych należy także udzielanie wsparcia i pomocy posiadaczom danych w dokonywaniu świadomych wyborów w zakresie wyrażania zgody na przetwarzanie ich danych, umożliwiając jednostkom dobrowolne gromadzenie danych dla obopólnej korzyści. Wprowadzenie instytucji niezależnego pośrednika między osobami, których dane dotyczą a podmiotami zbierającymi dane, pozwala także skuteczniej negocjować warunki wykorzystania danych zgodnie z wymogami bezpiecznego środowiska ich wymiany. Dzieje się tak ze względu na większą siłę przetargową wynikającą z agregacji danych w rękach jednego podmiotu (Data Trust Initiative, 2021). Jak bowiem wiadomo, dane zyskują na wartości dopiero w masie. O ile pojedyncze wycofanie zgody na ich przetwarzanie nie wpłynie negatywnie na model biznesowy platformy, utrata większej ilości rekordów byłaby dla serwisu realnym zagrożeniem (Paszcza, 2022).

Potrzeba przekazania użytkownikom większej kontroli w zakresie zasobów danych ich dotyczących została dostrzeżona przez unijnego regulatora. Jednak przewidziane w unijnych aktach prawnych (Data Governance Act; Data Act) reguły uwzględniają małą elastyczność w zakresie swobody działalności pośredników. Jak wynika z projektu rozporządzenia w sprawie europejskiego zarządzania danymi, dostawcy usług udostępniania danych mają za zadanie działać jedynie jako pośrednicy w transakcjach i nie mogą wykorzystywać udostępnianych danych do żadnych innych celów. Tym samym, brak możliwości czerpania zysków z zarządzania danymi stawia pod znakiem zapytania występowanie pośredników w realnym świecie. Atrakcyjny model biznesowy jest bowiem kluczowy dla pojawienia się na rynku nowych podmiotów oraz wykształcenia konkurencyjnych, wysokiej jakości usług.

Dla jakich rodzajów danych?

- dane handlowe
- dane płatnicze
- dane z inteligentnych urządzeń
- dane o lokalizacji
- dane społecznościowe
- adresy IP

Jakie formy?

Data trusts - instytucje, które w imieniu danej osoby zarządzają jej danymi na wzór obecnego w prawie anglosaskim “funduszu powierniczego”. Relacje ustanawiane są na podstawie powtarzalnych ram występujących na gruncie prawa kontraktowego (Mehta, Dawande i Mu, 2022). Podmiot powierzający przekazuje dane powiernikowi, które następnie mogą być wykorzystywane przez osobę trzecią, czyli beneficjenta. Data trusts dzielą się na te, które przechowują dane i te, które zarządzają indywidualnymi i zbiorowymi prawami dostępu do danych. Model ten można porównać do bibliotek lub zbiorów umożliwiających cyfrowy dostęp do treści, które mają służyć określonej społeczności i chronić zasoby przed nieuprawnionym dostępem (Artyushina, 2021). Pod względem odpowiedzialności za dane koncepcja ta porównywana jest także do występujących w rzeczywistym świecie profesji obciążonych tajemnicą zawodową (np. prawniczych, lekarskich). Podmioty zarządzające danymi zyskują dostęp do osobistych, potencjalnie wrażliwych informacji, ale równocześnie są prawnie zobowiązane do działania w najlepszym interesie beneficjentów ich usług (Artyushina, 2021). Widocznym jest więc, że w kontekście zaufania obowiązek powierniczy wiąże się z dużym stopniem bezstronności, rozwagi, przejrzystości i lojalności (Delacroix i Lawrence, 2019). Jeżeli w przypadku *data trustu* stosowane jest prawo powiernicze, powiernik jest zobligowany prawnie do lojalności i staranności względem beneficjenta. Pozostawanie w zgodzie z tymi zobowiązaniami wymaga natomiast, aby *trust* danych był niezależny, co może uniemożliwić występowanie tej instytucji w formie firmy nastawionej na zysk (Mehta, Dawande i Mu, 2022).

Przykładem takiej formy współdzielenia danych jest organizacją non-profit PlaceFund, która działa jako trust danych geoprzestrzennych, promujący wykorzystanie danych, w celu prostowania kwestii związanych z prawami własności do ziemi, niezrównoważonym użytkowaniem gruntów i zmianami klimatu. Podstawą ambicji PlaceFund jest stworzenie zaufanej przestrzeni dla danych geoprzestrzennych, które będą przetwarzane w sposób zrównoważony, a następnie będą udostępniane społecznościom lokalnym.

Spółdzielnie danych (data cooperatives) - propozycja instytucji opisana szczególnie w Data Governance Act jako jedna z usług współdzielenia danych. Zapewniając nadzór i przejrzystość, spółdzielnie danych mają umożliwić osobom fizycznym skuteczniejsze korzystanie z praw przysługujących im na mocy RODO. Dzięki ich usługom możliwe stałoby się wzmocnienie pozycji osób fizycznych w relacjach z platformami oraz wspieranie użytkowników w dokonywaniu świadomych wyborów w zakresie wyrażania zgody na wykorzystywanie ich danych (Komisja Europejska, 2021). Instytucje te miałyby również za zadanie ulepszać warunki oferowane osobom, których dane dotyczą oraz rozwiązywać spory dotyczące kilku osób, których dane dotyczą w ramach grupy (Bayamlioglu, 2021).

Do przykładów istniejących spółdzielni możemy zaliczyć Driver's Seat, czyli spółdzielnię, która agreguje dane związane z pracą smartfonów kierowców aktywnych w obszarze gig-economy. Swash natomiast gromadzi dane osób surfujących po sieci oraz pozwala użytkownikom Internetu, programistom i firmom na wykorzystywanie danych do tworzenia nowych wartości dzięki innowacyjnym mechanizmom monetyzacji danych i spółdzielczym ramom rozwoju.

Wykorzystywanie podmiotów świadczących usługi pośrednictwa danych może potencjalnie zaadresować różne kwestie społeczno-ekonomiczne, a ponadto poprawić pozycję jednostki w relacji z podmiotami cyfrowymi. Przede wszystkim, agregując pojedyncze dane, instytucja taka, jak spółdzielnia danych wzmacnia swoją siłę przetargową i tym samym może uzyskać korzystniejsze warunki zbierania danych (Mehta, Dawande i Mu, 2022). Jednak poza korzyściami płynącymi z powierzania cyfrowych zasobów spółdzielniom, czy szerzej - pośrednikom danych osobowych, zaobserwować można także bariery i wątpliwości związane z omawianym modelem.

Korzyści

- Sprawowanie kontroli nad danymi przez użytkownika

Założeniem korzystania z usług pośredników danych osobowych jest zwiększenie indywidualnej sprawczości i decyzyjności osób fizycznych w zakresie tego, co dzieje się z ich “śladami cyfrowymi”. Podmioty, których dane dotyczą mogą mieć także kontrolę nad jakością oraz ilością danych, którymi się dzielą (Mehta, Dawande i Mookerjee 2021).

- Wysoki wskaźnik bezpieczeństwa danych

Nadrzędnym celem pośrednictwa świadczonego przez spółdzielnię jest zapewnianie bezpiecznego i godnego zaufania środowiska, w którym osoby prawne lub osoby fizyczne (posiadacze danych) będą mogły udostępniać swoje dane. Data trusty charakteryzują się natomiast wysokim wskaźnikiem bezpieczeństwa ze względu na zobowiązania wynikające z ich “powierniczego” charakteru.

- Większa siła przetargowa

Dane osobowe jednostki nie mają same w sobie dużej wartości (Pentland i Hardjono, 2020). Z tego względu siła negocjacyjna pojedynczego użytkownika jest niewielka, co bywa wykorzystywane przez platformy stosujące politykę ”*take it or leave it”.* Internauci, nie dysponując możliwością ingerencji w regulaminy platformy, często zmuszeni są zaakceptować warunki stawiane przez stronę, choć nie zawsze są one dla nich

korzystne. Wprowadzenie instytucji pośrednika danych osobowych jest szansą na zmianę tego paradygmatu. Przyjmując, że dane zyskują na wartości w masie, można spodziewać się, że *trusty* bądź spółdzielnie dysponujące większymi zasobami danych będą w lepszej pozycji do tego, by dyktować warunki platformom oraz żądać od nich bardziej zrównoważonego przetwarzania danych.

- Odciążenie posiadaczy danych

Zamiast angażować się w relacje z poszczególnymi jednostkami, użytkownicy danych mogą zawrzeć pojedynczą umowę ze spółdzielnią, która reguluje dostęp i warunki korzystania z danych (Mehta, Dawande i Mookerjee, 2021). *Data trust* zwalnia natomiast podmiot z konieczności podejmowania najważniejszych decyzji w zakresie jego danych osobowych, równocześnie zapewniając, że wszelkie operacje pozostawać będą w zgodzie z wymogami prywatności i bezpieczeństwa.

Bariery

- Niejasność modelu biznesowego
- Brak świadomości w zakresie korzyści płynących z dzielenia się danymi
- Niejednolite standardy
- Brak odpowiedniej rządowej strategii w zakresie dzielenia się danymi

Proponowane rozwiązania

1) Opracowanie atrakcyjnego modelu biznesowego

Jak już zostało wyżej wspomniane, zgodnie z założeniami unijnego projektu Rozporządzenia ws. Zarządzania Danymi (Data Governance Act), dostawcy usług nie powinni wykorzystywać udostępnionych danych do innych celów niż samo pośrednictwo. Nie mogą więc oni czerpać zysków z danych, na przykład sprzedając je innym podmiotom (Rada UE i Rada Europejska, 2021). Co więcej, model biznesowy pośredników gwarantować ma brak niewłaściwych zachęt dla osób fizycznych by udostępniać do przetwarzania większą ilość danych, niż to leży w ich własnym interesie (Komisja Europejska, 2020b). Tym samym, pośrednicy danych osobowych mają ograniczone możliwości monetyzacji informacji cyfrowych im przekazanych. Zwrócili na to uwagę uczestnicy warsztatów. Jak zostało wskazane, **koncept, w którym niejasnym jest w jaki sposób instytucje pośrednictwa danych mogłyby na siebie zarabiać stawia pod znakiem zapytania efektywność tego modelu współdzielenia danych.**

Pośrednicy danych w praktyce mogę występować w różnych formach, począwszy od podmiotów prywatnych, a skończywszy na organizacjach non-profit i instytucjach publicznych. Ich struktura, motywacja i zarządzanie uzależnione jest od odmiennych uwarunkowań i celów zapisanych w statucie organizacji bądź strategii firmy. Z tego względu model biznesowy dostawcy usług powinien być dopasowany do konkretnego przypadku. W przypadku podmiotów działających w formie non-profit dobrym pomysłem byłoby uruchomienie programu małych grantów dla pośredników danych (prowadzących działalność pożytku publicznego oraz spółdzielni) na budowę podstawowej infrastruktury współdzielenia danych, utrzymanie jej prawidłowego funkcjonowania oraz tworzenie zasad korzystania z usług serwisu. Podmioty prywatne powinny mieć natomiast możliwość uzyskiwania korzyści finansowych za oferowane usługi pośrednictwa, chociażby poprzez

pobieranie opłat za dołączenie do zbudowanego przez nie ekosystemu danych (Janssen i Singh, 2022).

2) **Zwiększanie świadomości w zakresie dzielenia się danymi**

Uczestnicy warsztatów wskazali na niedostateczną wiedzę społeczeństwa w zakresie korzyści płynących z dzielenia się danymi. Świadomość zalet przekazywania danych niezależnym, sprawdzonym pośrednikom jest natomiast kluczowa, aby przezwyciężyć lęki oraz niechęć użytkowników do nowych instytucji. Dlatego koniecznym jest położenie większego nacisku na wysokiej jakości edukację w obszarze cyfryzacji (szkoły podstawowe, licea ogólnokształcące), jak również działania propagujące spółdzielczy charakter danych (np. kampanie promujące; dofinansowania na szkolenia pracowników w firmach).

3) **Wypracowanie jednolitego standardu wymiany danych**

Problemem w zakresie efektywnego współdzielenia danych, który podczas warsztatów podnoszony był kilkukrotnie, jest brak jednolitych standardów wymiany danych. Z tego względu dobrym pomysłem byłoby wyodrębnienie z administracji rządowej podmiotu, który mógłby zajmować się wyznaczaniem norm i wymagań oraz certyfikacją podmiotów zamierzających pełnić usługi pośrednictwa pod kątem zgodności ich działania z odgórnie ustalonymi standardami.

4) **Zaplanowanie efektywnej strategii**

O ile w rankingu otwartości danych Polska zajmuje czwarte miejsce na tle całej Unii Europejskiej (Van Hesteren et al, 2021), o tyle w przypadku indeksu gospodarki cyfrowej i społeczeństwa cyfrowego (DESI) na 2021 rok, Polska plasuje się dopiero na 24. miejscu wśród 27 państw członkowskich (Komisja Europejska, 2021). Niski wskaźnik “ucyfrowienia” naszego kraju przekłada się także na utrudnienia we wdrażaniu takich koncepcji jak współdzielenie danych. Wskazane byłoby więc przyjrzenie się strategiom dotyczącym otwierania danych publicznych w celu zaczerpnięcia inspiracji oraz zidentyfikowania dobrych praktyk, które mogłyby znaleźć zastosowanie także w przypadku uwspólniania danych osobowych. Mowa tutaj o zwróceniu uwagi na środki (kapitałowe, ludzkie, organizacyjne, informacyjne) jakie zostały wykorzystane przy otwieraniu danych publicznych, a następnie na wykorzystaniu tej wiedzy do nakreślenia kierunków działań w zakresie budowania instytucji współdzielenia danych.

4.2 Wspólnice danych biznesowych (tzw. wirtualne składnice)

Szacuje się, że globalna wartość sztucznej inteligencji w 2020 r. miała wynieść około 2,65 biliona dolarów, a między samym rokiem 2018 i 2019, liczba organizacji, które wdrożyły AI, wzrosła z 4% do 14% (Goasduff, 2019).Powodem, dla którego wykorzystywanie urządzeń *business intelligence* przez przedsiębiorców staje się coraz bardziej powszechne jest fakt, że wykładniczy postęp jaki dokonuje się w infrastrukturze obliczeniowej daje szereg możliwości do poprawy efektywności firmy, obniżenia kosztów, ulepszania produktów. Budowanie dokładnych modeli predykcyjnych oraz tworzenie sztucznej inteligencji za pomocą uczenia maszynowego wymaga jednak dostępu do ogromnych zbiorów danych. O ile w przypadku dużych koncernów zwykle nie stanowi to problemu, firmy z sektora MŚP, start-upy, lokalni producenci, rolnicy w dalszym ciągu pozostają w tyle, popadając tym samym w **dług innowacyjny**. Z tego względu, tak ważne jest stworzenie warunków do

współdzielenia danych pomiędzy wyżej wymienionymi podmiotami. Jednym z rozwiązań może być stworzenie omówionej przez nas podczas warsztatów badawczych przestrzeni dla danych biznesowych.

Wspólnica (wspólna składnica) **danych biznesowych** przewiduje formę współpracy, w ramach której aktorzy gospodarczy wymieniają się swoimi danymi w celu tworzenia zbiorowej wartości (Data Collaboratives, 2021). Przyjęta w publikacji **nowa nazwa** dla wirtualnej składnicy danych wynika z dostrzeżenia przez organizatorów warsztatów utraty relewantności podziału na dane osobowe i nieosobowe, zaś **konieczności skupienia się na ich funkcjach i celach wykorzystania.** Hasło "wspólnica" ma za zadanie podkreślać spółdzielczy charakter tego modelu. Natomiast odwołanie do "danych biznesowych" służy wskazaniu szerokiego zakresu danych oraz możliwości administrowania nimi (zarówno przez podmioty prywatne, jak i operatorów publicznych).

Głównym założeniem modelu jest wykorzystywanie pozytywnych efektów sieciowych, a więc zachęcanie przedsiębiorstw (dużych firm, MŚP, start-upów) do nawiązywania partnerstw i tworzenia wspólnie bardziej złożonych produktów oraz powiązań. Chodzi także o wykreowanie pewnego wzorca wymiany danych, opartego na swobodnym przepływie informacji cyfrowych w ramach federacji zainteresowanych podmiotów. Komunikacja i współpraca mogłaby odbywać się za pośrednictwem połączonych platform posiadających jednolity standard dla współdzielonych danych. Tak funkcjonujący system znajdziemy chociażby w przypadku istniejącej już formy współpracy pomiędzy przedsiębiorcami - International Data Spaces (wcześniej Industrial Data Space (IDS). Innym ciekawym przykładem formy uwspólniania danych rolniczych jest Ethiopian Commodity Exchange - giełda, która zrzesza rolników i dostarcza im dane potrzebne do podnoszenia jakości płodów rolnych oraz optymalizacji upraw (Verhlust, Young i Srinivasan 2022).

Dla jakich rodzajów danych?

- dane przemysłowe (dane wytwarzane przez maszyny; związane z konserwacją maszyn; dane o łańcuchach dostaw)
- dane biznesowe (dane zawierające liczbę wizyt oraz czas spędzony na konkretnej stronie internetowej; analizy big data; dane o logistyce)
- dane rolnicze (dane z czujników; dane z ciągników; dane meteorologiczne)
- dane finansowe

Jakie formy?

Pośrednik Zaufany - firmy z sektora prywatnego udostępniają dane partnerom z sektora publicznego, obywatelskiego i akademickiego. Dane mogą być również przekazywane za pośrednictwem zewnętrznych firm analitycznych, które udostępniają dane podmiotom zarządzającym danymi w celu przeprowadzenia analizy i udostępnienia jej wyników podmiotom zewnętrznym, co pozwala na uzyskanie wyników z danych bez ujawniania informacji wrażliwych (Verhlust, Young i Srinivasan, 2022). Wymiana danych za pośrednictwem zaufanego podmiotu mogłaby odbywać się na zamkniętej platformie, utworzonej przez niezależnego pośrednika (Komisja Europejska, 2018). Zaletą omawianej formy współdzielenia byłaby możliwość obniżenia kosztów przyszłych transakcji oraz ujednolicenia sposobu wymiany danych pomiędzy przedsiębiorcami. Standardy na rynku mogą być bowiem narzucane nie tylko przez organy regulacyjne, lecz także przez pośredników przedsiębiorczych występujących pod postacią firm bądź międzynarodowych organizacji (Baron et al, 2019)

Łączenie danych (Data Pooling) - Interesariusze z różnych sektorów przystępują do "puli danych", aby dzielić się zasobami. Publiczne pule danych pozwalają partnerom na otwarty dostęp do danych i ich niezależne wykorzystanie, natomiast prywatne pule danych ograniczają dostęp do informacji i ich wkład. Rozwiązanie to zbliżone jest do znanego prawu europejskiemu “patent pooling”, czyli modelu, w którym wynalazcy udostępniają swoje patenty za określoną z góry cenę, a dany projekt może zostać nabyty przez każdy zainteresowany podmiot. Łączenie danych mogłoby przybrać formę platformy dla danych przemysłowych. Byłoby to podejście oparte na współpracy ograniczonej grupy firm oraz budowaniu przez nie wspólnej strategii. Wiązałoby się ono z dobrowolnym dołączaniem przedsiębiorców do zamkniętych, bezpiecznych i cyfrowych środowisk oraz przekazywaniem danych w celu wspierania rozwoju nowych produktów/usług. Dane mogłyby być udostępniane odpłatnie bądź nieodpłatnie (The Ministry of Electronics & Information Technology, Government of India, 2020). *Data pooling* występuje już oraz ma szczególne znaczenie w rolnictwie. Dzięki łączeniu danych dotyczących pól uprawnych, maszyn, pogody możliwe jest wykorzystywanie narzędzi tzw. smart-farming, co w dalszej kolejności przekłada się na bardziej efektywne wykorzystanie zasobów oraz na wyższe plony (Schubert i Harari Dayan, 2020).

***Na podstawie warsztatu przeprowadzonego z przedstawicielami sektora rolnictwa**

Agencja sprawująca nadzór nad wspólnicą - propozycja zaangażowania organizacji działającej w sektorze rolniczym w tworzenie oraz administrowanie wspólnicą danych. Takim podmiotem mógłby być Krajowy Ośrodek Wsparcia Rolnictwa, który jako państwowa agencja wykonawcza znajduje się “blisko państwa” (co mogłoby ułatwić współpracę oraz nadzór nad wdrażanym projektem), z drugiej zaś strony (dzięki terenowym oddziałom, po jednym w każdym województwie) KOWR byłby w stanie utrzymywać bieżący kontakt z rolnikami oraz promować wdrażanie pewnych rozwiązań na szczeblu lokalnym.

Współpraca stron - forma polegająca na tworzeniu powiązań pomiędzy partnerami biznesowymi poprzez łączenie posiadanych przez nich zasobów. Mogłaby ona polegać na zawiązywaniu spółdzielni - dobrowolnych zrzeszeń nieograniczonej liczby podmiotów, które w interesie swoich członków prowadzą wspólną działalność na rzecz wymiany danych pochodzących z rolnictwa. Innym wskazanym przez uczestników warsztatów rozwiązaniem dla wspólnicy danych rolniczych było zobowiązanie się do współpracy stron na podstawie umowy partnerskiej. Mogłaby ona przybierać formę długoterminowej umowy zawartej pomiędzy zainteresowanymi podmiotami, której celem byłoby stworzenie składników infrastruktury umożliwiającej świadczenie usług współdzielenia danych oraz wspólne czerpanie z zysków.

Korzyści

- Wyrównywanie szans

Dostęp do wspólnic danych rolniczych wzmocni pozycję rolników indywidualnych i przedsiębiorstw rolno-spożywczych, które w większości przynależą do sektora MŚP i w obrębie swojej organizacji, nie dysponują dostępem do szerokich zasobów cyfrowych i

narzędzi analitycznych. Przedsiębiorstwa te powinny mieć jednak pewność, że mogą faktycznie skorzystać na udostępnianiu i wymianie danych, unikając zagrożeń ze strony największych firm na rynku (Nagel i Lycklama, 2021).

- Zwiększona dostępność zróżnicowanych danych do uczenia maszynowego

Zaawansowane aplikacje odgrywają fundamentalną rolę w procesach biznesowych oraz krytycznych gałęziach przemysłu, m.in. w rolnictwie. Dostępność cennych zbiorów danych ma kluczowe znaczenie dla tworzenia wartości dodanej za pośrednictwem sztucznej inteligencji w takich obszarach rolnictwa, jak. Sfederowane analizy rozproszonych danych dają natomiast możliwość uwspólniania uzyskanych wyników bez konieczności udostępniania oryginalnych danych. Możliwe jest więc zapewnienie równowagi pomiędzy prywatnością, autonomią, ochroną własności intelektualnej (Big Data Value Association, 2019)

- Budowanie nowych modeli biznesowych opartych na danych

Najbardziej innowacyjne modele biznesowe oparte na danych wykazują szeroką gamę możliwości budowania wartości ekonomicznej - od bezpośredniej monetyzacji danych po budowanie usług opartych na dostępie do platformy na podstawie subskrypcji. W kontekście danych rolniczych.

- Wzrost skali produktywności dla całej gospodarki oraz poprawa konkurencji

Pomimo możliwości wielokrotnego wykorzystania danych, ze względu na ich duże znaczenie w gospodarce cyfrowej do tej pory pozostawały one strzeżone prawami własności intelektualnej i tajemnicami handlowymi największych platform. Obowiązek współdzielenia danych wprowadzony dla najbardziej newralgicznych społecznie sektorów podniesie efektywność produkcji wśród mniejszych przedsiębiorstw, które do tej pory nie mogły konkurować z największymi dostawcami, posiadającymi ogromne zasoby danych zamknięte w prywatnych silosach. By przywrócić konkurencyjność, konieczne jest obniżenie kosztów dostępu do danych dla tych, którzy dotychczas nie mieli takiej możliwości.

- Poprawa efektywności wykorzystania dostępnych zasobów, większa produktywność i zrównoważone ekologicznie rozwiązania

Narzędzia sztucznej inteligencji dla rolnictwa oparte na dużych zbiorach danych pochodzących z czujników i sensorów pozwalają na wykorzystywanie rozwiązań sprzyjających bardziej zrównoważonemu rolnictwu. Jak wskazują eksperci, dzięki tzw. *smart-farming* do roku 2050 rolnicy mogliby zwiększyć produkcję żywności aż o 70% przy równoczesnym obniżeniu kosztów produkcji oraz ograniczeniu eksploatacji środowiska naturalnego (Nayyar, A., Puri, V., 2016).

Bariery

- Obawy względem dzielenia się danymi
- Niechęć do wprowadzania zmian (powszechna szczególnie wśród starszego pokolenia rolników)
- Nieufność w stosunku do nowych technologii
- Internetowe białe plamy
- Brak interoperacyjności silnej systemów i standardów wymiany danych

- Obawy o utratę konkurencyjności; powoływanie się na tajemnicę przedsiębiorstwa

Proponowane rozwiązania

1) **Podnoszenie umiejętności cyfrowych oraz wyposażenie rolników w narzędzia analityczne**

Wraz z budowaniem infrastruktury wspólnicy ważnym jest przekazywane rolnikom umiejętności cyfrowych, by faktycznie mogli oni osiągać korzyści z dostępu do informacji cyfrowych. Dane pozostają bowiem bezużyteczne, jeśli nie są poprawnie interpretowane i przekształcane w znaczące decyzje biznesowe. Dlatego aby czerpać z gromadzonych danych koniecznym jest nie tylko zapewnienie dostępu do narzędzi służących ich analizie, ale także wyposażenie rolników w odpowiednie kompetencje do obsługi tych urządzeń (Nagel i Lycklama, 2021).

2) **Oferowanie obietnicy uczestnictwa w zysku**

Uczestnicy warsztatu wskazali, iż rolnicy rzadko kiedy widzą korzyści płynące z otwierania danych zebranych w obrębie ich przedsiębiorstwa. Z tego względu trafionym pomysłem wydaje się być tworzenie pozytywnych zachęt dla współdzielenia danych, np. poprzez oferowanie obietnicy w zyskach. Jako przykład podane zostały próbki produktów stworzonych na podstawie danych dostępnych we wspólnicy.

3) **Ujednolicenie standardów oraz zapewnienie interoperacyjności silnej**

Jedna z najważniejszych barier dotyczy ograniczonej interoperacyjności i braku otwartości systemów technicznych w urządzeniach rolniczych, co utrudnia rolnikom wybór dostawców nowych technologii (Komisja Europejska, 2020c). Zostało to również zauważone podczas warsztatów. Uczestnicy wskazali, że różni producenci urządzeń rolniczych przewidują niejednakowe standardy dla zbierania danych. Co więcej, często systemy oraz bazy danych nie są ze sobą kompatybilne. Zwiększona interoperacyjność silna pozwoliłaby natomiast na intensywniejsze udostępnianie danych i generowanie wynikającej z tego wiedzy.

4) Angażowanie lokalnej społeczności w tworzenie wspólnic

Według uczestników warsztatu, powodzenie we wdrażaniu nowych koncepcji na obszarach wiejskich jest ściśle skorelowane z tym, jak proponowane rozwiązania odbierane są przez społeczność lokalną. W przypadku wspólnic dla danych rolniczych warto byłoby więc zaangażować w ich tworzenie różne podmioty aktywne na wsiach - nie tylko potencjalnych beneficjatnów wspólnicy (rolników), ale także społeczno-zawodowe organizacje (np. kółka rolnicze; koła gospodyń wiejskich; rolnicze zrzeszenia branżowe) oraz przedstawicieli lokalnej wiary.

5) Tworzenie programów pilotażowych

Jak zostało zauważone podczas warsztatów, wdrażanie innowacyjnych rozwiązań powinno być poprzedzone programami pilotażowymi. Stanowiska testowe pozwoliłyby byłoby na sprawdzanie i porównywanie nowych technologii. Dopiero po wypracowaniu sprawdzonych narzędzi wskazane byłoby ich wdrażanie na większą skalę.

4.3 Publiczne wspólnice danych

Model wspólnic danych osobowych stanowi alternatywę dla korporacyjnych silosów danych, zapewniając uczciwą kooperację między zaangażowanymi podmiotami i świadome uspołecznianie wytwarzanej wartości. Z natury rzeczy dane cyfrowe są bowiem dobrem wspólnym - nie podlegają wyczerpaniu i mogą być dowolnie replikowane. Głównym celem wspólnicy jest stworzenie ekosystemu zaufania dla prospołecznego wykorzystania danych w oparciu o zasady współzarządzania i ustaloną hierarchię wartości. To instytucja publiczna, która w swoim założeniu odchodzi od podziału na twórców wartości/ekstraktujących wartość na rzecz stworzenia banków rozwoju czy środków służących do integracji wiedzy o gospodarce informacyjnej (Zygmuntowski, 2020b).

Z wartości płynącej z agregacji tych danych korzysta przede wszystkim określona społeczność - lokalna, regionalna czy też paneuropejska. Ważnym aspektem funkcjonowania wspólnic jest model współzarządczy i inkluzywność - dzięki możliwości uzyskania dostępu do danych, które wcześniej były w posiadaniu jedynie największych firm, poza zwykłymi obywatelami zarówno małe i średnie przedsiębiorstwa, jak i administracja publiczna będą miały możliwość ulepszenia swoich wewnętrznych procesów dzięki bardziej dogłębnej analizie danych, przy jednoczesnym dbaniu o ochronę interesu publicznego.

Dla jakich rodzajów danych?

- dane osobowe o znaczeniu publicznym (np. dane zdrowotne)
- dane wytwarzane w systemie usług publicznych
- dane społecznościowe (np. z mediów społecznościowych)
- dane administracyjne
- dane samorządowe

Jakie formy?

Współzarządzane rejestry publiczno-prywatne - w tym modelu wspólnica danych to zaufana przestrzeń wymiany danych utworzona w interesie publicznym, chroniąca europejskie wartości i prawa dzięki aktywnemu uczestnictwu społeczności w zarządzaniu danymi. Model współzarządzania jest w tym zakresie ważny ze względu na brak wystarczającej siły negocjacyjnej pojedynczego podmiotu danych wobec znacznie większych dostawców platform cyfrowych (Zygmuntowski, 2020a). Wprowadzenie jednej, silnej instytucji stojącej na straży interesów jednostek i właściwego wykorzystywania ich danych pochodzących z różnych źródeł wzmacnia pozycję osób, których dane dotyczą wobec podmiotów chcących te dane wykorzystać.

Opis struktury:

Podstawową zasadą wspólnic danych osobowych powinna być transparentność zamiarów podmiotu chcącego wykorzystać zebrane dane do własnych badań. Ich wyniki powinny zostać udostępnione publicznie, a ochrona interesów jednostek wymaga, by każdorazowo przeprowadzać ocenę skutku algorytmu. Podmioty, które już na początkowym etapie nie spełniałyby ustanowionych zasad etycznych nie uzyskiwałyby dostępu. Co również ważne, aby zapewnić odpowiedni poziom zabezpieczeń danych, postuluje się by zgodnie z koncepcją “move algorithm to data” to podmioty zewnętrzne przekazywały swój algorytm do baz wspólnicy, unikając transferowania danych poza zaufaną i bezpieczną infrastrukturę (Hardjono i Pentland, 2019). Algorytm ten, po przesłaniu przez odpowiedni interfejs,

prowadziłby obliczenia bezpośrednio na danych zgromadzonych w bazie wspólnicy, uzyskując później same wyniki (Zygmuntowski, 2020a).

Z uwagi na osobowy charakter danych gromadzonych we wspólnicach, konieczne jest zapewnienie odpowiedniego poziomu ochrony prywatności i bezpieczeństwa danych. Istnieje wiele metod technicznych zapewniających odpowiedni poziom poufności dla danych, m.in. szyfrowanie homomorficzne polegające na przeprowadzaniu obliczeń na zaszyfrowanej treści czy prywatność różnicowa, pozwalająca na przechowywanie informacji o grupach w zestawie danych w taki sposób, by nie było konieczne ujawnianie danych pojedynczej osoby (Zygmuntowski, 2020a). Pod względem bezpieczeństwa, ważnym aspektem jest również wybór odpowiedniego dostawcy infrastruktury chmurowej. Firmyposiadające siedzibę w USA podlegają tamtejszym przepisom o nadzorze, zezwalającym agendom rządowym (np. FBI) na praktycznie nieograniczony dostęp do danych nie-obywateli USA w pewnych określonych przypadkach (Konarski, 2020). Aby uniknąć możliwości podlegania pod te regulacje, rekomendowane jest wybieranie europejskich usługodawców, których zarówno siedziby, jak i centra danych zlokalizowane są na terenie EOG.

Utrzymanie złożonej architektury technicznej wymaga odpowiednich środków finansowych. Ze względu na neutralny charakter przedsięwzięcia, wybrany model biznesowy powinien być w miarę możliwości samowystarczalny. Wspólnica może utrzymywać się przede wszystkim z opłat licencyjnych uiszczanych za umożliwienie dostępu do danych przez API. Powinna również wprowadzać kryteria rozróżniania wysokości opłat, które byłyby niższe dla podmiotów działających niekomercyjnie, prowadzących badania naukowe i inicjatywy społeczne i odpowiednio wyższe dla tych, którzy wykorzystują dane dla rozwoju własnego biznesu (Zygmuntowski, 2020a). Z uwagi na prospołeczny charakter wspólnic, pozyskiwanie środków z grantów organizowanych przez publiczne instytucje czy organizacje pozarządowe również wydaje się być dobrym rozwiązaniem.

Korzyści:

- Podnoszenie zaufania do współdzielenia danych

Zgodnie z raportem Polskiego Instytutu Ekonomicznego, co do zasady Polacy nie są chętni do dzielenia się swoimi danymi. Mniej niż połowa respondentów (45,2%) byłaby gotowa udostępnić dane o swoich nawykach zdrowotnych na potrzeby publicznego programu profilaktycznego (Grzeszak et al., 2020). Taki stan rzeczy wynika przede wszystkim z obawy przed przetwarzaniem danych w złej wierze i możliwością wykorzystania wyników przetwarzania przeciwko osobie, której dane dotyczą. Związany z tym strach przed inwigilacją organów publicznych w sposób szczególny ujawnił się w kontekście reakcji opinii publicznej na projekt rejestru ciąż. Zgodnie z nowelizacją rozporządzenia Ministra Zdrowia z dnia 26 czerwca 2020 roku w sprawie szczegółowego zakresu danych zdarzenia medycznego przetwarzanego w systemie informacji oraz sposobu i terminów przekazywania tych danych do Systemu Informacji Medycznej, wszystkie podmioty świadczące usługi medyczne będą zobowiązane do przekazywania danych o pacjentkach w ciąży do systemu informacji medycznej. Chociaż powody zmian były słuszne (np. przepisywanie leków kobietom w ciąży, skorzystanie z zapisów do lekarza poza kolejnością, ratowanie życia), ze względu na kontekst polityczny związany z zachodzeniem w ciążę w Polsce wiele osób domniemywało, że w rzeczywistości regulacja nastawiona jest na sprawowanie kontroli nad obywatelkami (PAP, 2021).

Tak nieprzychylna reakcja na pomysł udostępniania informacji wynika z braku wcześniejszej praktyki współdzielenia danych i silnego przekonania o możliwej utracie prywatności w sieci. Zaangażowanie wszystkich zainteresowanych stron w proces zarządzania wspólnicami danych mogłoby stopniowo przywrócić zaufanie do inicjatyw o charakterze publicznym, nastawionych na kreowanie Wspólnej Wartości Społecznej.

- Poprawa jakości predykcji systemów AI - im więcej danych tym lepiej; pozytywne efekty zewnętrzne z agregacji danych

Zgodnie z założeniami "Polityki rozwoju sztucznej inteligencji w Polsce" szacuje się, że rozwój sztucznej inteligencji w Polsce poprawi dynamikę PKB o nawet 2,65 pp w skali każdego roku. W ciągu najbliższych ośmiu lat, AI pozwoli na zautomatyzowanie ok. 49% czasu pracy w Polsce, generując jednocześnie lepiej wynagradzane miejsca pracy w sektorach strategicznych (GovTech Polska, 2020). Aby jednak móc właściwie działać w obszarach zidentyfikowanych jako najistotniejsze, kluczowym czynnikiem rozwoju sztucznej inteligencji jest dostarczenie jej jak największej ilości jakościowych danych, które będą kompletne, odpowiednio opisane i etykietowane (Kawalec, 2021). Im więcej informacji na temat określonego zjawiska dostarczamy algorytmowi, tym lepsza będzie wygenerowana przez system predykcja. Instytucja wspólnic danych, ze względu na jej przewagę płynącą z agregacji dużej ilości danych, stanowi doskonały potencjał do wykorzystania przez modele AI.

- Pozytywne efekty sieciowe wynikające ze współpracy różnych podmiotów

Wybuch pandemii koronawirusa w 2020 roku spowodował, że odkładana przez wiele podmiotów transformacja technologiczna, stała się wręcz nieodzowna do przetrwania kryzysu. Ponad połowa polskich firm przyspieszyła transformację cyfrową w czasie pandemii, ale ze wszystkich możliwych narzędzi technologicznych najmniej popularne były narzędzia do analizowania dużych zbiorów danych oraz systemy przewidujące (Ernst & Young, 2021). Mogło to wynikać ze zbyt wysokich kosztów takich rozwiązań czy niekompletności zbiorów do właściwego wykorzystania przez modele sztucznej inteligencji. Gromadzenie danych różnych podmiotów mogłoby skłonić firmy, uczelnie jak i sektor publiczny do połączenia posiadanych zasobów (finansowych, intelektualnych i organizacyjnych) i współpracy w ramach określonych sektorów gospodarki dla odnalezienia jak najbardziej efektywnych rozwiązań.

Przykład projektu: Wspólnica Danych Zdrowotnych

Największa krajowa baza danych dotyczących zdrowia. Gromadzone w jej ramach dane służą podnoszeniu jakości systemu opieki zdrowotnej, projektowaniu nowoczesnych rozwiązań telemedycznych, a także prowadzeniu przełomowych badań naukowych. Źródła danych trafiających do wspólnicy to Narodowy Fundusz Zdrowia, Państwowa Inspekcja Sanitarna, Centrum Systemów Informacyjnych Ochrony Zdrowia, prywatne sieci medyczne, rejestry podmiotów leczniczych (placówki medyczne, szpitale itd.), zakłady ubezpieczeń i reasekuracji, aptek, a także inteligentnych urządzeń, czy aplikacji zdrowotnych i medycznych. Ze względu na posiadaną infrastrukturę i doświadczenie w zakresie gromadzenia danych z różnych systemów i bezpiecznego zarządzania posiadanymi zasobami, potencjalnym operatorem wspólnicy mogłoby być Centrum e-Zdrowia (CeZ)*.

***Na podstawie warsztatu przeprowadzonego z przedstawicielami sektora medycznego**

Bariery

- Brak ujednoliconych standardów;
- Wielość wykorzystywanych systemów w sektorze zdrowia*;
- Odmienna kultura zbierania danych w różnych placówkach*;
- Niemożliwość zapewnienia całkowitego utajnienia informacji osobowych*;
- Nieufność do udostępniania danych dotyczących zdrowia;
- Brak wzajemnego zaufania aktorów uczestniczących w procesie współdzielenia danych*;
- Brak zrozumienia relacji pomiędzy “interesem publicznym” a korzyścią indywidualną;
- Brak jednoznacznie określonych definicji celów naukowych*;
- Brak definicji legalnej danych dotyczących zdrowia i ich relacji wobec danych medycznych;
- Zamknięcie danych w bazach prywatnych firm;
- Brak wyraźnego modelu biznesowego*;
- Brak koordynacji i współpracy między dedykowanymi ministerstwami oraz instytucjami*;
- Brak łączy o odpowiedniej mocy w szpitalach*;

Proponowane rozwiązania

1) Interwencja ustawowa w zakresie standaryzacji danych

Rozporządzenie Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności, minimalnych wymagań dla rejestrów publicznych i wymiany informacji w postaci elektronicznej oraz minimalnych wymagań dla systemów teleinformatycznych (zwanym dalej „Rozporządzeniem KRI”) zawiera szczegółowe wymogi dotyczące zachowania odpowiednich standardów jedynie dla podmiotów sektora publicznego. Chociaż 68% placówek medycznych posiada rozwiązania IT pozwalające na prowadzenie dokumentacji w postaci elektronicznej, 93% ankietowanych wskazuje formę papierową jako najpopularniejszą metodę wymiany informacji pomiędzy podmiotami Co więcej, prawie 70% placówek nie wprowadza takich dokumentów następnie do systemu (Centrum e-Zdrowia, 2021).

Aby zmienić ten stan rzeczy, należy przede wszystkim zmienić przyzwyczajenia pracowników ochrony zdrowia w zakresie wprowadzania danych do systemu, a tym samym - dzielenia się nimi. W początkowej fazie istnienia wspólnicy, ważne będzie przełamanie istniejącego aksjomatu myślenia w zakresie udostępniania informacji. Należy wprowadzić ustawowy obowiązek umieszczania danych we wspólnicy, nałożony na konkretne podmioty zarówno prywatnej, jak i publicznej ochrony zdrowia. Ponadto, konieczne jest wprowadzenie odpowiedniej standaryzacji formatów w sektorze medycznym. Już teraz istnieją formaty stosowane powszechnie przez różne placówki, m.in. standardy wymiany informacji HL7 (Health Level Seven) czy DICOM (Digital Imaging and Communications in Medicine) - norma ujednolicająca wymianę i interpretację danych medycznych reprezentujących lub związanych z obrazami diagnostycznymi w medycynie. Ponieważ jednak wspólnica danych jest całkowicie nowym projektem, należy zbudować świadomość i nowe nawyki w sektorze zdrowia, rozpoczynając od sankcjonowanego obowiązku stosowania ustalonych formatów.

2) Wspólne zakupy oprogramowań na wzór wspólnych zakupów leków

Chociaż w stosunku do roku 2018, budżet wydatków na digitalizację sektora ochrony zdrowia wzrósł niemal dwukrotnie, Ministerstwo Zdrowia nie posiada informacji ani kontroli nad tym, jakie systemy informatyczne wybierają publiczne placówki ochrony zdrowia. Mają one jednak obowiązek zapewnienia interoperacyjności słabej z centralną architekturą zdrowia cyfrowego (Minister Zdrowia, 2022). Aby jednak wzmocnić bezpieczeństwo danych trafiających do wspólnicy, warto uodpornić systemy na możliwe ataki cybernetyczne, wybierając we wspólnym przetargu najlepsze możliwe oprogramowanie dla placówek medycznych.

3) **Stopniowe budowanie wzajemnego zaufania poprzez maksymalną transparentność celów i otwarty dialog**

Najważniejszą częścią wspólnic danych jest społeczność, która dzieli się swoimi danymi na rzecz osiągnięcia wielowymiarowej korzyści. Budowa zaufania społecznego powinna stać się jednym z najważniejszych elementów etycznego projektowania wspólnic danych. Ponieważ ma służyć dobru wspólnemu, do dyskusji na temat zasad funkcjonowania wspólnicy powinno się włączać wszystkie strony biorące udział w jej współtworzeniu, już na etapie projektowania. W związku z tym, wszelkie plany legislacyjne powinny być komunikowane z odpowiednim wyprzedzeniem i jak najszerszym zasięgiem społecznym - tak, by w sposób skuteczny mogły dotrzeć do każdego interesariusza.

Ze względu na swoją specyfikę i wrażliwy charakter, projekt ten powinien być jak najbardziej neutralny politycznie. Według najnowszych badań Edelman Trust Barometer mierzących poziom zaufania obywateli do poszczególnych sektorów, rząd i media znajdują się za biznesem i organizacjami pozarządowymi (Edelman Trust Barometer, 2022). Należy więc w jak największym stopniu włączyć organizacje typu NGO w cały proces wdrażania projektu wspólnicy. To również one, we współpracy z innymi podmiotami, powinny przewodniczyć działaniom promującym ich ideę za pomocą publicznych kampanii. Komunikacja społeczna powinna zawierać jak najwięcej przykładów technicznych poświadczających bezpieczeństwo i odporność technologiczną wspólnic przed możliwymi wyciekami danych. Ponadto, dzięki organizacji warsztatów wyjaśniających w jasny i przystępny sposób funkcjonowanie wspólnic, ich model biznesowy i możliwość wykorzystania danych, możliwe będzie stopniowo przywrócenie zaufania społecznego do dzielenia się danymi, a także przełamanie dychotomii interesu publicznego i interesu jednostki, postrzeganych jako dwóch, zupełnie opozycyjnych interesów.

4) **Ułatwienie możliwości zbierania danych na “cele naukowe” i inne, wyraźnie określone cele w interesie publicznym**

Zgodnie z art. 26 ust. 4 Ustawy z dnia 6 listopada 2008 r. o prawach pacjenta i Rzeczniku Praw Pacjenta, dokumentacja medyczna może być udostępniona także szkole wyższej lub instytutowi badawczemu do wykorzystania w celach naukowych, bez ujawniania nazwiska iinnych danych umożliwiających identyfikację osoby, której dokumentacja dotyczy. Ze względu na istniejący już potencjał organizacyjno-techniczny Centrum e-Zdrowia, zlecanie zadania tworzenia wspólnic w ramach szkół wyższych czy niektórych instytutów, mogłoby hamować proces współdzielenia danych i nie być dość efektywne. Warto zatem rozważyć poszerzenie katalogu podmiotów, którym można udostępniać dokumentację na podstawie przepisów zawartych w niniejszej ustawie i ustanowić zasady, na podstawie których taki dostęp byłby udzielany.

Z uwagi na niejasny charakter “celów naukowych”, rekomendowanym jest również zdefiniowanie tego pojęcia i poszerzenie katalogu możliwości wykorzystywania danych dotyczących zdrowia również na inne uzasadnione cele w interesie publicznym.

5) **Państwo jako facylitator procesu**

Publiczny charakter wspólnic danych wymaga maksymalnej inkluzywności w procesie ich tworzenia. Z tego względu, rola administracji publicznej powinna ograniczać się do koordynowania procesu zakładania rad społecznych, które to powinny podejmować decyzje w zakresie funkcjonowania wspólnic - w imię zasady “from decision maker to decision ‘taker’ “od najniższych szczebli.

6) **Ulepszenie technologii w każdej części procesu**

Chociaż postępująca cyfryzacja ochrony zdrowia przyniosła wiele pozytywnych zmian w zakresie przyspieszenia pewnych procesów, wiele szpitali w dalszym ciągu boryka się z problemem słabych łączy, a przez to trudnościami z szybkim przekazywaniu dużych plików. Sukces współdzielenia danych jest w dużej mierze uzależniony nie tylko od zdobycia odpowiedniego zaufania społecznego czy odpowiednich regulacji, ale również od możliwości technicznych podmiotów będących częścią całego łańcucha współdzielenia danych.

4.4 Metody zarządzania danymi w zależności od stopnia ich wrażliwości

4.4.1. Dane nieosobowe

Istnieją dane, które z różnych przyczyn mają poufny charakter, stanowiąc np. tajemnicę przedsiębiorstwa lub dane dotyczące zdrowia pacjentów. Wymagają on przez to szczególnej ochrony.

Dane nieosobowe rozumiane są jako informacje elektroniczne inne niż te wskazane w RODO - tj. nie dotyczące informacji o zidentyfikowanej lub możliwej do zidentyfikowania osobie fizycznej. Często nie są generowane przez człowieka, choć mogą być przez niego zbierane, przetwarzane i wykorzystywane. Do ich przykładów zaliczamy informacje cyfrowe wytwarzane przez maszyny (pochodzące z różnego rodzaju czujników i sensorów) i produkty elektroniczne (np. zanonimizowane zbiory BigData), jak i dane meteorologiczne i przyrodnicze. Często dostęp nich jest w pełni otwarty, jak chociażby w przypadku danych wskazujących natężenie ruchu na drogach; innym razem mogą one podlegać konkretnej organizacji bądź przedsiębiorcy np. twórcy oprogramowania czy właścicielowi maszyny To właśnie w tym drugim przypadku najczęściej dochodzi do zamknięcia danych. Wśród argumentów, na które firmy powołują się odmawiając dostępu do zagregowanych przez nie zbiorów danych najczęściej wymieniane są kwestie prywatności przedsiębiorstwa (tajemnic handlowych) oraz praw własności intelektualnej.

Jeżeli chodzi o tajemnicę przedsiębiorstwa, to w prawie polskim zagadnienia jej dotyczące zostały uregulowane przede wszystkim w ustawie z dnia 16 kwietnia 1993 r. o zwalczaniu nieuczciwej konkurencji („u.z.n.k.”). Z art. 11 ust. 2 u.z.n.k. wynika, że przez tajemnicę przedsiębiorstwa rozumie się informacje techniczne, technologiczne, organizacyjne przedsiębiorstwa lub inne informacje posiadające wartość gospodarczą, które jako całość lub w szczególnym zestawieniu i zbiorze ich elementów nie są powszechnie znane osobom zwykle zajmującym się tym rodzajem informacji albo nie są łatwo dostępne dla takich osób, o ile uprawniony do korzystania z informacji lub rozporządzania nimi podjął, przy

zachowaniu należytej staranności, działania w celu utrzymania ich w poufności. W przypadku danych nieosobowych mogą być to więc informacje pochodzące ze stron internetowych, urządzeń czy czujników maszyn, przyjmujące postać np. tekstu, liczb lub obrazów, ustrukturyzowane lub nieuporządkowane. Aby skorzystać z ochrony, dana informacja musi jednak zostać odpowiednio zabezpieczona przed dostępem osób nieuprawnionych i posiadać wartość gospodarczą lub należeć do zbioru lub zestawienia, który taką wartość posiada.

Jeżeli chodzi o prawa własności intelektualnej, to część danych w gospodarce cyfrowej takich, jak towary cyfrowe (utwory muzyczne, e-książki i oprogramowania) faktycznie podlegają ochronie na gruncie prawa autorskiego. W stosunku do tych danych ich twórcom przysługują wyłączne prawa majątkowe. Jednak wiele informacji cyfrowych, zwłaszcza tych generowanych maszynowo, nie spełnia wymogów ochrony praw autorskich (Kerber, 2016). Podobnie prezentuje się sytuacja dotycząca prawa *sui generis* do baz danych - o ile należą się one producentom baz danych, o tyle często uprawnienia z nich wynikające są nadużywane przez administratorów danych. Wynika to z faktu, iż wspomniana już ochrona *sui generis* przysługuje jedynie wtedy, kiedy w związku z bazą danych poczyniona została istotna inwestycja dla konieczna dla uzyskania, weryfikacji lub prezentacji tejże bazy (art. 7(1) dyrektywy 96/9). Ochrona bazy danych chroni więc niejako inwestycję dokonaną w celu gromadzenia i porządkowania już istniejących informacji cyfrowych, nie zaś samo wytwarzanie czy zbieranie danych (Kerber, 2016). W kontekście Big Data warto więc zauważyć, że o ile duże zbiory mogą być (pośrednio) chronione zarówno przez prawo autorskie, jak i w ramach reżimu *sui generis*, o tyle ochrona ta nigdy nie obejmie zawartości baz danych, czyli danych "samych w sobie" (Żyrek, 2022).

Jak wynika z powyższego, dane nieosobowe mogą podlegać specjalnym wymogom dotyczących poufności, a tym samym mogą wymagać specjalnego zarządzania nimi. Jednak nagminne powoływanie się przedsiębiorców na tajemnicę przedsiębiorstwa bądź prawa własności intelektualnej względem śladów cyfrowych pochodzących z maszyn czy czujników stanowi duże nadużycie. W ramach data economy postuluje się, żeby dane były dostępne dla jak największej ilości podmiotów - publicznych i prywatnych; dużych i małych; nowo powstałych i tych ugruntowanej na rynku pozycji (Komisja Europejska, 2020b). Ma to pozwolić na zwiększenie innowacji oraz wzrost gospodarczy wszystkich aktorów włączonych we współdzielenie. Warto bowiem podkreślić, że z perspektywy ekonomicznej dane nie są konkurencyjne (The Ministry of Electronics & Information Technology, Government of India, 2020). Tym samym, w przeciwieństwie do typowych paliw kopalnych, których wykorzystanie jest ograniczone od dostępności zasobów, czerpanie korzyści z danych przez jednego przedsiębiorcę nie wyklucza ich użyteczności dla innego podmiotu (Paszcza, 2022).

4.4.2 Dane wrażliwe

Szczególnie ważnymi, ale też wymagającymi wzmożonej ochrony danymi osobowymi są tzw. dane wrażliwe. Wynika to z faktu, że ich przetwarzanie może być dość poważną ingerencją w prywatną, czy nawet intymną sferę życia człowieka, stanowiąc jednocześnie podstawę do dyskryminacji takiej osoby (Fajgielski, 2021). Wśród kategorii wymienionych w RODO znajdują się dane dotyczące pochodzenia rasowego/etnicznego, poglądów politycznych, przekonań religijnych i światopoglądowych, kwestia przynależności do związków zawodowych, a także dane genetyczne, biometryczne przetwarzane w celu jednoznacznego zidentyfikowania osoby fizycznej lub dane dotyczące zdrowia, seksualności lub orientacji seksualnej danej osoby. W przeciwieństwie do poprzedzającej RODO

dyrektywy 95/46/WE, państwa członkowskie nie mogą rozszerzyć katalogu danych sensytywnych i objąć innych rodzajów takim samym reżimem, jaki posiadają one w RODO. Co do niektórych danych natomiast, mogą one przyjąć wyjątkowe regulacje chronioną poszczególne interesy jednostek - dotyczy to np. wyrażonego w Prawie bankowym obowiązku tajemnicy bankowej obejmującej informacje dotyczące czynności bankowej, uzyskane w czasie negocjacji, w trakcie zawierania i realizacji umowy danych (Prawo bankowe).

Szczególnie ważne w kontekście tworzenia wspólnic są dane dotyczące zdrowia i wymagana dla ich przetwarzania ochrona. Poza danymi w oczywisty sposób związanymi ze zdrowiem (np. informacje z Elektronicznej Dokumentacji Medycznej), istnieje szereg danych, które w zestawieniu z innymi mogą stanowić podstawę do wysnucia wniosków na temat czyjegoś stanu zdrowotnego. Ta trudność w rozgraniczeniu dotyczy w szczególności opasek typu *fitbit* i aplikacji sportowych, które mogą mierzyć nie tylko stan kondycji danej osoby, ale również jej tętno czy jakość snu.

Istnieją trzy możliwości rozstrzygnięcia, czy konkretne dane będą podlegać reżimowi danych o szczególnym charakterze z RODO:

- **Podejście kontekstowe** - dla określenia rodzaju danych bierze się pod uwagę interesy administratora oraz potencjalnych odbiorców danych, warunki przetwarzania danych, a także możliwe konsekwencje ich przetwarzania;
- **Podejście celowościowe** - skoncentrowane jedynie na jasno określonych intencjach przetwarzania, jakie miał administrator danych;
- **Podejście mieszane** - cel ma nadrzędne znaczenie, ale jeżeli celem nie było wyciągnięcie wniosków o wrażliwym charakterze lub uzyskanie danych, które mogłyby ujawnić sensytywne informacje o danej osobie, należy wziąć jeszcze pod uwagę kontekst i to, co można było racjonalnie przewidzieć, że dane te mogą ujawnić wrażliwe informacje o osobach, których dane dotyczą.

Podejście mieszane, choć zapewnia największe bezpieczeństwo osób, których dane dotyczą, może przysparzać licznych trudności twórcom oprogramowań i aplikacji. Co do zasady bowiem, przetwarzanie danych wrażliwych jest zakazane, chyba że zachodzi jedna z dziewięciu przesłanek wymienionych w RODO. Wśród tych, które mogą mieć znaczenie w przypadku danych dotyczących zdrowia i współdzielenia danych znajdują się przede wszystkim:

- **Wyraźna zgoda podmiotu, którego dane dotyczą**

Zgoda na przetwarzanie danych wrażliwych nie powinna mieć charakteru dorozumianego - oznacza to, że dana osoba w oczywisty sposób powinna wyrazić swoje przyzwolenie na przetwarzanie konkretnych danych osobowych we wskazanych przez administratora celach. Zbieranie zezwoleń bez określenia powodów i sposobu wykorzystania wyników danych jest niezgodne z zasadą przejrzystości wyrażoną w RODO. Administrator powinien również wykorzystywać jedynie te dane, które są stosowne do danego celu i ograniczone do tego co niezbędne (zasada minimalizacji danych). Nie powinien on korzystać z całego posiadanego zbioru, lecz wybierać jedynie te zestawy, które są mu potrzebne do przetwarzania. Biorąc pod uwagę wspólnicę danych zdrowotnych, której ogromne zasoby mogą być wykorzystywane przez rozmaite podmioty, wyzwaniem może być określenie z wyprzedzeniem wszystkich celów, a w związku z tym - w legalny sposób zebrać zgody od osób, których dane dotyczą (Najbuk et al, 2020).

Ciekawym rozwiązaniem może być fińska propozycja regulacji typu opt-in/opt-out, w której to dzielenie się danymi na cele wspólne odbywa się w sposób domyślny, a użytkownicy mogą w razie takiej potrzeby zrezygnować z dalszego udostępniania swoich danych. Ze względu na rolę interesu publicznego w rozwoju narzędzi służacych do ulepszania ochrony zdrowia, takie rozwiązanie wydaje się być najbardziej efektywnym.

Niezbędność przetwarzania do celów badań naukowych

Zgodnie z art. 26 ust. 4 Prawa pacjenta, dostęp do dokumentacji medycznej w celach naukowych może być przyznany szkole wyższej lub instytutowi. Zgodnie z motywem 159 RODO, do "celów naukowych" należą:

- rozwój technologiczny i demonstracja;
- badania podstawowe;
- badania stosowane;
- badania finansowane ze środków prywatnych

Jak wskazano w motywie, wyrażenie „do celów badań naukowych" powinno obejmować także badania prowadzone w interesie publicznym w dziedzinie zdrowia publicznego. W tym kontekście, RODO wskazuje również konieczność stworzenia europejskiej przestrzeni badawczej. Motywy zawarte w RODO stanowią jednak jedynie pewien kierunek interpretacyjny i nie mogą być podstawą prawną dla ewentualnych uzasadnień zbierania danych.

Bezpieczeństwo danych wrażliwych

Samo spełnienie przesłanki zezwalającej na przetwarzanie danych wrażliwych nie jest wystarczające. W każdym przypadku, konieczne jest zapewnienie zarówno bezpieczeństwa wewnętrznego systemu (polegającego na zapewnieniu odpowiedniego utajnienia danych), oraz zewnętrznego, dotyczącego odporności na niespełniajace europejskich wymogów ochrony danych osobowych regulacji państw trzecich.

Bezpieczeństwo wewnętrzne może być zapewnione przede wszystkim poprzez anonimizację lub pseudonimizację danych. Pojęcia te nie są tożsame - w uproszczeniu, w wyniku anonimizacji danych dane tracą swój przymiot osobowości w nieodwracalny sposób. Takie dane nie podlegają już wówczas reżimowi prawnemu z RODO.

Pseudonimizacja to natomiast zabieg uniemożliwiający zidentyfikowanie danej osoby bez dostępu do kluczy do tak zaszyfrowanych danych, przechowywanych bezpiecznie w innym miejscu. Skuteczna pseudonimizacja danych zdrowotnych nie jest jednak w pełni osiągalna*. Wynika to z faktu, że klucze do odszyfrowania danych zawsze pozostają w oddzielnej bazie jakiegoś podmiotu, który ma możliwość ich odczytania.

Z uwagi na utworzenie nowych sposobów przechowywania danych i pojawienie się nowych zagrożeń dla gromadzonych danych, ENISA wydała dokument zawierający zalecenia co do inżynierii ochrony danych, wskazujących techniczne i organizacyjne procesy, które już na etapie projektowania i domyślnej fazy tworzenia struktury chroniącej dane, pozwolą na zapewnienie odpowiedniego poziom bezpieczeństwa. Poza anonimizacją i pseudonimizacją ENISA wymienia m.in. maskowanie danych, czyli funkcję, za pomocą której ukrywa się prawdziwą wartość danych. Do jej rodzajów zalicza się zaufane środowisko przetwarzania, szyfrowanie homomorficzne (umożliwiające obliczenia na zaszyfrowanych danych bez ich odszyfrowywania), czy stosunkowo nowa metoda danych syntetycznych, polegająca na

tworzeniu zbioru danych przypominających rzeczywiste dane, ale różniące się od nich tym, że nie dotyczą żadnej prawdziwej osoby fizycznej (Kaczmarek, 2022). ENISA wskazuje również rekomendowane technologie w zakresie ograniczania nieuprawnionego dostępu do danych, takie jak bezpieczne poświadczanie uwierzytelniania, które pozwala na weryfikację podmiotów ubiegających się o dostęp bez konieczności ujawniania danych o nich samych.

Prawdziwym wyzwaniem jest jednak zapewnienie odporności na zagrożenia zewnętrzne, a więc dotyczące wzmocnienia suwerenności cyfrowej Polski na arenie międzynarodowej. Konieczność ta wynika z rozbieżności w poziomie ochrony danych osobowych zapewnianych przez regulacje unijne i regulacje państw trzecich. Dotyczy to zwłaszcza przepisów amerykańskich, których niejasny zakres obowiązywania był powodem złożenia skargi przez Maximiliana Schremsa do irlandzkiego organu ochrony danych osobowych. Za jej pośrednictwem ustalono, że transfer danych do Stanów Zjednoczonych na podstawie dotychczasowo obowiązującej "Tarczy Prywatności", jest zabroniony. Jako przyczynę takiego rozstrzygnięcia TSUE podał przepisy będące podstawą do funkcjonowania programów nadzoru wywiadowczego. Pozostają one bez ograniczeń co do możliwości ingerowania w prawo do prywatności osób niebędących obywatelami USA. Ponadto, w niektórych przypadkach nie zostało ustanowione prawo do zaskarżenia decyzji sądów, które mogą przyznawać niektórym organom uprawnienia do prowadzenia inwigilacji osób spoza USA.

Chociaż po wyroku TSUE w sprawie *Schrems II* możliwy jest transfer danych do USA na podstawie standardowych klauzul umownych, w dalszym ciągu wydawane są decyzje orzekające, że korzystanie z dostawców posiadających siedzibę Stanach Zjednoczonych jest niezgodne z RODO i z wymogiem spełnienia takich samych waunków ochrony danych osobowych. Z powodu niechęci do zmiany przepisów po stronie organów amerykańskich, konieczne jest rozważenie wyboru jedynie lokalnych, europejskich rozwiązań informatycznych - tak, aby mieć pewność, że dane obywateli zarówno pod względem ich odpowiedniego zaszyfrowania, jak i przechowywania w infrastrukturze, nie są w żaden sposób zagrożone.

5. Podsumowanie

Konieczne jest dokonanie odpowiednich zmian, polegających na *odtowarowieniu* danych i traktowaniu ich raczej jako dobro wspólne, zarządzane w imieniu właścicieli przez wyjęte poza logikę zysku, niekomercyjne instytucje publiczne. W świecie postępującej digitalizacji, to dane stanowią nasz najcenniejszy wspólny zasób. Nadrzędnym celem powinno być wytworzenie Wspólnej Wartości Społecznej (Shared Social Value) - w taki sposób, by poprzez dzielenie się danymi wspierać holistyczny rozwój społeczeństwa. Musimy bardziej świadomie wykorzystywać posiadane zasoby, nie pozwalając im marnować się przez zamknięcie w silosach organizacji czy dostawców technologicznych. Ze względu jednak na różny stopień wrażliwości takich danych, należy zwracać uwagę nie tylko na możliwe modele ale również na to, z jakimi danymi w konkretnym przypadku mamy do czynienia - taka informacja, w połączeniu z wiedzą na temat modelu zarządzania, może pozwolić na wybór najkorzystniejszego modelu współdzielenia.

Najważniejsze zadanie dla rozwoju Polski to zerwać z modelem kapitalizmu kognitywnego, w którym ludzie nie są podmiotami wskazującymi cele współdzielenia danych, lecz pełnią funkcję darmowych pracowników korporacji technologicznych. Aby jednak w ogóle móc

mówić o jakimkolwiek “współdzieleniu” trzeba zaangażować podmioty i ekspertów z możliwie jak największej liczby sektorów - tak, aby poprzez rozmowę doprowadzić do wypracowania najefektywniejszych rozwiązań dla wszystkich zainteresowanych.

Bibliografia

Alemanno A. (2018) Big Data for Good: Unlocking Privately-Held Data to the Benefit of the Many [online], European Journal of Risk REgulation, [Dostęp: 11.05.2022], Dostępny w : https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/big-data-for-good-unlocking-privatelyheld-data-to-the-benefit-of-the-many/C739E1DE223088FD3D761466DCDA2EFE

Artyushina, A. (2021) The future of data trusts and the global race to dominate AI [online], Bennett Institute For Public Policy, [Dostęp: 08.04.2022], Dostępny w: https://www.bennettinstitute.cam.ac.uk/blog/data-trusts1/

Baron, J., Contreras, J. L., Husovec, M., Larouche, P., Thumm, N. (2019) Making the Rules: The Governance of Standard Development Organisations and their Policies on Intellectual Property Rights, JRC Science for Policy Report, EUR 29655 EN (March 2019); ISBN 978-92-76-00023-5 , University of Utah College of Law Research Paper No. 308, TILEC Discussion Paper No. 2019-021, Available at SSRN: https://ssrn.com/abstract=3364722

Bayamlioglu, E. (2021) Data cooperative: a new intermediary on the horizon [online], KU Leuven, [Dostęp: 07.04.2022], Dostępny w: https://www.law.kuleuven.be/citip/blog/data-cooperative-a-new-intermediary-on-the-horizon/

Big Data Value Association (2019) Towards a European Data Sharing Space. Enabling data exchange and unlocking AI potential, BDVA Position Paper, Kwiecień 2019

Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H. (2018) Przemysł + Gospodarka oparta o dane, Ministerstwo Cyfryzacji, [Dostęp: 14.05.2022], Dostępny w: Gospodarka oparta o dane - Gov.plhttps://www.gov.pl › documents › Gospodarka+O...

Data Collaboratives (2021) Data Collaboratives [online], GovLab, [Dostęp: 08.04.2022], Dostępny w: https://datacollaboratives.org/

Data Trust Initiative (2021) Data trusts: international perspectives on the development of data institutions, DTA, Working Paper 2

Delacroix, S., Lawrence, N. D. (2019) Bottom-up data Trusts: distributing the ‘one size fits all’ approach to data governane, International Data Privacy Law, Volume 9, Issue 4, 236-252

Domeyer, A., Hieronimus, S., Klier, J., Weber, T. (2021) Government data management for the digital age [online], McKinsey & Company, [Dostęp: 07.04.2022], Dostępny w: https://www.mckinsey.com/industries/public-and-social-sector/our-insights/government-data-management-for-the-digital-age

Edelman Trust Barometer (2022) Wyniki najnowszego badania zaufania Edelman Trust Barometer 2022 [online], publicrelations.pl, [Dostęp: 15.05.2022], Dostępny w: https://publicrelations.pl/wyniki-najnowszego-badania-zaufania-edelman-trust-barometer-2022/

Empirica (2022) MonitorEHR [online], Empirica, [Dostęp: 12.05.2022], Dostępny w: https://empirica.com/project/details/?projectid=291

Ernst & Young (2021) 57% polskich firm przyspieszyło transformację cyfrową w czasie pandemii, a jeden na pięciu uważa, że w ich firmach transformacja jest zaawansowana [online], EY Polska, [Dostęp: 14.05.2022], Dostępny w:https://www.ey.com/pl_pl/news/2021/03/badanie-ey-transformacja-cyfrowa

European Data Protection Board (2020) Wytyczne w sprawie przetwarzania danych dotyczących zdrowia do celów badań naukowych w kontekście pandemii COVID-19 [online], EDPB, [Dostęp 14.05.2022], Dostępny w: https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_pl.pdf

Fajgielski, P. (2021) Komentarz do rozporządzenia nr 2016/679 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), [w:] Ogólne rozporządzenie o ochronie danych. Ustawa o ochronie danych osobowych. Komentarz, wyd. II

Foroohar R., (2019) Don't Be Evil: How Big Tech Betrayed Its Founding Principles -- and All of Us, Penguin Books Ltd, ISBN 13: 9781984824004

Goasduff, L. (2019) Top Trends on the Garnter Hype Cycle for Artificial Intelligence, 2019 [online], Gartner, [Dostęp: 13.05.2022], Dostępny w: https://www.gartner.com/smarterwithgartner/top-trends-on-the-gartner-hype-cycle-for-artificial-intelligence-2019

GovTech Polska (2020) Polityka rozwoju AI w Polsce przyjęta przez Radę Ministrów - co dalej? [online], gov.pl, [Dostęp 14.05.2022], Dostępny w: https://www.gov.pl/web/govtech/polityka-rozwoju-ai-w-polsce-przyjeta-przez-rade-ministrow--co-dalej

Grzeszak, J., Łukasik, K., Święcicki, I. (2021) Ile warte są nasze dane?, Polski Instytut Ekonomiczny, Warszawa

Hardjono, T., Pentland, S. (2018) Open Algorithms for Identity Federation, Proc IEEE Future of Information and Communication Conference, Singapur, Kwiecień 2018, https://arxiv.org/pdf/1705.10880.pdf

Hardjono T., Pentland S., (2019) Data Cooperatives: Towards a Foundation for Decentralized Personal Data Management, MIT Connection Science, Massachusetts Institute of Technology https://doi.org/10.48550/arXiv.1905.08819

Janssen, H., Singh, J. (2022) Data intermediary, Internet Policy Review, 11(1) https://doi.org/10.14763/2022.1.1644

Jemielniak, D., Przegalińska, A. (2020) Społeczeństwo współpracy, Wydawnictwo Naukowe Scholar, Warszawa, ISBN: 978-83-66470-04 -0

Jessop, B. (2007) State Power: A Strategic-Relational Approach, Polity, Cambridge

Kaczmarek, A. (2022) Inżynieria ochrony danych wg ENISA [online], TKP, [Dostęp: 16.05.2022], Dostępny w: https://www.traple.pl/2022/04/06/inzynieria-ochrony-danych-wg-enisa/

Kaplan, A., Haenlein, M. (2019) Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence, Business Horizons, Vol. 62 Issue 1, January - February 2019, 15-25

Kawalec, J. (2021) Sztuczna inteligencja - wyścig o naszą wolność [online], Pomorski Przegląd Gospodarczy, [Dostęp: 14.05.2022], Dostępny w: https://ppg.ibngr.pl/pomorski-przeglad-gospodarczy/sztuczna-inteligencja-wyscig-o-nasza-wolnosc

Kerber, W., A New (Intellectual) Property Right for Non-Personal Data? An Economic Analysis (October 24, 2016). Gewerblicher Rechtsschutz und Urheberrecht, Internationaler Teil (GRUR Int), 11/2016, 989-999

Komisja Europejska (2018) Staff Working Document - Guidance on sharing private sector data in the European data economy [online], Komisja Europejska, [Dostęp: 13.05.2022], Dostępny w: https://digital-strategy.ec.europa.eu/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy

Komisja Europejska (2020a) Europejska strategia w zakresie danych [online], Komisja Europejska, [Dostęp: 11.05.2022], Dostępny w: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_pl

Komisja Europejska (2020b) Rozporządzenie Parlamentu Europejskiego i Rady 2020/0340 z dnia 25 listopada 2020 r w sprawie europejskiego zarządzania danymi (akt w sprawie zarządzania danymi) stanowiące uzupełnienie Dyrektywy Parlamentu Europejskiego i Rady (UE) 2019/1024 z dnia 20

czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystywania informacji sektora publicznego

Komisja Europejska (2020c) Horizon 2020, Work Programme 2018-2020 Information and Communication Technologies, European Commission Decision C(2020)4029, 17 czerwca 2020

Komisja Europejska (2021a) The Digital Economy and Society Index [online], Komisja Europejska, [Dostęp: 12.05.2022], Dostępny w: https://digital-strategy.ec.europa.eu/en/policies/desi

Komisja Europejska (2022) Cyfrowe dane i usługi dotyczące zdrowia - europejska przestrzeń danych dotyczących zdrowia [online], Komisja Europejska, [Dostęp: 11.05.2022], Dostępny w: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12663-Cyfrowe-dane-i-us%C5%82ugi-dotyczace-zdrowia-europejska-przestrzen-danych-dotyczacych-zdrowia_pl

Mayer-Schönberger V., Ramge T. (2022) Access Rules: Freeing Data from Big Tech for a Better Future, University of California Press; First edition (April 26, 2022) ISBN-13: 978-0520387737

Małobęcka-Szwast, I. (2021) Data Governance Act - o krok bliżej do łatwiejszego dzielenia się danymi [online], newtech.law, [Dostęp: 11.05.2022], Dostępny w: https://newtech.law/pl/data-governance-act-o-krok-blizej-do-latwiejszego-dzielenia-sie-danymi/

Małobęcka-Szwast, I. (2022) Projekt Aktu w sprawie danych (Data Act) - kolejne ułatwienia w zakresie dzielenia się danymi [online], newtech.aw [Dostęp: 11.05.2022], Dostępny w: https://newtech.law/pl/projekt-aktu-w-sprawie-danych-data-act-kolejne-ulatwienia-w-zakresie-dzielenia-sie-danymi/

Mehta, S., Dawande, M., Mookerjee, V. (2021) Can data cooperatives sustain themselves? [online], LSE, [Dostęp: 14.05.2022], Dostępny w: https://blogs.lse.ac.uk/businessreview/2021/08/02/can-data-cooperatives-sustain-themselves/

Mehta, S., Dawande, M., Mu, L. (2022) The key to designing sustainable data cooperatives [online], Światowe Forum Ekonomiczne, [Dostęp: 14.05.2022], Dostępny w: https://www.weforum.org/agenda/2022/02/the-key-to-designing-sustainable-data-cooperatives/

Minister Zdrowia (2022) Odpowiedź na interpelację nr 3092 Posła Roberta Kwiatkowskiego w sprawie dostępu do informacji dotyczących cyfryzacji służby zdrowia [online], Ministerstwo Zdrowia, [Dostęp: 15.05.2022], Dostępny w: https://interpelacje.sejm.gov.pl/interpelacje9.nsf/0/E8019F514173502EC12587EC0040E718/%24File/ODP_K9INT30932.pdf

Nagel, L., Lycklama D. (2021) Design Principles for Data Spaces. Position Paper. Version 1.0. Berlin

Najbuk, P., Pachocki, J., Kruczyk-Gonciarz, A., Kaźmierczyk, P. Lorent, R. (2020). Wykorzystanie danych medycznych w celu rozwoju AI w Polsce i w celu prowadzenia badań naukowych. Raport Regulacyjny, DZP, Warszawa

Nayyar, A., & Puri, V. (2016, September) Smart farming: IoT based smart sensors agriculture stick for live temperature and moisture monitoring using Arduino, cloud computing & solar technology. In *Proc. of The International Conference on Communication and Computing Systems (ICCCS-2016)* (pp. 9781315364094-121). [Dostęp:14.05.2022], Dostępny w: https://www.researchgate.net/profile/Anand-Nayyar/publication/313804002_Smart_farming_IoT_based_smart_sensors_agriculture_stick_for_live_temperature_and_moisture_monitoring_using_Arduino_cloud_computing_solar_technology/links/59d9f67c0f7e9b12b36d66f8/Smart-farming-IoT-based-smart-sensors-agriculture-stick-for-live-temperature-and-moisture-monitoring-using-Arduino-cloud-computing-solar-technology.pdf.

Nowoczesna Polska, Lekcja - Cyfrowy świat [online], Edukacja medialna, [Dostęp: 10.05.2022], https://edukacjamedialna.edu.pl/lekcje/cyfrowy-slad/

PAP (2021) Rejestr ciąż. Ministerstwo Zdrowia: “Chodzi o względy medyczne” [online], Dziennik Gazeta Prawna, [Dostęp: 14.05.2022], Dostępny w: https://serwisy.gazetaprawna.pl/zdrowie/artykuly/8299619,rejestr-ciaz-ministerstwo-zdrowia-chodzi-o-wzgledy-medyczne.html

Paszcza, B. (2022) Dwa wielkie wyzwania e-gospodarki: kontrola nad danymi i legislacja interoperacyjności [online], Klub Jagielloński, [Dostęp: 13.05.2022], Dostępny w: https://klubjagiellonski.pl/2022/04/22/dwa-wielkie-wyzwania-e-gospodarki-kontrola-nad-danymi-i-legislacja-interoperacyjnosci/

Petland, A., Hardjono, T. (2020) Data Cooperatives [online], Work in Progress MIT, [Dostęp: 15.05.2022], Dostępny w: https://wip.mitpress.mit.edu/pub/pnxgvubq/release/2

Rada Unii Europejskiej i Rada Europejska (2021) EU looks to make data sharing easier: Council Agrees position on Data Governance Act [online], Rada UE i Rada Europejska, Press Release, [Dostęp: 14.05.2022], Dostępny w: https://www.consilium.europa.eu/en/press/press-releases/2021/10/01/eu-looks-to-make-data-sharing-easier-council-agrees-position-on-data-governance-act/

Rozporządzenie 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 swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych), OJ L 119, 4.5.2016, p. 1–88, art. 9.

Schubert, S., Harari Dayan, F. (2020) When is data pooling anticompetitive? [online], Freshfields Bruckhaus Deringer, [Dostęp: 13.05.2022], Dostępny w: https://technologyquotient.freshfields.com/post/102glxx/when-is-data-pooling-anticompetitive

Swant. M (2019), People Are Becoming More Reluctant To Share Personal Data, Survey Reveals [online], Forbes, [Dostęp: 10.05.2022], https://www.forbes.com/sites/martyswant/2019/08/15/people-are-becoming-more-reluctant-to-share-personal-data-survey-reveals/?sh=66b3889b1ed1

The Ministry of Electronics & Information Technology, Government of India (2020) Report by the Committee of Experts on Non-Personal Data Governance Framework, 111972/2020/CL&ES

Ustawa z dnia 29 sierpnia 1997 r. - Prawo bankowe, Dz.U. 1997 nr 140 poz. 939

Ustawa z dnia 6 listopada 2008 r. o prawach pacjenta i Rzeczniku Praw Pacjenta, Dz. U 2009 nr 52 poz. 417

Wawrzyniak, B., Zygmuntowski, J. J., Lamański, F. (2020) Polska suwerenna cyfrowo. Regulacje na rzecz sprawiedliwej i konkurencyjnej gospodarki cyfrowej. Instrat Policy Paper 06/2020

Van Hesteren, D., Van Knippenberg, L., Weyzen, R., Huyer, E., Cecconi, G. (2021), Open Data Maturity Report 2021, data.europa.eu, Publications Office of the European Union

Verhlust,S., Young, A., Srinivasan, P. (2022) An Introduction to Data Cooperatives [online], GovLab, [Dostęp: 08.04.2022], Dostępny w: https://datacollaboratives.org/introduction.html#section1

Zygmuntowski, J. J. (2020a) Wspólnice danych: Alternatywny model zarządzania danymi, Raport projektu: SpołTech, Centrum Cyfrowe

Zygmuntowski, J. J. (2020b) Kapitalizm Sieci, Stowarzyszenie Rozruch, ISBN: 978-83-957-6720-3

Zygmuntowski J. J., Chojecka K., & Roy, S.A., (2020), Podatek cyfrowy od gigantów. Ekspertyza w zakresie wprowadzenia w Polsce podatku cyfrowego (DST).

Zygmuntowski, J. J., Zoboli, L., Nemitz, P. F. (2021). Embedding European values in data governance: a case for public data commons. Internet Policy Review, 10(3). https://doi.org/10.14763/2021.3.1572

Żyrek, A. (2022) Big Data cz. 1, Big Data a prawo autorskie i ochrona sui generis baz danych [online], B&K, [Dostęp: 14.05.2022], Dostępny w: https://bartakalinski.pl/artykuly/big-data-cz-i-big-data-a-prawo-autorskie-i-ochrona-sui-generis-baz-danych/