Uwolnić potencjał danych.

Zarządzanie danymi jako zasobem współdzielonym

Blanka Wawrzyniak Marta Musidłowska Jan J. Zygmuntowski

Warszawa | sierpień 2022

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Uwolnić potencjał danych.

Zarządzanie danymi jako zasobem współdzielonym

Raport przygotowany dla Cyfryzacji KPRM
Blanka Wawrzyniak
Marta Musidłowska
Jan J. Zygmuntowski

Warszawa, sierpień 2022

Rekomendujemy cytowanie:
Musidłowska M., Wawrzyniak B., Zygmuntowski J.J., *Uwolnić potencjał danych. Zarządzanie danymi jako zasobem współdzielonym*. Raport przygotowany na podstawie projektu Roberta Kroplewskiego na zamówienie Cyfryzacji KPRM.

Autorzy:
Blanka Wawrzyniak,
Marta Musidłowska,
Jan J. Zygmuntowski

Inicjator projektu: Robert Kroplewski, r.pr., Pełnomocnik Ministra Cyfryzacji do spraw społeczeństwa informacyjnego

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

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

Kierownik projektu: Robert Kroplewski, r.pr., Pełnomocnik Ministra Cyfryzacji do spraw społeczeństwa informacyjnego

Opracowano dla: Cyfryzacja KPRM

Warszawa, sierpień 2022

Nota wstępna:
Niniejsza publikacja stanowi przedmiot zamówienia realizowanego na zlecenie Kancelarii Prezesa Rady Ministrów (KPRM) oraz obejmuje kompleksowe omówienie tematu współdzielenia danych z wyróżnieniem czterech zidentyfikowanych przez autorów wyzwań tj. współdzielenia danych w modelu: 1) Pośredników danych osobowych; 2) Wirtualnych wspólnic dla danych nieosobowych; 3) Publicznych wspólnic dla danych osobowych; oraz 4) zarządzania danymi o szczególnej wrażliwości.

GOV.pl CYFRYZACJA

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Spis treści

Słownik pojęć

stworzony na potrzeby niniejszej publikacji

Bazy danych

zorganizowany zbiór usystematyzowanych informacji (danych), gromadzonych według określonych reguł; zwykle przechowywany w systemie komputerowym w formie elektronicznej. Np. dane w najpopularniejszych typach baz danych stosowanych obecnie są umieszczone w wierszach i kolumnach szeregu tabel, co usprawnia przetwarzanie danych i tworzenie dotyczących ich zapytań, ułatwia dostęp do danych oraz udzielanie dostępu do nich, a także umożliwia zarządzanie i sterowanie nimi, ich modyfikowanie, aktualizowanie i organizowanie (Oracle Polska).

Big Data

inaczej *analizy wielkich zbiorów danych*; są to zbiory informacji o dużej objętości, dużej zmienności lub dużej różnorodności, które wymagają nowych form przetwarzania w celu wspomagania podejmowania decyzji, odkrywania nowych zjawisk oraz optymalizacji procesów (Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H., 2018).

Dane dotyczące zdrowia /Dane zdrowotne

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, M., Maśniak, L., Kroplewski, R., Romaniec, H., 2018). W odniesieniu do współdzielenia danych, dług innowacyjny oznacza **utratę** przez polskich przedsiębiorców **potencjalnych zysków bądź ponoszenie dodatkowych kosztów wynikających z braku dostępu do 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 **format 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).

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 lub organizacjom do gromadzenia danych dotyczących tychże podmiotów bądź danych wygenerowanych z użytkowanych przez te jednostki/organizacje 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).

Repozytorium danych

miejsce służące do przechowywania i porządkowania dokumentów przeznaczonych do udostępniania.

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.

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, dla autonomicznego kształtowania szans (Wawrzyniak, Zygmuntowski i Lamański, 2020).

Sztuczna inteligencja

system maszynowy, który jest w stanie wpływać na środowisko poprzez wytwarzanie danych wyjściowych (przewidywań, zaleceń lub decyzji) dla danego zestawu celów. Wykorzystuje on informacje pochodzące z maszyn lub od ludzi do (i) postrzegania rzeczywistych lub wirtualnych środowisk; (ii) tworzenia abstrakcji tego postrzegania w postaci modeli poprzez analizę w sposób zautomatyzowany (np. za pomocą uczenia maszynowego) lub ręczny; oraz (iii) wykorzystywania wnioskowania z modelu do formułowania opcji wyników (OECD, 2019).

Ślad cyfrowy

niepowtarzalny zestaw możliwych do prześledzenia działań, czynności, wypowiedzi i komunikatów cyfrowych danej osoby, które przejawiają się w Internecie lub urządzeniach cyfrowych. Ślady cyfrowe można sklasyfikować jako pasywne lub aktywne. Pierwsze z nich składają się z aktywności użytkownika podczas przeglądania stron internetowych oraz informacji zapisanych w plikach cookie. Te drugie są często tworzone celowo przez użytkownika w celu udostępniania informacji na stronach internetowych lub w mediach społecznościowych. Choć termin ten odnosi się zazwyczaj do osoby, to ślad cyfrowy może również dotyczyć firmy, organizacji lub korporacji.

Wirtualna składnica danych

forma współpracy w ramach organizacji zrzeszonych podmiotów; pewien wzorzec zachowania polegający na dzieleniu się danymi w ramach federacji przedsiębiorców. W zaufanym środowisku cyfrowym (przestrzeniach danych) na zasadzie wzajemności w ramach przyjętej wzorcowej logiki dostępności danych.

Wspólnica danych

zaufana instytucja współdzielenia danych zgodnie z interesem publicznym/wspólnym. Może przybierać formy przestrzeni dla danych zarówno nieosobowych, jak i osobowych (w tym tych o szczególnym charakterze). **Wirtualna wspólnica danych** oznacza przestrzeń dla współdzielenia danych opartą na współpracy w ramach federacji zrzeszonych podmiotów; przewiduje budowanie rozproszonych repozytoriów danych w których członkowie zapewniają dostęp do danych na uznanych wspólnie zasadach technicznych, organizacyjnych i prawnych (Kroplewski, R., 2020). Koncepcja powstała na kanwie Wirtualnej składnicy danych, jednak termin “wspólnica” ma za zadanie uwypuklać wspólny charakter dostępu do danych, “wirtualność” wspólnicy ma natomiast podkreślać jej cyfrową formę oraz federacyjny model wymiany danych za pomocą rozproszonych repozytoriów (w odróżnieniu od repozytorium centralnego, w którym przechowywane są wspólne dane). Celem jest umożliwianie działań z zakresu analizy biznesowej (Business Intelligence, BI), w szczególności analityki, najczęściej na danych przemysłowych, rolniczych i biznesowych.

Publiczna wspólnica danych oznacza podmiot publiczny lub prywatno-publiczny (hybrydowy) zarządzający danymi osobowymi z różnych źródeł o znaczeniu publicznym (np. dane zdrowotne). Przy zapewnieniu odpowiednich zabezpieczeń prywatności, instytucja ta przyznaje różnym podmiotom (zarówno publicznym, jak i prywatnym) możliwość dostępu do zgromadzonych danych na określonych zasadach.

Termin pochodzi od “wspólnej składnicy danych”.

Współdzielenie danych

zbiór praktyk, technologii, elementów kulturowych i ram prawnych mających za zadanie sprzyjać dzieleniu się wartością danych przez podmioty. Współdzielenie może obejmować różne sposoby przetwarzania danych, w tym udostępnianie i powierzanie (na gruncie RODO), ale również dzielenie się dostępem do danych (bez transferu).

Skróty i objaśnienia

B2G	ang. Business to Government – współdzielenie danych prywatnych (biznesowych) z administracją publiczną
ENISA	ang. European Union Agency for Cybersecurity – Agencja Unii Europejskiej ds. Cyberbezpieczeństwa
GAFA	największe amerykańskie firmy z branży IT: Google, Amazon, Facebook, Apple
IT	ang. Information technology – branża zajmująca się zastosowaniami technologii obliczeniowych
KOWR	Krajowy Ośrodek Wsparcia Rolnictwa
MŚP	Małe i średnie przedsiębiorstwa
NGO	ang. Non-governmental organisation – organizacja pozarządowa
TSUE	Trybunał Sprawiedliwości Unii Europejskiej
UE	Unia Europejska

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 i gospodarcza. 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.

W 2020 roku, zarówno przedstawiciele GAFA jak i pozostali technologiczni giganci (Microsoft, Tencent i Alibaba), posiadali nieco ponad 300 zależnych spółek usytuowanych w rajach podatkowych, wykorzystywanych do uzyskiwania korzystniejszych warunków obciążeń publicznoprawnych (Wawrzyniak, B., Iwanowski, D., 2021). Co więcej, firmy te przez wiele lat prowadzenia działalności twierdziły, że nie dotyczą ich jakiekolwiek regulacje krajowe, ponieważ nie da się jednoznacznie ustalić, gdzie znajduje się faktyczna alokacja ich kapitału i gdzie jest generowana wartość z procesowanych danych. 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. Znane są również liczne naruszenia przepisów ochrony konkurencji, dotyczące m.in. algorytmicznej dyskryminacji porównywarek cen innych niż Google Shopping w wyszukiwarce, czy wpływanie na wolność słowa w Internecie poprzez tendencyjne blokowanie wybranych treści o charakterze politycznym. Dochodziło również do sytuacji, w których dane służyły do profilowania użytkowników wzmacniającego krzywdzące stereotypy i nierówności czy radykalizowania niektórych grup społecznych, co prowadziło do destabilizacji porządku publicznego (Zygmuntowski, 2020a).

Efektem tego stanu rzeczy była zmiana paradygmatu myślenia w stronę zwiększenia ochrony prywatności w sieci. Chociaż ochrona danych osobowych stanowi atrybut dóbr osobistych, ze względu na cyfrowy format zawartych w nich informacji i możliwość ich spieniężenia, dane zaczęto traktować jak chronione prawem własności (prawem majątkowym), co nie znajduje podstawy w prawie stanowionym. 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 posiadających jedynie właściwości osobiste oraz podlegających jedynie dyspozycji 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 cech o charakterze ekonomicznym 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 na rzecz ich ponownego wykorzystania w celu zapewnienia ogólnego społecznego wzrostu i wytworzenia wspólnej i publicznej wartości (Creating Shared/Public Value).

Chociaż więc niektórzy określają dane jako “nową ropę” (np. The Economist), w rzeczywistości posiadają one fundamentalnie odmienne cechy. Dane nieprzetworzone na przykład, mogą być przyrównane do “powietrza” jako zasobu istniejącego w środowisku naturalnym człowieka (np. w zakresie danych powstających w wyniku funkcjonowania smart cities, inteligentnych urządzeń (IoT) czy też transportu autonomicznego). Całkowita wolność w zakresie dostępu do danych powinna dotyczyć **danych publicznych czy danych pochodzących ze środowiska naturalnego człowieka. Dane firm, szczególnie małych i średnich przedsiębiorstw, czy też dane dotyczące zdrowia obywateli, ze względu na swoją specyfikę, powinny być możliwe do wykorzystania w oparciu o wypracowane podejście danego sektora (Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H., 2018). To właśnie jest celem niniejszego raportu.**

Cyfrowe dane nie są konkurencyjne w konsumpcji, ponieważ ich zasobów nie da się wyczerpać poprzez wielokrotną eksploatację (Zygmuntowski, 2020a). 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 i dzielenia się nimi, ale też brać pod uwagę możliwość wielokrotnego wykorzystywania informacji w rozmaitych celach. Istnieją bowiem dane, które z różnych przyczyn mają poufny charakter, stanowiąc tajemnicę przedsiębiorstwa lub dane dotyczące zdrowia pacjentów, przez co wymagają szczególnej ochrony. Tajemnicę przedsiębiorstwa mogą stanowić informacje mające wartość gospodarczą, które jako całość lub w szczególnym zestawieniu nie są powszechnie znane osobom zwykle zajmującym się tym rodzajem informacji albo nie są łatwo dostępne dla takich osób, o ile uprawniony podjął odpowiednie działania w celu utrzymania ich w poufności. Dane te mogą mieć osobowy (np. listy klientów) oraz nieosobowy charakter (np. sposoby ulepszania produktu). W przypadku tych pierwszych, zastosowanie będą miały zasady wynikające z przepisów RODO – stopień ochrony poszczególnych danych osobowych natomiast będzie zależeć od tego, czy przetwarzane dane stanowią “dane wrażliwe” czy też nie. W przypadku tej drugiej kategorii, tj. danych nieosobowych, nie stosuje się przepisów z RODO – mogą natomiast być objęte regulacjami dotyczącymi nieuczciwej konkurencji.

Z potencjału danych powinni zatem korzystać nie tylko giganci technologiczni, ale także (albo przede wszystkim) organy administracji, dostawcy usług publicznych, ośrodki naukowo-badawcze, twórcy w dziedzinie kultury, sztuki i innowacji, organizacje pozarządowe i nietechnologiczne MŚP (a więc wykorzystujące technologie jedynie pomocniczo, w związku z realizacją innego rodzaju działalności komercyjnej). Ignorując bowiem potęgę informacji cyfrowej, podmioty te pozbawiłyby się dostępu do nowych, inteligentnych rozwiązań, a tym samym groziłoby im popadnięcie w dług technologiczny i wynikające z niego straty.

Najbardziej perspektywiczną strategią dla nowoczesnych gospodarek jest zarządzanie **danymi jako wspólnym zasobem**, o charakterystyce infrastruktury (“środek dla wielu aktorów do wielu celów”), 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 współdzielenia danych to całkowicie nowa koncepcja, ich powstanie może być porównane do niegdyś innowacyjnego pomysłu zakładania banków czy spółdzielni, bez których funkcjonowanie współczesnej gospodarki jest trudne do wyobrażenia.

2. Stan gospodarki opartej na danych w Polsce

Wartość danych w Polsce szacowana jest na 6,2 mld euro. W ciągu najbliższych 3 lat wartość ta może osiągnąć nawet do 12 mld euro (Bożykowski et al., 2019). Istotność danych okazuje się znamienna dla wzrostu gospodarczego – w przypadku Polski ogólna produktywność związana z danymi plasuje się na poziomie 92%, dużo powyżej średniej europejskiej. Oznacza to, że korzyść ekonomiczna ze zwiększenia wykorzystania danych w Polsce będzie szczególnie wymierna. Jednocześnie intensywność wykorzystania danych, też nieosobowych, w Polskiej gospodarce okazała się znamienna, gdyż polskie PKB zależy w 46% od transgranicznego przepływu danych nieosobowych (Koloch, G., Grobelna, K., Zakrzewska-Szlichtyng, K., Kamiński, B., Kaszyński, D., 2017).

Zauważenie potencjału danych dla rozwoju polskiej gospodarki dało podwaliny do przyjęcia rozmaitych strategii politycznych, rekomendacji i zmian regulacyjnych w dziedzinie ich ponownego wykorzystania dla poszczególnych technologii. Już dokumenty z 2013 roku wskazywały na konieczność otwarcia dostępu do informacji publicznej dla przedsiębiorstw i obywateli poprzez udostępnianie danych i dokumentów do ponownego wykorzystywania (Ministerstwo Gospodarki, 2013).

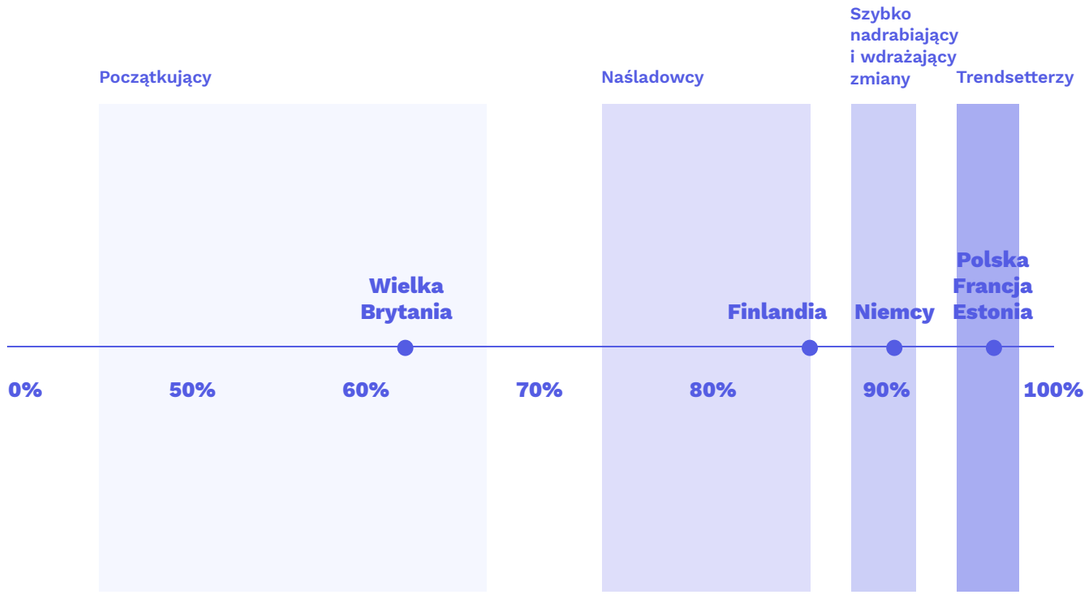
W kolejnych latach, polskie dokumenty strategiczne w coraz większym stopniu wskazywały na potrzebę skierowania działalności regulacyjnej i wdrożenia innowacyjnych zmian w zakresie cyfryzacji, dostępu i efektywnego wykorzystania gromadzonych danych. Począwszy od programu otwierania danych publicznych, wyraz prawa do informacji na miarę postępu technologicznego (Ministerstwo Cyfryzacji, 2016), poprzez Strategię na rzecz Odpowiedzialnego Rozwoju, której pierwszoplanowym celem było zapewnienie trwałego wzrostu gospodarczego opartego o wiedzę, dane i doskonałość organizacyjną (Strategia na Rzecz Odpowiedzialnego Rozwoju, 2017), polskie rekomendacje w coraz bardziej kompleksowy sposób dążyły do ukształtowania jednolitego podejścia do technologii informacyjnej.

Jeśli chodzi o dane nieosobowe, kierunek polskiej strategii w tym zakresie nadał raport “Przemysł Plus – Gospodarka oparta o dane” z 2018 roku. Wskazano w nim 5 kluczowych filarów dla rozwoju gospodarczego opartego o wykorzystanie danych – **dostęp do danych, zaawansowane umiejętności ich przetwarzania, cyfryzacja przemysłu, łączność oraz zaufanie uczestników i procesów w jej ramach** (Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H., 2018). W dokumencie wskazano, że kluczowym zadaniem w najbliższym czasie jest cyfryzacja przemysłu, dotycząca zarówno cyfryzacji obiektów używanych w przemyśle, jak również oparcia o dane procesów biznesnowych **czy ponownego użycia i łączenia strumieni danych w ramach i pomiędzy sektorami**.

Po licznych dyskusjach prowadzonych na poziomie unijnym, kolejnym ważnym krokiem podkreślającym potencjał danych dla rozwoju rozmaitych sektorów funkcjonowania państwa, była “Polityka dla rozwoju sztucznej inteligencji w Polsce” (Rada Ministrów, 2020). Dokument wskazuje, jak ważny jest etap pozyskiwania, gromadzenia, analizy i świadomego wykorzystywania danych dla rozwoju sztucznej inteligencji w obszarach takich jak społeczeństwo, innowacyjne firmy, nauka, edukacja, współpraca międzynarodowa i sektor publiczny.

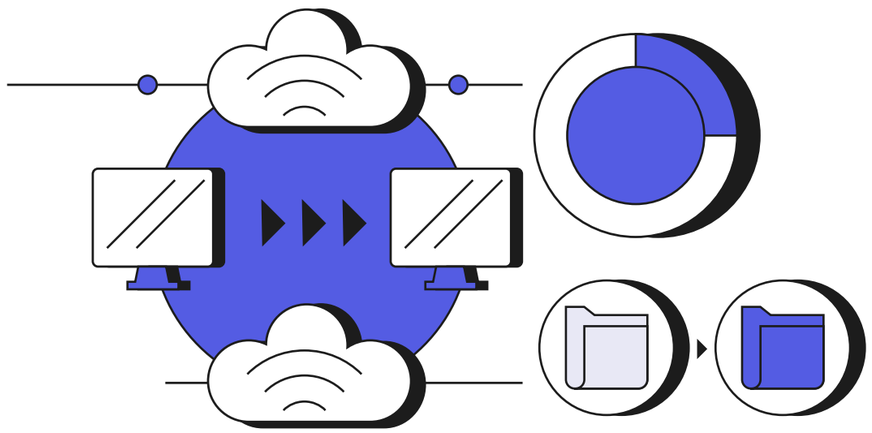
Pomimo sukcesów cyfryzacji usług publicznych w ostatnich latach (m.in. mObywatel, Internetowe Konto Pacjenta), Polska plasuje się na końcu rankingu Digital Economy and Society Index przygotowanego przez Komisję Europejską w 2021 roku (Komisja Europejska, 2021). Wynika to z wielu różnych czynników, które 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.

WYKRES 1. Dojrzałość otwartych danych w UE

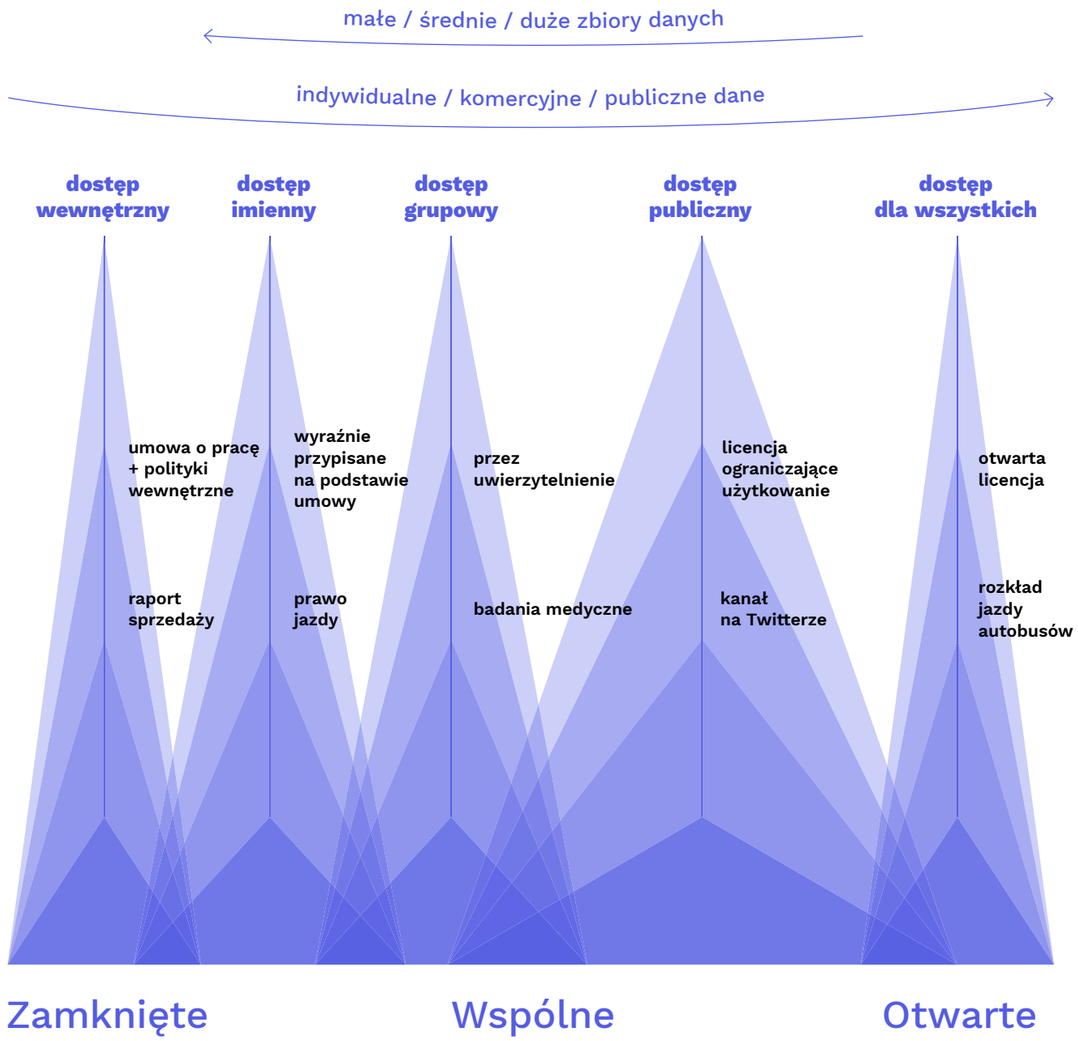


Sukces polskich projektów cyfryzacji administracji publicznej i ochrony zdrowia nie powinien jednak przysłonić ich niedoskonałości i barier, z którymi należy się zmierzyć. Internetowe Konto Pacjenta i związana z nim Elektroniczna Dokumentacja Medyczna (EDM) w dalszym ciągu zapewniają jedynie interoperacyjność 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 EDM 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. Jednocześnie nie można bagatelizować niskiego poziomu zaufania, co dobitnie pokazała społeczna reakcja na wdrażanie Zintegrowanej Platformy Analitycznej oraz EDM. Aby dokonać modernizacji dotychczasowego modelu zarządzania cyfrowymi informacjami, konieczne jest zidentyfikowanie występujących barier, określenie priorytetów oraz opracowanie wytycznych i standardów współdzielenia danych w zaufanych przestrzeniach.



WYKRES 2. Przekrój otwartości danych



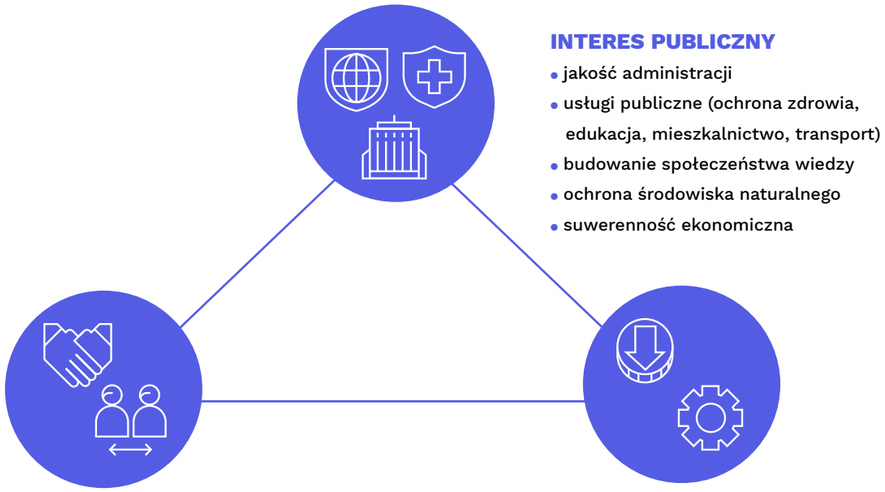
Źródło: The Open Data Institute's (ODI) Data Spectrum

3. 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 "własności danych osobowych", pozwalając na ukształtowanie się rynku ich komercyjnej wymiany. Takie podejście prowadzi jednak do zamykania danych w prywatnych silosach i utraty potencjału płynącego z agregacji i dalszego (ponownego) wykorzystywania. Współdzielenie danych, aby zmaksymalizować korzyści, musi nie tylko uniknąć pułapki utowarowienia danych, ale też pułapki indywidualnej własności.

Zarządzanie danymi prezentuje się raczej jako trylemat, który musi być rozstrzygany w zaufanej, bezpiecznej przestrzeni współpracy – zarówno w kontekście infrastruktury, jak i procedur instytucji i rozwiązań prawnych.

WYKRES 3. Trylemat zarządzania danymi



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 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 and communities (np. optymalizowanie konsumpcji energii poprzez inteligentne taryfy miejskie; efektywne zarządzanie transportem miejskim, , inteligentne łączenie usług ośrodków miejskich i wiejskich;
- 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);
- spersonalizowana edukacja;

Przedsiębiorcy i biznes:

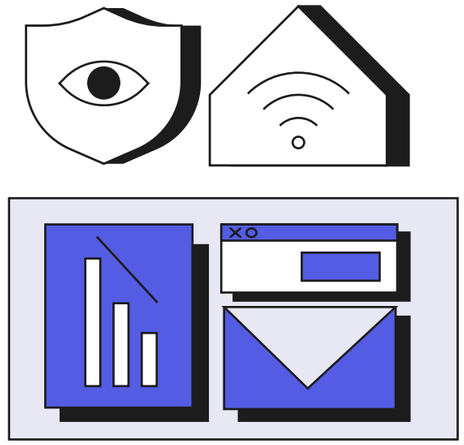
- zwiększenie dostępności danych handlowych przydatnych do prognozowania trendów;
- identyfikacja przewag konkurencyjnych;
- opracowywanie nowych produktów i usług, w tym wysokotechnologicznych;
- budowanie strategii cenowych i analiza rynku;
- optymalizacja procesu obsługi klienta;
- obniżenie kosztów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

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;
- poprawa relacji między konsumentem a różnymi uczestnikami łańcucha wartości;

Nauka, 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

4.1 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 dzięki staraniom państw członkowskich grupy D9+, w tym Polski. 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 rezultacie 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 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. Dostawcy usług przetwarzania danych oraz operatorzy przestrzeni, w których dane się znajdą, zostaną również zobowiązani do przestrzegania reguł dotyczących interoperacyjności silnej formatów rejestrowania danych (Małobęcka-Szwast, 2021). Jednocześnie projekt w sposób niekorzystny ogranicza dostęp sektora publicznego do danych podmiotów prywatnych jedynie do wymagających nadzwyczajnej interwencji przypadków (np. w sytuacjach klęski żywiołowej) oraz zawęża zakres użycia danych przez stronę trzecią, której użytkownik może je przekazać.

Pierwszym przedsięwzięciem o zasięgu sektorowym z planowanych blisko 10 jest projekt stworzenia **wspólnej 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 zakłada powołanie organu "data access body" na szczeblu krajowym, którego zadaniem byłoby udzielanie pozwoleń na dostęp do 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.

Jednym z głównych celów projektu jest wprowadzenie wymogów dotyczących infrastruktury i interoperacyjności silnej obejmujących cały obszar Unii Europejskiej (Bertuzzi & Fortuna, 2022). Projekt jest spójny z wcześniej wymienionymi Aktem w sprawie zarządzania danymi oraz Aktem w sprawie danych, 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 interoperacyjnością silną.

4.2 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, Francja, a także Japonia.

Poniżej przedstawiamy niektóre rozwiązania:

XRoad

kraj pochodzenia: Estonia

- 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;
- Wysokie zabezpieczenia np. nakładki na dane poszczególnych pacjentów, które ujawniają tylko niezbędne informacje.

Open Data Institute

kraj pochodzenia: Wielka Brytania

- 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;
- Rosnące zainteresowanie wymianą pomiędzy przedsiębiorstwami a organami rządowymi (B2G);
- 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;

- Badania w zakresie różnych krajowych porządków legislacyjnych i podejść do modeli dzielenia się danymi – wniosek, że we wszystkich jurysdykcjach pojawiające się debaty na temat rozwoju trustów danych wskazują na znaczenie oparcia się na lokalnych uwarunkowaniach w celu sprostania nowym wyzwaniom w zakresie zarządzania, dostosowania struktur do odpowiedniego kontekstu i celu oraz kontrolowania napięć między interesami indywidualnymi i zbiorowymi.

Daten-Treuhänder

kraj pochodzenia: Niemcy

- Planowany pośrednik dla danych przemysłu motoryzacyjnego oraz ubezpieczycieli;
- Dane udostępniane do pracy nad zwiększeniem bezpieczeństwa na drogach oraz usprawnianiem budowy i naprawy maszyn;
- Dostęp do danych dla sektora publicznego, ubezpieczycieli, związków inspekcji technicznej i centrum serwisowych.

Bundesdruckerei

kraj pochodzenia: Niemcy

- Przedsiębiorstwo państwowe produkujące dokumenty i urządzenia do ich weryfikacji, oferujące także usługi pośredniczenia w przekazywaniu danych;
- Autoryzowanie dostępu do danych, łączenie danych w większe zbiory, monitorowanie zbiorów danych i ich jakości, analiza danych;
- Zabezpieczanie danych poprzez ich pseudonimizację lub anonimizację;
- Użytkownicy CenTrust: Instytut Roberta Kocha (np. baza danych o szczepieniach przeciwko COVID-19), agencje sektora ochrony zdrowia.

Finlandia

kraj pochodzenia: Finlandia

- Fiński Urząd ds. Pozwoleń na Udostępnianie Danych Społecznych i Zdrowotnych;
- Publiczna instytucja zarządzająca dostępem do różnych zbiorów danych 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 i ch kompletnością i odpowiednim wprowadzaniem).

Health Data Hub

kraj pochodzenia: Francja

- 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).

Smart Data Platform with Trust

kraj pochodzenia: Japonia

- Usprawnienie wymiany danych między azjatyckimi firmami;
- Projekt wspierany przez Ministerstwo Gospodarki, Handlu i Przemysłu (METI);
- Informacje dla przedsiębiorstw o zapasach produktów i części, potencjalnych zakłóceniach w zaopatrzeniu itd.;
- Możliwość decydowania przez firmy o tym, jakie dane zamierzają udostępniać;
- W zależności od charakteru konkretnych danych (np. stanowiących tajemnicę przedsiębiorstwa czy chronionych na podstawie prawa autorskiego), możliwość udostępniania danych na różnych warunkach;
- Rozwój infrastruktury wymiany danych na poziomie regionalnym.

5. Modele współdzielenia danych – warsztaty badawcze

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 wspólnic dla danych nieosobowych, biznesowych; publicznych wspólnic dla danych osobowych. Czwarty blok poświęcony był analizie sposobów zarządzania danymi w zależności od stopnia ich wrażliwości.

WYKRES 4. Cele zarządzania i sposoby zarządzania

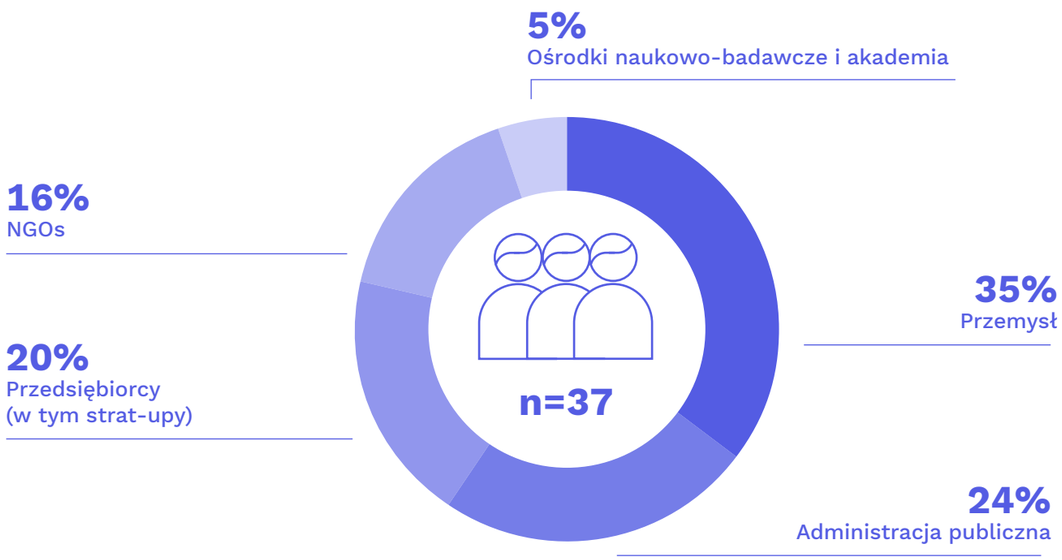
Kontrola interesariuszy / sposoby zarządzania

Alokacja wartości/cele zarządzania	**indywidualna / oparte na uprawnieniach**	**instytucjonalna / oparte na zaufaniu**
prywatna / zysk	Osobiste kapsuły danych	Wirtualne wspólnice danych
publiczna / dobro wspólne	Pośrednicy danych osobowych	Publiczne wspólnice danych

Podział wypracowany podczas warsztatów “Uwolnić potencjał danych”; inspirowany 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 zajmujących się współdzieleniem danych z zagranicy.

WYKRES 5. Uczestnicy warsztatów



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 ze współdzielenia danych, a także bariery wynikające z analizy SLEPT, które potencjalnie hamować mogą współdzielenie danych 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. Przy zachowaniu kompleksowego podejścia do kwestii współdzielenia danych, uwaga autorów skupiła się przede wszystkim na **wspólnicy przemysłowo-rolnej** 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.

5.1 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 pojedyn-

cze 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

propozycja instytucji opisana szczególnie w Data Governance Act jako jedna z usług współdzielenia danych. Zapewniając nadzór i przejrzystość, spółdzielnie danych mają umożliwić osobom fizycznym nie tylko korzystanie z praw danych, ale również udział w zarządzaniu spółdzielnią, zgodnie z zasadą 1 osoba – 1 głos. 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. Z kolei szwajcarska Midata czy hiszpański Salus.coop gromadzą dane zdrowotne i pozwalają spółdzielcom decydować do jakich badań chcą je przekazać i na jakich warunkach.

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.

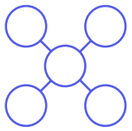
KORZYŚCI

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 (BARIERA EKONOMICZNA)

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.**

BRAK ŚWIADOMOŚCI W ZAKRESIE KORZYŚCI PŁYNĄCYCH Z DZIELENIA SIĘ DANYMI (BARIERA SPOŁECZNO-KULTUROWA)

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). Tak niski wskaźnik wynika m.in. z niewystarczającego poziomu cyfryzacji społeczeństwa. Brak odpowiedniej wiedzy w zakresie technologii przejawia się natomiast nieufnością oraz nieświadomością obywateli w zakresie innowacyjnych koncepcji i rozwiązań takich, jak chociażby współdzielenie danych.

NIEJEDNOLITE STANDARDY (BARIERA TECHNOLOGICZNA)

Zarówno podmioty prywatne, jak i placówki administracji publicznej w Polsce posługują się odmiennymi formatami oraz przewidują niejednolite standardy wymiany danych, utrudniając tym samym ich efektywne współdzielenie.

BRAK ODPOWIEDNIEJ RZĄDOWEJ STRATEGII W ZAKRESIE DZIELENIA SIĘ DANYMI (BARIERA POLITYCZNA/STRATEGICZNA)

Pomimo wysokiego poziomu otwartości danych publicznych, współdzielenie danych pomiędzy podmiotami działającymi zarówno w przestrzeni publicznej, jak i prywatnej jest ograniczone. Równocześnie brakuje na tę chwilę rządowej strategii, na wzór tej dot. Open Data.

PROPONOWANE ROZWIĄZANIA

1

INKUBACJA ATRAKCYJNEGO MODELU BIZNESOWEGO

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 wspólny charakter danych (np. kampanie promujące dofinansowania na szkolenia pracowników w firmach).

3

WYPRACOWANIE JEDNOLITEGO STANDARDU WYMIANY DANYCH

Propozycja wyodrębnienia 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

Niski wskaźnik “ucyfrowienia” naszego kraju przekłada 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.

5.2 Wirtualne wspólnice danych

Wykorzystywanie urządzeń *business intelligence* przez przedsiębiorców staje się coraz bardziej powszechne ze względu na możliwości jakie oferują narzędzia AI w zakresie poprawy efektywności firmy, obniżenia kosztów, podnoszenia jakości 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.

Wirtualna wspólnica (wspólna składnica) **danych** przewiduje formę współpracy, w ramach której aktorzy gospodarczy wymieniają się dostępem do danych (poprzez API; technologię blockchain) w celu tworzenia zbiorowej wartości (Data Collaboratives, 2021). Przyjęta w publikacji **nowa nazwa** dla dawnej 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**. Termin “wspólnica” ma za zadanie uwypuklać wspólny charakter dostępu do danych, “wirtualność” wspólnicy ma natomiast podkreślać jej cyfrową formę oraz federacyjny model wymiany danych za pomocą rozproszonych repozytoriów (w odróżnieniu od repozytorium centralnego, w którym przechowywane są wspólne dane).

Głównym założeniem tego konceptu jest zachęcanie przedsiębiorstw (dużych firm, MŚP, start-upów) do nawiązywania partnerstw w celu wykreowania 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. Następnie, dzięki wykorzystaniu pozytywnych efektów sieciowych możliwe byłoby tworzenie wspólnie bardziej złożonych produktów oraz powiązań.

Tak funkcjonujące systemy 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 (Verhulst, Young i Srinivasan 2022).

DLA JAKICH RODZAJÓW DANYCH?

- dane przemysłowe (wewnętrzne ciągi produkcyjne lub usługowe przedsiębiorstw; dane wytwarzane przez maszyny; związane z konserwacją maszyn; dane o łańcuchach dostaw)
- dane biznesowe (dane zawierające liczbę wizyt oraz czas spędzony na stronie internetowej; analizy big data; dane o logistyce)
- dane rolnicze (dane z czujników; dane z ciągników; dane meteorologiczne)
- dane logistyczne i transportowe
- dane finansowe (transakcje B2B; transakcje C2B)

JAKIE FORMY?

ZAUFANY PODMIOT

Zakłada wybranie lub stworzenie podmiotu pośredniczącego (potencjalnie kontrolowanego przez strony współdzielenia) który dba o standardy techniczne i warunki współpracy, jednak nie przechowuje danych. Tworzenie oraz promowanie zaufanych podmiotów mogłoby odbywać się również w drodze zacieśniania współpracy w ramach UE; Sojuszu Północnoatlantyckiego; krajów Trójmorza; Grupy Wyszehradzkiej, czy Trójkąta Weimarskiego. Dzięki takiemu rozwiązaniu możliwe byłoby tworzenie zaufanych podmiotów nie tylko pomiędzy polskimi przedsiębiorcami, lecz także pomiędzy podmiotami zagranicznymi (Rada Ministrów, 2020). Zaletą umiędzynarodowienia byłaby możliwość łączenia znacznie większej ilości danych w celu generowania wartości; obniżenie kosztów przyszłych transakcji oraz ujednolicenie sposobu wymiany danych pomiędzy przedsiębiorcami na arenie ponadnarodowej. Standardy na rynku mogą być bowiem narzucane lub innowacyjnie wprowadzane nie tylko przez organy regulacyjne, lecz także przez pośredników, w tym międzynarodowych organizacji (Baron et al, 2019).

W przypadku sektorów przemysłowych i rolnych sugerowana jest współpraca stron polegająca na zawiązywaniu spółdzielni danych – tj. dobrowolnych zrzeszeń nieograniczonej liczby podmiotów, które w interesie swoich członków prowadziłyby wspólną działalność na rzecz współdzielenia danych pochodzących z wielu gospodarstw. Alternatywnie rolę zaufanego podmiotu może pełnić agencja publiczna, wspierając powstanie wirtualnej wspólnicy. 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. Trzecim modelem może być konsorcjum kontraktowe z jednolitym regulaminem i standardem dzielenia się danymi w ramach konsorcjum, z zasadą wzajemności pomiędzy członkami co do logiki dostępności do danych.

WSPÓŁPRACA KONTRAKTOWA

Forma polegająca na tworzeniu powiązań pomiędzy partnerami biznesowymi poprzez łączenie posiadanych przez nich zasobów. Mogłaby ona polegać na zobowiązywaniu się stron na podstawie kontraktów prywatnoprawnych (formy kooperacji; współdziałania kilku przedsiębiorstw z ich liderem na czele), jak również umów partnerskich o charakterze publiczno-prywatnym (realizowanych na mocy Ustawy o partnerstwie publiczno-prywatnym z dnia 19 grudnia 2008 r.). Rozwiązanie zbliżone do znanej prawnie puli patentowej, czyli modelu, w którym wynalazcy udostępniają swoje patenty za określoną z góry cenę, a dany projekt może zostać nabyty przez podmioty będące stroną umowy. W przypadku łączenia danych interesariusze z różnych sektorów zawieraliby umowy o udzielenie licencji uprawniającej do korzystania z „puli”, aby wzajemnie dzielić się zasobami w określonym przez siebie zakresie. Pomimo podobieństwa do pul patentowych, w przypadku data-pooling mielibyśmy do czynienia z dużo bardziej złożonymi formami współpracy. Wymagałyby one nie tylko kontraktów na udzielenie licencji dostępowych, ale także dodatkowych umów dotyczących technologii przetwarzania potrzebnej do łączenia przesyłanych danych (Schneider, G., 2020).

W przypadku danych przemysłowych łączenie danych mogłoby dokonywać się poprzez budowanie ekosystemów branżowych i międzybranżowych na platformie internetowej. Model opierałby się na współpracy ograniczonej grupy firm oraz limitowanym dostępie przedsiębiorców do zamkniętych i bezpiecznych środowisk cyfrowych. Data pooling występuje już oraz ma szczególne znaczenie w rolnictwie, dlatego model ten jest godny rozważenia podczas tworzenia pilotaży dla wspólnicy danych rolniczych. Wskazuje się bowiem, że dzięki łączeniu w pule 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).

KORZYŚCI

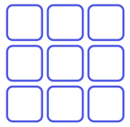
WYRÓWNYWANIE SZANS

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 w tym także w rolnictwie. Algorytmy podejmowania decyzji, systemy Digital Farming, teledetekcja,

precyzyjne aplikowanie nawozów na podstawie analiz gleby i wielkości plonów, śledzenie w czasie rzeczywistym plonu i oceny efektu nawożenia pozwalają nie tylko o dbanie o poziom produktywności i stanu całego przedsiębiorstwa, lecz także o monitorowanie kondycji pojedynczych roślin (Rada Ministrów, 2020). **Sfederowane analizy rozproszonych danych dają możliwość uwspólniania uzyskanych z narzędzi AI wyników bez konieczności udostępniania oryginalnych danych**. Tym samym, możliwe jest zagwarantowanie bezpieczeństwa danych pochodzących z pojedynczych gospodarstw oraz zapewnienie równowagi pomiędzy prywatnością, autonomią, ochroną własności intelektualnej (Big Data Value Association, 2019) i wolnością.



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.

WZROST SKALI PRODUKTYWNOŚCI DLA CAŁEJ GOSPODARKI ORAZ POPRAWA KONKURENCJI

Jak już zostało wykazane, wielokrotne wykorzystanie danych daje ogrom możliwości. Jednak do tej pory najcenniejsze informacje cyfrowe ograniczane są **tajemnicami handlowymi największych platform – pomimo iż w większości nie stanowią one ich wytworu, a jedynie reprezentują serię zjawisk zachodzących w przestrzeni internetowej**. Obowiązek współdzielenia danych wprowadzony dla najbardziej newralgicznych społecznie sektorów mógłby zatem podnieść 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). Dane pochodzące z rolnictwa mogłyby być wykorzystywane także przez państwowe instytucje badawcze do uzyskiwania dokładniejszych wyników o stanie gospodarw w kraju oraz pozwalałyby lepiej zrozumieć lokalne praktyki rolnicze, co w dalszej kolejności pozwoliłoby efektywniej planować politykę państwa w zakresie rolnictwa (GPAI, 2021).

Na podstawie warsztatu przeprowadzonego z przedstawicielami sektora przemysłowo-rolnego

BARIERY

OBAWY WZGLĘDEM DZIELENIA SIĘ DANYMI (BARIERA SPOŁECZNO-KULTUROWA)

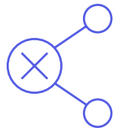
Uczestnicy warsztatów wskazali na pojawiający się wśród przedsiębiorców strach przed wyciekami informacji, a także niewłaściwym wykorzystaniem danych przez podmioty zewnętrzne. W przypadku rozwiązań chmurowych, rolnicy mają obawy przede wszystkim o bezpieczeństwo swoich danych, ze względu na przekonanie, że lepiej przechowywać dane w przestrzeni swojego komputera.

NIECHĘĆ DO WPROWADZANIA ZMIAN POWSZECHNA SZCZEGÓLNIE WŚRÓD STARSZEGO POKOLENIA ROLNIKÓW (BARIERA SPOŁECZNO-KULTUROWA)

Zauważono, iż w rolnictwie panuje silne przywiązanie do tradycji, które skutkuje sceptycznym podejściem do nowych sposobów zarządzania gospodarstwem oraz wykorzystywania innowacyjnych rozwiązań.

NIEUFNOŚĆ W STOSUNKU DO NOWYCH TECHNOLOGII (SPOŁECZNO-KULTUROWE)

Rolnictwo to obszar charakteryzujący się niewielkim nasyceniem technologią w porównaniu do innych sektorów gospodarki. Jak wynika z badań przeprowadzonych przez Uniwersytet Rolniczy w Krakowie, nawet wśród młodych rolników, dostrzegających potencjał w unowocześnianiu gospodarstw, widoczna jest zachowawczość przed wdrażaniem innowacji na szeroką skalę (B. Kiełbasa, J. Puchała, 2015).



INTERNETOWE BIAŁE PLAMY ORAZ BRAK ODPOWIEDNIEJ INFRASTRUKTURY (TECHNICZNE)

Ponad połowa osób (55 proc.), które nigdy nie korzystały z sieci, mieszka na obszarach wiejskich (Bartol, A., Herbst, J., Pierścińska, A., 2021). Wśród wszystkich mieszkańców wsi grupą szczególnie zagrożoną wykluczeniem cyfrowym są osoby starsze. Tym samym, struktura demograficzna rolników (niewielka liczba młodych), jak i położenie gospodarstw na obszarach wiejskich może skutkować trudnościami w sięganiu po nowe rozwiązania cyfrowe. Nawet jeżeli rolnicy są chętni, aby wykorzystywać najnowocześniejsze technologie, napotykają oni problemy związane z niedostatecznym pokryciem kraju Internetem szerokopasmowym, rozproszenie baz danych, brak odpowiednich narzędzi do ich obsługi, czy niedostateczną liczbę czujników.

BRAK INTEROPERACYJNOŚCI SILNEJ SYSTEMÓW I STANDARDÓW WYMIANY DANYCH (TECHNICZNE)

Uczestnicy warsztatów zwrócili uwagę na odmienne sposoby zbierania danych, zróżnicowane formaty agregowania danych wykorzystywane przez maszyny oraz brak interoperacyjności pomiędzy systemami.

OBAWY O UTRATĘ KONKURENCYJNOŚCI; POWOŁYWANIE SIĘ NA TAJEMNICĘ PRZEDSIĘBIORSTWA (SPOŁECZNO-KULTUROWE/PRAWNE)

Widocznym jest, że wśród przedsiębiorców dalej panuje przekonanie, że udostępnienie ich danych innym podmiotom działającym w tym samym bądź innym sektorze może skutkować dla nich utratą konkurencyjności i pogorszeniem się ich pozycji na rynku.

5.3 Publiczne wspólnice danych

Model publicznych wspólnic danych stanowi alternatywę dla korporacyjnych silosów danych, zapewniając uczciwą kooperację między zaangażowanymi podmiotami i świadome uspołecznianie wytwarzanej wartości. Dane cyfrowe najpierw bowiem są zapisem rzeczywistości (cyfrową kopią), jednak w środowisku informacyjnym zyskują naturalne cechy dóbr wspólnych – mogą być niezwykle prosto replikowane, przy czym trzeba dbać o ich zrównoważone użycie. Głównym celem publicznej wspólnicy danych jest stworzenie ekosystemu zaufania dla prospołecznego i proinnowacyjnego wykorzystania danych w oparciu o zasady współzarządzania i ustaloną hierarchię wartości (Zygmuntowski, 2020a). Jest to zatem instytucja swoistego banku danych czy infrastruktury dla społeczeństwa informacyjnego, łącząca państwową wagę z oddolnym udziałem we wspólnej kreacji wartości.

Z wartości płynącej z agregacji tych danych korzysta przede wszystkim określona społeczność – lokalna, regionalna czy też paneuropejska. 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. Ważnym aspektem funkcjonowania publicznych wspólnic danych jest model współzarządczy, inkluzywność i partycypacja. Doskonałym przykładem takiego działania może być model brytyjskiego NICE (National Institute for Health and Care Excellence), pozarządowej instytucji publicznej ustalającej wytyczne zdrowotne przy udziale rad otwartych dla interesariuszy.

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?

WSPÓŁZARZĄDZANA INSTYTUCJA PUBLICZNA

Operatorem repozytorium, infrastruktury i standardów jest agencja publiczna z zakresu danego sektora, przykładowo Centrum eZdrowia dla danych zdrowotnych. Operator dba również o jakość danych, promuje ich dalsze współdzielenie i udziela zezwoleń na dostęp (na wzór europejskiej koncepcji *data access body*). Współzarządzanie (np. w postaci rady nadzorczej) gwarantuje, że dostęp do danych, warunki techniczne i kierunki rozwoju wspólnicy danych są ustalane przez reprezentację społeczną i ryzyko nadużycia przez państwo jest minimalizowane. Wprowadzenie jednej, silnej instytucji stojącej na straży zarówno interesu publicznego, jak i praw osób, których dane dotyczą, wzmacnia pozycję obywateli, potencjalnie umożliwiając im korzystanie ze swoich praw nie w formie zgłoszeń naruszeń *(ex post*), ale bezpośrednio w formie usługi publicznej (Zygmuntowski et al., 2021).

WSPÓŁZARZĄDZANY OŚRODEK NAUKOWY

Instytucje badawcze, uczelnie prywatne i publiczne oraz centra B+R już dzisiaj zarządzają repozytoriami naukowymi i mają rozbudowane kompetencje w zakresie zarządzania danymi. Cieszą się również znacznym zaufaniem społecznym, a w kontekście niektórych typów wrażliwych danych (np. danych medycznych) mają uprawnienia dostępu większe niż nienaukowe

przedsiębiorstwa gospodarcze. Dlatego operatorem wspólnicy może być ośrodek naukowy, jednak również przy założeniu, że nadzór nad danymi będzie sprawowany przez pozanaukowe grono reprezentujące innych interesariuszy, w tym samą administrację publiczną.

OPIS STRUKTURY

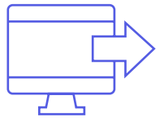
Podstawową zasadą wspólnic danych osobowych powinna być transparentność zamiarów podmiotu chcącego wykorzystać zebrane dane do własnych badań. Ich wyniki powinny zostać udostępnione publicznie lub na innych zasadach chroniących interes publiczny, a ochrona interesów jednostek wymaga, by każdorazowo przeprowadzać ocenę skutku algorytmu. Podmioty, które już na początkowym etapie nie spełniałyby ustanowionych zasad etycznych i reguł wzajemności 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.

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. Firmy posiadające 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 zarządów, jak i centra danych zlokalizowane są na terenie EOG. W braku takiej możliwości, ważne jest ustanowienie odpowiednich, nadrzędnych reguł swobodnego przepływu danych między różnymi podmiotami podlegającymi rozmaitym porządkom prawnym, ustanawiając sprawnie działający, międzynarodowy metasystem oceny zgodności zarówno regulacji wewnętrznych danego dostawcy, jak i przepisów krajowych z ustalonymi zasadami.

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 jeśli dotyczy

to danych wyższego rzędu niż dane surowe czy ustrukturyzowane. 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 szeroko rozumiany, publiczny charakter wspólnic, pozyskiwanie środków z grantów organizowanych przez publiczne instytucje czy organizacje pozarządowe również wydaje się być dobrym rozwiązaniem.

KORZYŚCI

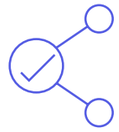


PODNOSZENIE ZAUFANIA DO WSPÓŁDZIELENIA DANYCH

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 zdrowia

BARIERY

BRAK UJEDNOLICONYCH STANDARDÓW (TECHNICZNE)

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).

WIELOŚĆ WYKORZYSTYWANYCH SYSTEMÓW W SEKTORZE ZDROWIA* (TECHNICZNE)

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). Nie istnieje natomiast żadna regulacja dotycząca uspójnienia standardów dla danych zdrowotnych i wspierająca utworzenie tzw. interoperacyjności silnej w ramach całego sektora ochrony zdrowia, a więc zarówno placówek publicznych jak i prywatnych. Z uwagi na brak kompatybilności systemów, pacjenci często zmuszeni są do samodzielnego przenoszenia danych o stanie swojego zdrowia z placówek prywatnych do publicznych i odwrotnie, generując tym samym wyższe koszty diganostyki i swojego leczenia.

ODMIENNA KULTURA ZBIERANIA DANYCH W RÓŻNYCH PLACÓWKACH* (SPOŁECZNO-KULTUROWE/TECHNICZNE)

Brak odpowiednich regulacji dotyczących standardów i opisywania konkretnych przypadków skutkuje często odmiennymi sposobami zbierania, oznaczania i systematyzowania danych w placówkach ochrony zdrowia. Przykładem tego stanu rzeczy może być sytuacja odmiennego oznaczania krwi, która kiedyś miała miejsce w szpitalach i mogła powodować liczne konfuzje.

NIEMOŻLIWOŚĆ ZAPEWNIENIA CAŁKOWITEGO UTAJNIENIA INFORMACJI OSOBOWYCH* (TECHNICZNE)

Ze względu na specyfikę pseudonimizacji danych, a więc techniki szyfrowania danych za pomocą odrębnie trzymanych haseł, niemożliwe jest zapewnienie stuprocentowego bezpieczeństwa. Jeżeli ktoś posiada dostęp do kluczy umożliwiających "odkodowanie" danych tego rodzaju, zawsze będzie istniało ryzyko i pewna doza prawdopodobieństwa, że dane te staną się całkowicie dostępne dla nieuprawnionych osób. Ponadto, ponieważ dane dotyczące zdrowia są często złożone z wielu różnych informacji, trudno jest ocenić, w braku których z nich dalsza identyfikacja podmiotu danych nie jest już możliwa, a które w zestawieniu z innymi posiadanymi danymi umożliwiają wskazanie osoby, której dane dotyczą.

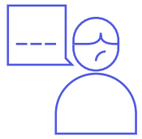
NIEUFNOŚĆ DO UDOSTĘPNIANIA DANYCH DOTYCZĄCYCH ZDROWIA (SPOŁECZNO-KULTUROWE)

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, J., Łukasik, K., Święcicki, I., 2021). 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 skłonnych było domniemywać, że w rzeczywistości regulacja nastawiona jest na sprawowanie kontroli nad obywatelkami (PAP, 2021).

BRAK WZAJEMNEGO ZAUFANIA AKTORÓW UCZESTNICZĄCYCH W PROCESIE WSPÓŁDZIELENIA DANYCH* (SPOŁECZNO-KULTUROWE)

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). Co więcej, pomimo prób współpracy między tymi podmiotami, również nie darzą się one wzajemnym zaufaniem.



BRAK ZROZUMIENIA RELACJI POMIĘDZY “INTERESEM PUBLICZNYM” A PRYWATNYM ORAZ PŁYNĄCA Z TEGO TRUDNOŚĆ W OKREŚLENIU INDYWIDUALNEJ KORZYŚCI WSPÓŁDZIELENIA DANYCH (SPOŁECZNO-KULTUROWE)

Interes publiczny, zarówno w prawie jak i w ogólnym odbiorze społecznym, kojarzony jest zwykle z ograniczaniem podstawowych praw i wolności jednostek na rzecz dobra przysługującemu określonej wspólnocie (np. bezpieczeństwa, zdrowia lub porządku publicznego). Ponadto, wielu osobom interes publiczny kojarzy się z władzą rządową i możliwością zaistnienia niepożądanej inwigilacji ze strony organów publicznych. Wobec tak zarysowanej panoramy skojarzeń, wśród wielu osób rodzi się pytanie o to, jaka jest korzyść dla pojedynczego obywatela z dzielenia się swoimi danymi z szerokim gronem nieznanych podmiotów. Jak pokazują bowiem badania przeprowadzone przez Polski Instytut Ekonomiczny, Polacy najchętniej dzielą się danymi ze swoimi najbliższymi i pielęgniarką (Grzeszak, J., Łukasik, K., Święcicki, I., 2021).

BRAK JEDNOZNACZNIE OKREŚLONYCH DEFINICJI CELÓW NAUKOWYCH* (PRAWNE)

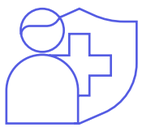
Chociaż art. 9 ust. 2 lit. j) RODO wskazuje na “cele naukowe” jako jedną z podstaw przetwarzania szczególnych kategorii danych, nie wskazuje ich definicji legalnej. Zgodnie z motywem 159 RODO, do “celów naukowych” należą:

- 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, dopuszczając również ośrodki prywatne. 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.

Wobec pozostawienia państwom członkowskim pewnej swobody regulacyjnej, w polskim ustawodawstwie przyjęto szersze pojęcie odwołujące się do prac badawczo-rozwojowych i działalności naukowej obejmującej m.in. badania naukowe oraz prace rozwojowe (Prawo o szkolnictwie wyższym i nauce). Brak możliwości określenia, jak polska definicja odnosi się do tej zawartej w RODO sprawia, że określenie, które z wyznaczonych celów należą do dozwolonych celów naukowych, staje się znacznie utrudniona.

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 i innych danych umożliwiających identyfikację osoby, której dokumentacja dotyczy. Oznacza to, że inne podmioty, takie jak prywatne przedsiębiorstwa i organizacje niewymienione w wykazie, nie mogą korzystać z możliwości ograniczenia stosowania RODO na tych samych zasadach, które przysługują uczelniom i instytutom badawczym (a więc z pominięciem art. 15 RODO traktującego o prawie dostępu do danych osobom, których dane dotyczą) (Najbuk, P., Pachocki, J., Kruczyk-Gonciarz, A., Kaźmierczyk, P. Lorent, R. 2020). Podmioty prywatne mogą jednak korzystać z art. 9 ust. 2 lit j RODO i prowadzić działalność badawczą, nie pomijając przy tym jednak innych przepisów.



BRAK DEFINICJI LEGALNEJ DANYCH DOTYCZĄCYCH ZDROWIA I ICH RELACJI WOBEC DANYCH MEDYCZNYCH (PRAWNE)

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. Brak jednoznacznej definicji, zarówno w prawie unijnym jak i prawie polskim powoduje, że bardzo wiele danych będących na pograniczu “zwy-

kłych” danych i tych o szczególnym charakterze, może zostać automatycznie traktowane z mniejszą dozą ostrożności, niż byłoby to w przypadku danych wrażliwych.

ZAMKNIĘCIE DANYCH W BAZACH PRYWATNYCH FIRM (PRAWNE)

Na rynek medyczny z coraz większym powodzeniem wchodzą duże koncerny technologiczny (m.in. Amazon, Google, Apple), oferując swoim użytkownikom wygodne rozwiązania służące do pomiarów codziennych czynności i wyciągania na tej podstawie wniosków o stanie ich zdrowia. Firmy te robią to jednak dla swoich własnych, komercyjnych celów, nie chcąc się dzielić danymi z samorządami i innymi podmiotami działającymi na rzecz dobra publicznego.

5.4 Metody zarządzania danymi szczególnej wrażliwości

5.4.1. Dane nieosobowe

Tym samym, czerpanie korzyści z cyfrowych zbiorów przez jednego przedsiębiorcę nie wyklucza ich użyteczności dla innego podmiotu (Paszcza, 2022). Jednak pomimo możliwości wielokrotnego wykorzystywania danych, w dalszym ciągu ich przekazywanie i re-użycie napotyka wiele barier.

Dane nieosobowe rozumiane są jako informacje elektroniczne inne niż te wskazane w RODO – tj. nie dotyczą one 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. Bywa, że dostęp do 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. Jednak w odniesieniu do kwestii dostępności danych nieosobowych, wskazuje się, że w przypadku, gdy ich źródłem jest otoczenie/środowisko naturalne bądź jeśli zostały one zanonimizowane (np. dane dot. konsumentów), nie powinny być one ograniczane i monopolizowane (Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H. (2017). Dane pochodzące z wewnętrznych ciągów produkcyjnych choć często stanowią przedmiot posiadania przedsiębiorstw, mogłyby być poddawane szerszemu wykorzystywaniu poprzez re-use przemysłowy z uwagi na ich pobudzający wpływ na całość ekosystemu gospodarczego (Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H. (2017).

Bywa jednak, iż przedsiębiorcy odmawiają dostępu do zagregowanych przez nich zbiorów danych powołując się najczęściej na kwestie prywatności (tajemnic handlowych) oraz praw własności intelektualnej. Jeżeli chodzi o tajemnicę przedsiębiorstwa, to w przypadku danych nieosobowych możemy mówić o informacjach pochodzących ze stron internetowych, urządzeń czy czujników maszyn, przyjmujących postać np. tekstu, liczb lub obrazów, ustrukturyzowanych lub nieuporządkowanych. Skorzystanie z prawa do ochrony tych danych wymaga jednak **odpowiedniego zabezpieczenia zbiorów** przed dostępem osób nieuprawnionych i posiadania wartości gospodarczej lub przynależności do zbioru lub zestawienia, który taką wartość reprezentuje[1]. Pomimo, iż część informacji cyfrowych znajdujących się w zasobach firm faktycznie spełnia wyżej wymienione przesłanki, wielu przedsiębiorców nadużywa uprawnień dotyczących tajemnic handlowych wynikających z ustawy o zwalczaniu nieuczciwej konkurencji hamując tym samym rozwój nie tylko ich własnej firmy, ale także całej gospodarki.

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 przysługuje ono producentom baz danych, o tyle często uprawnienia z niego wynikające są nadużywane przez administratorów danych. Wspomniana już ochrona *sui generis* przysługuje bowiem 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).

Podsumowując, dane z systemów wewnętrznych firm, co do zasady będą przedmiotem ochrony tych przedsiębiorstw. Jednak z uwagi na ich duży potencjał gospodarczy oraz możliwość rozwijania łańcuchów wartości wskutek wzajemnej wymiany informacji, koncepcja wzajemnego udzielania dostępu do danych powinna być propagowana i wdrażana na szeroką skalę (Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H. 2017). Należy także przyjrzeć się występującym barierom dla wymiany danych pomiędzy róż-

1 W prawie polskim zagadnienia dotyczące tajemnicy przedsiębiorstwa uregulowane zostały 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

nymi aktorami gospodarczymi. Nadużywanie przez przedsiębiorców prawa do ochrony baz danych oraz praw własności intelektualnej najczęściej wynika z niskiej świadomości na temat korzyści płynących z współdzielenia danych w biznesie oraz obaw o utratę konkurencyjności firmy. Jednak pojawiają się też głosy przedstawicieli biznesu, którzy wskazują na niepewność w zakresie bezpieczeństwa udostępnianych danych oraz strach przed ich ewentualnym wyciekiem oraz niewłaściwym wykorzystaniem. Budowa systemu IT pozwalającego w efektywny sposób współdzielić dane przy równoczesnym zapewnianiu maksimum bezpieczeństwa i możliwości kontroli jest utrudniona ze względu na różne uwarunkowania prawne i techniczne baz danych (Dymek, Komnata, Kotulski, 2011). Jednak nie jest to zadanie niewykonalne. Dzięki różnego rodzaju rozwiązaniom technicznym możliwym jest bowiem zniwelowanie ryzyk związanych z wymianą danych, a także zapewnienie poufności danych oraz konkurencyjności poszczególnych podmiotów.

FEDERACYJNY MODEL WIRTUALNEJ WSPÓLNICY DANYCH BIZNESOWYCH

Analizowany w niniejszym raporcie model wirtualnej wspólnicy opierać miałby się na federacyjnej strukturze, co samo w sobie stanowiłoby swojego rodzaju zabezpieczenie. Tak zaprojektowany system pozwala bowiem na wzajemne udostępnianie sobie danych, lecz jedynie w zakresie wynikającym z potrzeb wspólnicy, przy zachowaniu wszelkich wymogów bezpieczeństwa oraz bez konieczności ujawniania szczegółów dotyczących budowy poszczególnych baz danych. Dodatkowo umożliwia on zapewnienie nadzoru nad procesami wymiany danych zarówno na poziomie wewnętrznym, jak i zewnętrznym (na poziomie wspólnicy). Koncepcję tę można porównać do bezpiecznego przesyłania danych poufnych pomiędzy organami ochrony prawa, która opiera się na trybie pytanie-odpowiedź. W trybie tym, dzięki wykorzystaniu ujednoliconych formularzy zapytań, organ pytający nie musi znać modelu i struktur baz danych organu pytanego (Dymek, Komnata, Kotulski, 2011). Podobnie mogłoby to wyglądać w przypadku wirtualnej wspólnicy danych biznezowych – przedsiębiorca przekazywałby podmiotowi obsługującemu wspólnicę jedynie wystandaryzowane, surowe dane, bez konieczności ujawniania wyników analizy danych przeprowadzanej na rzecz jego firmy.

UCZENIE FEDERACYJNE

Federated Learning jawi się jako nowy paradygmat współpracy i partnerstwa między przedsiębiorstwami. Umożliwia ono firmom udostępnianie ich zbiorów w „systemie zamkniętym" oraz pozwala na budowanie wspólnego, wydajnego modelu uczenia maszynowego bez konieczności faktycznej wymiany danych. Uczenie federacyjne zapewnia bowiem możliwość "trenowania" algorytmu przy równoczesnym przechowywaniu danych na poziomie urządzeń poszczególnych podmiotów. Przedsiębiorca zachowuje swoje prywatne, lokalne dane na urządzeniach firmowych, dzięki czemu minimalizuje obawy o ich bezpieczeństwo. Tym samym, ze względu na fakt, iż w przypadku uczenia federacyjnego, wszystkie dane wymagane do trenowania modelu pozostają pod ścisłym nadzorem organizacji, model ten może być stosowany w licznych sektorach, takich jak służba zdrowia, przemysł i e-commerce. Model ten może stanowić warstwę techniczną współpracy w ramach wirtualnych wspólnic danych.

PRZETWARZANIE BRZEGOWE

Przetwarzanie brzegowe to podejście oparte na rozproszeniu wynikającym z możliwości przetwarzanie danych bezpośrednio na inteligentnych urządzeniach, np. telefonach komórkowych i sieciach. Opracowywanie danych odbywa się lokalnie – "brama brzegowa" przetwarza dane z urządzenia, a dopiero po ich ustrukturyzowaniu, przesyła odpowiednie dane do serwera w celu ich dalszego przechowywania. Tym samym, dzięki takiemu rozwiązaniu możliwe jest przetwarzanie danych maksymalnie blisko "źródła", co pozwala na przyspieszenie reakcji urządzenia i zmniejszenie opóźnień w działaniu (Velotio Technologies, 2019). Należy jednak zaznaczyć, że ze względu na wielość indywidualnych urządzeń połączonych z serwerem, konieczne jest zachowanie szczególnych środków ostrożności. Aby firmy mogły zagwarantować pełne bezpieczeństwo danych powinny zapewniać szyfrowanie każdego elementu danych przechowywanych na urządzeniach firmy. Warto także, aby zawsze upewniały się, że łączność wykorzystuje wielostopniowe uwierzytelnianie oraz certyfikaty SSL/TLS lub podobne zabezpieczenia na poziomie korporacji.

Edge computing mogłoby znaleźć swoje zastosowanie w przypadku wspólnicy danych przemysłowych i rolniczych, ponieważ obejmuje szeroki zakres narzędzi cyfrowych, takich jak czujniki, sensory połączone w sieci (pozostają one połączone z centralnym serwerem oraz przesyłają informacje w czasie rzeczywistym). Na przetwarzaniu brzegowym mogłoby także znacznie skorzystać uczenie maszynowe wykorzystywane w przypadku wielu narzędzi biznesowych. Proponuje się bowiem, aby szkolenie algorytmów wykonywać w chmurze, a następnie przygotowane modele wdrażać na urządzeniach brzegowych w celu prognozowania zjawisk.

TECHNOLOGIA BLOCKCHAIN

Blockchain i technologie rozproszonego rejestru (DLT) umożliwiają organizacjom weryfikację stron i transparentne sprawdzenie uprawnień dostępu do danych dzięki zastosowaniu metod kryptograficznych (Bechtel, Buchholz, 2022). Korzystający z tego rozwiązania mogą przeglądać historię operacji zapisaną w rejestrze (np. udzielonych zgód i dostępów), jednak dostęp do danych wymaga posiadania odpowiednich uprawnień. Jakiekolwiek próby manipulacji czy przeprowadzenia nieautoryzowanej operacji są natychmiast wykrywane i odrzucane (system nie dopuszczają ich uwzględnienia w łańcuchu blokowym). Jedynym wektorem ataku może być manipulacja rejestru, ale to wymaga kontroli nad 51% węzłów potwierdzających rejestr, co jest dużo trudniejsze niż przejęcie kontroli nad pojedynczym serwerem. Co więcej, blockchain charakteryzuje się dużą odpornością na awarie infrastruktury IT, ze względu na fakt, iż dane są zapisywane jednocześnie w pamięci nie jednego, a wielu tysięcy serwerów. Jeżeli chodzi o wykorzystywanie tego rozwiązania technologicznego w biznesie, dobrym przykładem jest branża transportowa i logistyczna, w której dzięki systemowi możliwe jest prowadzenie rejestru kierowców, współdzielenie przejazdów, przechowywanie informacji o historii pojazdów. Możliwe jest także kontrolowanie wysyłek w czasie rzeczywistym oraz wymienianie informacji pomiędzy wszystkimi uczestnikami łańcucha dostaw przy równoczesnym zapewnieniu bezpieczeństwa i transparentności.

Technologia blockchain mogłaby więc zyskać szczególne znaczenie w kontekście projektowania wirtualnej wspólnicy dla danych biznesowych – dzięki łańcuchowi blokowemu firmom łatwiej byłoby zarządzać danymi, certyfikatami, stworzonymi wspólnie cyfrowymi produktami, a równocześnie zmniejszeniu uległyby obawy o utratę kontroli nad danymi.

5.4.2 Dane wrażliwe

Dane wrażliwe wymagają szczególnie wzmożonej ochrony ze względu na często poufny i intymny charakter informacji o osobie, której te dane dotyczą. Ich przetwarzanie może być dość poważną ingerencją w prywatną sferę życia człowieka, stanowiąc jednocześnie podstawę do dyskryminacji takiej osoby (Fajgielski, 2021).

RODO w sposób enumeratywny wymienia kategorie danych uznawane za szczególne kategorie danych:

- dane dotyczące pochodzenia rasowego/etnicznego,
- dane dotyczące poglądów politycznych,
- dane dotyczące przekonań religijnych i światopoglądowych,
- dane dotyczące przynależności do związków zawodowych,
- dane genetyczne,
- dane biometryczne przetwarzane w celu jednoznacznego zidentyfikowania osoby fizycznej
- dane dotyczące zdrowia,
- dane dotyczące seksualności lub orientacji seksualnej

W przeciwieństwie do poprzedzającej RODO dyrektywy o ochronie danych osobowych, państwa członkowskie nie mogą rozszerzyć katalogu tych danych i objąć innych rodzajów takim samym reżimem, jaki posiadają one w RODO. Co do niektórych danych natomiast, mogą one przyjąć wyjątkowe regulacje chronionące 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).

Co do zasady, przetwarzanie danych wrażliwych jest zakazane, chyba że zachodzi jedna z dziewięciu przesłanek wymienionych w RODO. Wśród tych, które mogą mieć znaczenie w przypadku modeli współdzielenia danych znajdują się przede wszystkim:

- **Wyraźna zgoda podmiotu**, którego dane dotyczą;
- **Niezbędność przetwarzania do wypełnienia obowiązków i wykonywania szczególnych praw przez administratora** lub osobę, której dane dotyczą, w dziedzinie prawa pracy, zabezpieczenia społecznego i ochrony socjalnej;
- **Niezbędność przetwarzania do ochrony żywotnych interesów osoby**, której dane dotyczą, lub innej osoby fizycznej, a osoba, której dane dotyczą, jest fizycznie lub prawnie niezdolna do wyrażenia zgody na przetwarzanie danych;
- **Niezbędność przetwarzania ze względów związanych z ważnym interesem publicznym** pod warunkiem, że są proporcjonalne do wyznaczonego celu, nie naruszają istoty prawa do ochrony danych i przewidują odpowiednie i konkretne środki ochrony praw podstawowych i interesów osoby, której dane dotyczą;
- **Niezbędność przetwarzania do celów archiwalnych w interesie publicznym, do celów badań naukowych lub historycznych lub do celów statystycznych** pod warunkiem, że są proporcjonalne do wyznaczonego celu, nie naruszają istoty prawa do ochrony danych i przewidują odpowiednie, konkretne środki ochrony praw podstawowych i interesów osoby, której dane dotyczą;

Istotną podstawą w kontekście modelu przetwarzania szczególnych kategorii danych, zwłaszcza w początkowej fazie tworzenia jego infrastruktury, jest przesłanka wyraźnej zgody podmiotu, którego dane dotyczą. Nie powinna ona 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.

Administrator powinien również wykorzystywać jedynie te dane, które są stosowne do danego celu i ograniczone do tego co niezbędne, wybierając te zestawy, które są mu potrzebne do przetwarzania (zasada minimalizacji danych). W modelu współdzielenia szczególnych kategorii danych niezwykle ważne będzie transparentne określenie planowanych przedsięwzięć i jasnych powodów, dla których dane są gromadzone jeszcze przed etapem wyrażania zgody przez osobę udostępniającą swoje dane. W przypadku chęci rozszerzenia celów, konieczne będzie również uprzednie poinformowanie o tym fakcie osób, których dane dotyczą.

Ciekawym rozwiązaniem jest **fińska regulacja typu 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 służących do ulepszania ochrony zdrowia, takie rozwiązanie wydaje się być najbardziej efektywnym.

Pomimo nałożonych przez RODO wymagań i wyzwań z tym związanych, możliwe jest zapewnienie takiej architektury modelu współdzielenia danych,

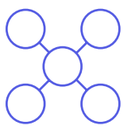
by zapewnić odpowiednią ochronę prywatności i bezpieczeństwo danych już w fazie projektowania (*privacy-by-design*). W tym celu należy jednak stworzyć warunki zapewniające odpowiednie **bezpieczeństwo wewnętrzne systemu** (polegające na odpowiednim utajnieniu danych), oraz **zewnętrzne**, dotyczące odporności systemu na podmioty niespełniające własnych standardów wyznaczonych przez dany model współdzielenia danych, oraz zasad wynikających z prawa unijnego.

CHMURA PRYWATNA

Dane wrażliwe wymagają zapewnienia wyższych standardów bezpieczeństwa zarówno w zakresie ich gromadzenia, jak i późniejszego wykorzystywania. Ze względu na ich wyjątkowo poufny charakter, ważna jest odpowiednia kontrola nad systemem i możliwość wyboru elementów infrastruktury i zabezpieczeń. Z tej przyczyny, rekomendowane jest przechowywanie danych wrażliwych w chmurze prywatnej, zapewniającej możliwość ograniczenia dostępu do jej zasobów.

Zgodnie z badaniami przeprowadzonymi przez VMware, głównym powodem wybierania chmury prywatnej przez organizacje jest możliwość kontroli nad danymi (VMware, 2017). Dzięki chmurze prywatnej, z jej zasobów mogą korzystać jedynie uprawnieni użytkownicy, którzy albo należą do organizacji zarządzającej gromadzonymi zasobami, albo otrzymali zezwolenie na korzystanie z danych w określony sposób. Pomimo braku ogólnej dostępności, chmura prywatna posiada kilka cech zbliżonych do chmury publicznej, m.in. elastyczność i skalowalność (OVH, 2018). Posiada też kilka własnych, szczególnych zalet, m.in. niski próg wejścia i brak nakładów inwestycyjnych. Jest to dobre rozwiązanie dla organizacji, które chcą korzystać z mocy obliczeniowej dostawcy usług chmurowych, przy jednoczesnej możliwości opracowania własnych, autorskich rozwiązań.

Wybierając infrastrukturę, należy mieć na uwadze również posiadane przez dostawcę i oferowane przez niego usługi certyfikaty zapewniające, że istnieje system zarządzania bezpieczeństwem informacji (SZBI) do zarządzania ryzykiem, podatnościami i ciągłości działania oraz systemu zarządzania informacjami o prywatności (OVH).



SCENTRALIZOWANA ORAZ ROZPROSZONA ARCHITEKTURA DANYCH

Pomimo powszechnego poglądu, że scentralizowany system baz danych oraz system rozproszony stanowią dwa rozłączne sposoby organizacji architektury zarządzania zasobami danych, w rzeczywistości mogą się one wzajemnie uzupełniać. Z jednej strony bowiem, system scentralizowany polega na przechowywaniu wszystkich danych w ramach jednej bazy, dzięki czemu jest łatwiejszy w zaprojektowaniu i zarządzaniu, a w razie jakiejkolwiek awarii systemu łatwiej jest odtworzyć jego stan na podstawie posiadanej kopii zapasowej zachowując możliwie najpełniejszą spójność. Z drugiej strony, posiadanie jednego dużego centrum danych dla dużego obszaru geograficznego narażone jest na awarie łączy komunikacyjnych. Z tego względu bardziej naturalnym rozwiązaniem w tej sytuacji byłoby stworzenie systemu rozproszonego, który dodatkowo wyróżnia większa moc przetwarzania, niezwykle ważna w kontekście przetwarzania dużych zbiorów danych (Wojciechowski, 1998).

Skonstruowanie architektury o charakterze hybrydowym, posiadającej elementy zaczerpnięte zarówno ze scentralizowanego jak i rozproszonego modelu jest rekomendowanym rozwiązaniem szczególnie w sytuacji, gdy przechowywanie danych na urządzeniach krańcowych może stanowić ryzyko dla ochrony prywatności lub jest znacznie utrudnione (Zygmuntowski, 2021a). Dzięki temu, dane mogą być gromadzone na jednym lub kilku serwerach należących do organizacji, strzegąc przyjętych warunków ich zbierania i przetwarzania.

MIT OPEN ALGORITHMS (OPAL)

Przetwarzanie danych wrażliwych wiąże się z występowaniem wzmożonego ryzyka nie tylko w braku zapewnienia odpowiednich zabezpieczeń, ale też w przypadku wycieku danych i ich trafienia w niepowołane ręce. W przypadku udzielaniu dostępu do bazy danych, podmioty zewnętrzne mają możliwość kopiowania danych na swój serwer, w konsekwencji czego może dojść do ich niekontrolowanego wykorzystania w złej wierze lub w sprzeczności z ustalonymi zasadami określonego modelu współdzielenia danych.

Aby zapobiec tej sytuacji, proponujemy odejście od udostępniania danych rozumianego jako przekazywanie danych do bazy innego podmiotu, na rzecz koncepcji ‘move algorithm to data’ stworzonej przez Massachusetts Institute of Technology. Koncepcja ta polega na przenoszeniu algorytmu należącego do podmiotu zewnętrznego do urządzeń krańcowych modeli współdzielenia danych (np. wspólnicy danych zdrowotnych) – w taki sposób, aby dane wrażliwe nigdy nie wydostały się poza bezpieczne środowisko repozytorium organizacji. Możliwe jest przyjęcie dodatkowych zabezpieczeń zapobiegających wypłynięciu danych z innej strony, stosując szyfrowanie homomorficzne, polegające na operowaniu na zaszyfrowanych danych bez konieczności ich “odtajniania” (Hardjono i Pentland, 2019).

Jednym z założeń przyjętych przez tę koncepcję jest również kwestia weryfikacji przesyłanych algorytmów – po to, by upewnić się, że są one “wolne” od uprzedzeń, dyskryminacji czy naruszeń prywatności. Zwracane właścicielowi algorytmu wyniki w postaci zagregowanych odpowiedzi muszą być na tyle dokładne, aby nie pozwalały odbiorcy na przeprowadzanie ataków korelacyjnych, które doprowadzą do ponownej identyfikacji osób. Jeżeli algorytm zmierza natomiast do uzyskania odpowiedzi specyficznych względem osoby, której dane dotyczą, operacje przeprowadzane za pośrednictwem tego algorytmu mogą być wykonywane wyłącznie po uzyskaniu potwierdzonej i w pełni świadomej zgody tej osoby (Hardjono, Pentland, 2019).

TECHNOLOGIA BLOCKCHAIN

Chociaż technologia blockchain służy do zdecentralizowanego gromadzenia danych, ze względu na ogólną dostępność prowadzonego rejestru i jego zawartości, nie jest dobrym rozwiązaniem dla przechowywania danych wrażliwych. Z uwagi jednak na jego transparentność i brak możliwości ingerencji we wprowadzone dane, blockchain mógłby posłużyć jako **rejestr udzielanych zezwoleń** – zarówno podmiotom zewnętrznym, przekazującym swoje algorytmy do modeli współdzielenia danych, jak również osób, które

udzieliły zgody na przetwarzanie ich danych. Pozwoliłoby to na zapewnienie społecznej kontroli nad udzielanymi dostępami i faktycznej przejrzystości jakiejkolwiek operacji na danych.

INŻYNIERIA OCHRONY DANYCH

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, spełniając najważniejsze cele tworzenia infrastruktury, do których należą integralność, poufność, dostępność, możliwość interwencji oraz brak powiązań (ENISA, 2022).

Na szczególną uwagę zasługują technologie ochrony prywatności (*privacy enhancing technologies*), które jako główne zasady zastosowania poszczególnych rozwiązań technicznych wyróżniają:

- **Zachowanie prawdy**: Celem inżynierii prywatności jest zachowanie prawdziwości danych przy jednoczesnym zmniejszeniu ich zdolności do ich identyfikacji. Cel ten może być osiągnięty na przykład poprzez **zmniejszenie szczegółowości danych** (np. z daty urodzenia do wieku). W ten sposób dane są nadal dokładne, ale w „zminimalizowany sposób", odpowiedni dla danego celu.
    - Wskazywaną metodą w tym zakresie jest **szyfrowanie danych**, polegające na przekształceniu jawnych i otwartych informacji w kryptogram, czyli zaszyfrowany tekst, który może być "odtajniony" za pomocą odrębnie trzymanych kluczy (Bitdefender, 2022). Technikę tą można uznać za zachowującą prawdziwość zbioru danych, ponieważ szyfrowanie zastosowane w odwrotnym kierunku w pełni przywraca oryginalne dane, nie wprowadzając do procesu żadnej niepewności (ENISA, 2022).
- **Zachowanie czytelności:** Dane są przechowywane w formacie, który „ma sens" dla administratora, bez ujawniania rzeczywistych atrybutów osób, których dane dotyczą.
    - Jedną z metod służących zachowaniu czytelności może być tzw. **prywatność różnicowa**, która polega na dodawaniu szumu do rzeczywistych danych niemającego dużego wpływu na ich ogólną użyteczność (Kaczmarek, 2022). Nie zmienia to ogólnego wrażenia i charakterystyki zbioru danych, ale zapewnia ich poufność.

Wśród innych metod zapewniających odpowiednie bezpieczeństwo danych znajdują się:

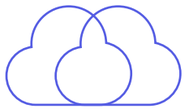
- **szyfrowanie homomorficzne** – umożliwia przeprowadzanie obliczeń na danych bez konieczności odszyfrowywania;

- **dane syntetyczne** – tworzenie danych w sposób mający przypominać dane rzeczywiste (np. rozkładem wartości zagregowanych), nie odnoszące się jednak do faktycznie istniejących osób fizycznych; ich opracowywanie ma na celu zmanipulowanie możliwości ponownej identyfikacji osób fizycznych.

OCHRONA PRZED WYCIEKIEM INFORMACJI (DATA-LOSS-PREVENTION)

Ze względu na przechowywanie w ramach jednej bazy różnych rodzajów danych ze względu na stopień ich wrażliwości, w praktyce niemożliwe jest samodzielne prześledzenie i wykrycie tych, które wymagają szczególnej ochrony. Rozwiązaniem przyspieszającym ten proces są systemy ochrony przed wyciekami informacji (*data loss prevention software*) wykorzystujące zestaw narzędzi i procesów stosowanych dla zapewnienia, że wrażliwe dane nie zostaną utracone, niewłaściwie wykorzystane lub udostępnione nieupoważnionym użytkownikom.

Oprogramowanie DLP klasyfikuje dane, wyróżniając między innymi poufne i krytyczne dane biznesowe, identyfikując naruszenia zasad zdefiniowanych przez daną organizację lub tych wynikających z ogólnie obowiązujących regulacji (np. HIPAA, PCI-DSS lub GDPR). Dla systemów zawierających dane wrażliwe system DLP może identyfikować, klasyfikować i oznaczać je w odpowiedni sposób, a także monitorować działania i zdarzenia związane z tymi danymi. Ponadto funkcje raportowania zapewniają szczegółowe informacje potrzebne do przeprowadzania audytów zgodności. Po zidentyfikowaniu naruszeń system wymusza środki zaradcze w postaci alertów, szyfrowania i innych działań ochronnych, aby uniemożliwić użytkownikom końcowym przypadkowe lub złośliwe udostępnianie danych, które mogą stanowić zagrożenie dla organizacji (De Groot, 2020).



COMPLIANCE DOSTAWCÓW USŁUG CHMUROWYCH

Prawdziwym wyzwaniem jest również 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 dotychczas 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 z wymogiem spełnienia takich samych warunków ochrony danych osobowych. Z powodu niechęci do zmiany przepisów po stronie organów amerykańskich, konieczne jest rozważenie wyboru 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. Inna opcją może być wyznaczenie standardów niezbędnych do spełnienia przez dostawcę chmurowego do realizacji usługi chmurowej, zawierającej między innymi wymogi dotyczące posiadania odpowiednich certyfikatów czy oddania przywileju utrzymywania kluczy do zaszyfrowanych danych instytucji korzystającej z tworzonej przez dostawców chmury (Datailsynet, 2022).

6. Wnioski i rekomendacje pilotażowe

Na podstawie przeprowadzonych warsztatów i badań eksperckich proponujemy przeprowadzić pilotaże dwóch modeli współdzielenia danych: wirtualnej wspólnicy danych przemysłowo-rolnych oraz publicznej wspólnicy danych zdrowotnych. Rola państwa w tym zakresie powinna polegać na całkowitym sfinansowaniu takich innowacji oraz koordynacji działań m.in. poprzez ustalanie ram prawnych oraz zasad ich funkcjonowania wykraczających poza same pilotaże, uwzględniających warstwę prawno-kontraktową, standardów technicznych i procesów organizacyjnych.

Źródło: Wykres opracowany na podstawie warsztatów badawczych „Uwolnić potencjał danych”.

6.1. Pilotaż wirtualnej wspólnicy danych przemysłowo-rolnych

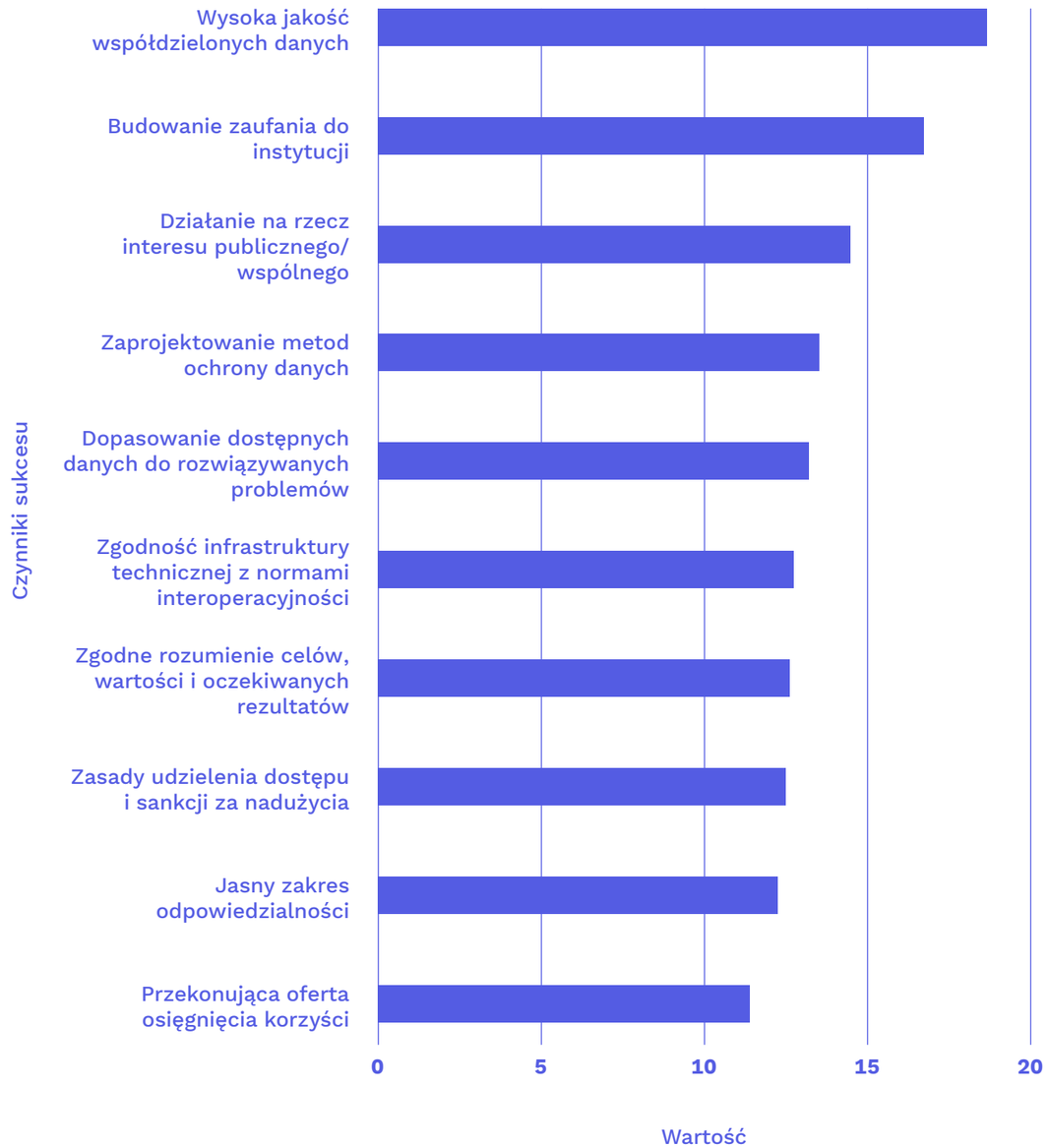
Nowe technologie odgrywają coraz ważniejszą rolę w światowym rolnictwie i przemyśle. Na efektywność produkcji i wydajność gospodarki wpływają bowiem zarówno standardowe czynniki (np. kwalifikacje pracowników), jak również zaawansowanie wykorzystywanej technologii oraz intensywność wykorzystania danych (Koloch G., Grobelna K., Zakrzewska-Szlichtyng K., Kamiński B., Kaszyński D. (2017). Kraje, którym zależy na konkurencyjności ich sektora rolno-spożywczego, logistycznego czy przemysłowego na rynku globalnym dokonują znacznych inwestycji w rozwiązania tzw. *smart farming* czy IoT. Aby nie pozostawać w tyle oraz nie narażać polskich gospodarstw na popadnięcie dług technologiczny, konieczny jest rozwój nowych instytucji w tym obszarze. Jest to szczególnie ważne w kontekście wyrównywania szans polskich rolników na arenie unijnej. Mimo wzrostu produktywności w ostatnich latach, w dalszym ciągu jest ona niższa niż w przodujących krajach Europy Zachodniej (i zbliżona jest do poziomu, który Francja i Niemcy osiągały w latach 70. ubiegłego wieku) (Miniszewski, M., 2021). Równocześnie wskazuje się, że możliwe jest podnoszenie produktywności poprzez wdrażanie innowacji w postaci nowych technologii w tym m.in. big data, rolnictwa precyzyjnego, technologii hodowli roślin.

Patrz s.62

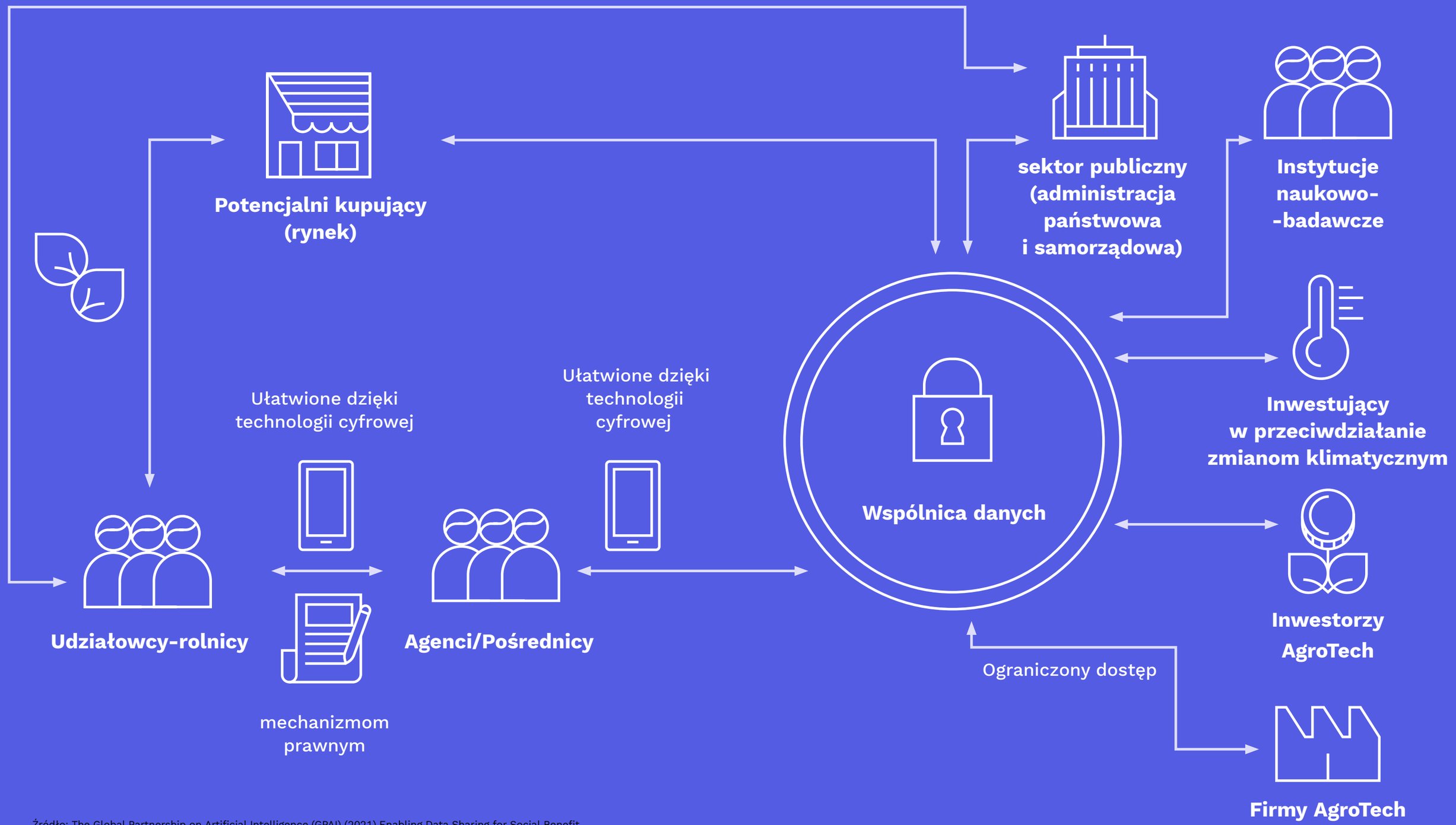
Warto zauważyć, że choć polski przemysł charakteryzuje się relatywnie niską intensywnością wykorzystania danych na tle innych państw europejskich, dane mają wysoki udział w produktywności gospodarki (nawet w porównaniu do krajów o wiodącej w UE intensywności wykorzystania danych) (Ministerstwo Rozwoju, 2020). Tym samym, „opóźnianie lub wręcz zaniechanie działań mających na celu rozwój warunków społeczno-gospodarczych sprzyjających rozwojowi gospodarki w wysokim stopniu opartej na danych powodować będzie, już w średnim okresie, istotne obniżanie możliwych do osiągnięcia korzyści ekonomicznych” (Koloch, G., Grobelna, K., Zakrzewska-Szlichtyng, K., Kamiński, B., Kaszyński, D., 2017).

W trakcie warsztatów zidentyfikowaliśmy 10 kluczowych czynników sukcesu, które wyznaczają główne wyzwania do rozwiązania w toku pilotażu. Uwagę zwracają też czynniki, które w toku warsztatów znacząco zyskały/straciły na znaczeniu w oczach interesariuszy.

WYKRES 6. 10 Kluczowych czynników sukcesu dla wirtualnej wspólnicy danych przemysłowo-rolnych

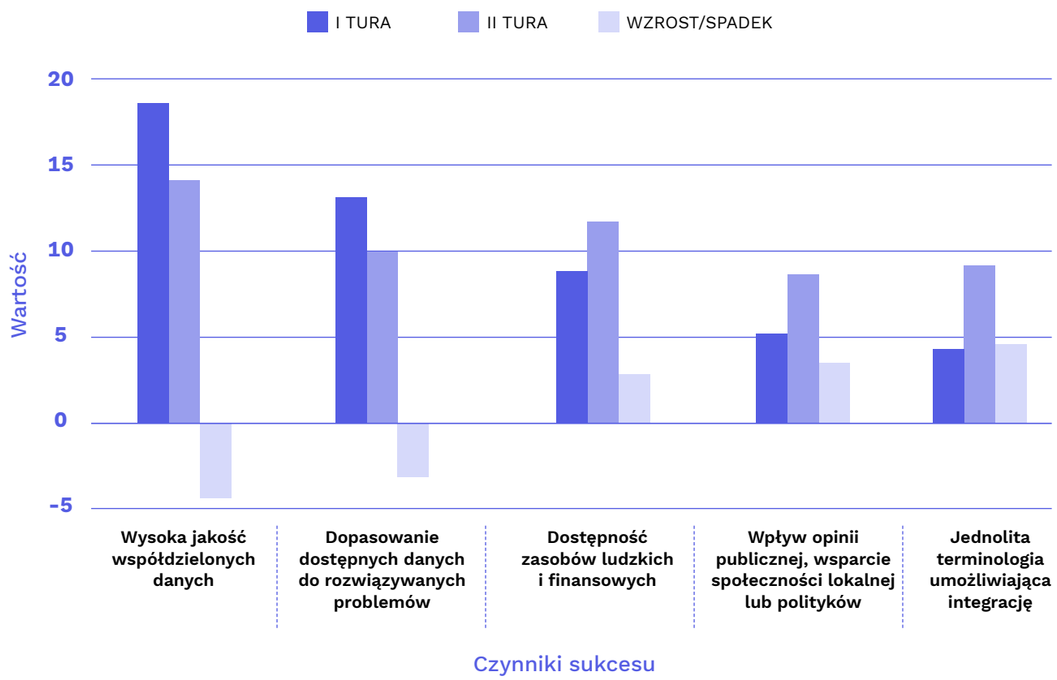


WYKRES 7. Przykładowy model wirtualnej składnicy danych



Źródło: The Global Partnership on Artificial Intelligence (GPAI) (2021) Enabling Data Sharing for Social Benefit Through Data Trusts: Data Trusts in Climate

WYKRES 8. 5 czynników, które uległy największej zmianie w toku warsztatów



PILOTAŻ PROPONUJEMY OPRZEĆ O KOLEJNE KROKI:

1

WYZNACZENIE PILOTAŻOWYCH ZBIORÓW DANYCH I STANDARDU WSPÓŁDZIELENIA

Ponieważ jakość danych i ich wolumen ma fundamentalne znaczenie dla powodzenia projektu, tworzący wspólnicę muszą zidentyfikować konkretne zbiory danych, które są dostępne, przetwarzane, oraz mają cechy czyniące je relatywnie prostszymi do współdzielenia niż zbiory wymagające intensywnego oczyszczania lub uzupełniania. Takie zbiory mogą mieć charakter bardziej przemysłowy lub rolny, mogą funkcjonować na pograniczu i pochodzić od podmiotów różnego rodzaju. Klucz funkcjonalny jest istotniejszy niż ścisłe wyznaczenie granicy sektorowej. Należy przy tym zdefiniować standard współdzielenia tych danych, tak, aby umożliwić interoperacyjność systemów.

2

WYBÓR FORMUŁY PRAWNEJ, CELÓW I STRATEGII WSPÓLNICY W TRANSPARENTNYM DIALOGU BRANŻOWYM

Rekomendowane formuły prawne to przede wszystkim spółdzielnia danych (jako zaufany podmiot), wskazanie agencji publicznej jako **zaufanego podmiotu lub współpraca kontraktowa**. Jednak budowa zaufania do współdzielenia danych powiedzie się, jeżeli wybór formy wirtualnej wspólnicy, ustalenie celów, wartości, metod ich osiągania, oraz oczekiwanych rezultatów zostanie wypracowane w dialogu o maksymalnie otwartym charakterze. Niezależnie od tego, czy wspólnica rozpocznie od zestawów danych przemysłowych czy rolnych, konsultacje powinny uwzględniać także przyszłych uczestników współdzielenia, oraz potencjalnych użytkowników danych (sektor nauki, analityki biznesowej, innowacji). Zaangażowanie stowarzyszeń branżowych umożliwiłoby wypracowanie warunków, na jakich powinna odbywać się współpraca w zakresie wymiany danych pomiędzy firmami; w jaki sposób powinny być przestrzegane reguły konkurencji (w szczególności w zakresie wymogów członkostwa, rodzaju udostępnianych danych, dostępu osób trzecich); w jaki sposób ramy własności intelektualnej wpływają na alokację praw do danych w zaufanych przestrzeniach danych. 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 obszarach wiejskich – nie tylko potencjalnych beneficjentów wspólnicy (rolników), ale także społeczno-zawodowe organizacje (np. kółka rolnicze; koła gospodyń wiejskich; rolnicze zrzeszenia branżowe) oraz przedstawicieli lokalnej wiary.

3

OPRACOWANIE ZASAD DOSTĘPU I METOD OCHRONY DANYCH

Tworzący wspólnicę muszą zdecydować jakie chcą stosować metody ochrony danych i przygotować zasady dostępu (np. w formie regulaminu i licencji użytkowania). Ponieważ dane przemysłowe i rolne nie są danymi wrażliwymi w rozumieniu RODO, a ich ochrona opiera się najczęściej na tajemnicy przedsiębiorstwa, rekomendujemy wykorzystanie federacyjnej (rozproszonej) architektury danych, potencjalnie z wykorzystaniem federacyjnego uczenia maszynowego w przetwarzaniu brzegowym, oraz zastosowanie technologii blockchain do rejestracji użycia. Podstawowym pytaniem jest jednak kto może udzielać zgody i czy wymagana jest ludzka weryfikacja użytkownika. Jeśli nie, zgoda może mieć miejsce automatycznie na podstawie ID bądź akceptacji licencji (np. z użyciem bramki w interfejsie API). Jeśli niezbędna jest każdorazowo zgoda, odpowiedni proces musi zostać uwzględniony w planach pilotażu.

4

OPRACOWANIE MODELU BIZNESOWEGO ZE WSPÓLNYMI KORZYŚCIAMI

Ponieważ uczestnicy współdzielenia danych wirtualnej wspólnicy są podmiotami prowadzącymi działalność gospodarczą, model biznesowy musi przewidywać jasno zdefiniowane wspólne korzyści. Z tego względu trafionym pomysłem wydaje się być tworzenie pozytywnych zachęt dla współdzielenia danych poprzez oferowanie obietnicy w zyskach pochodzących z wytworzonych produktów lub usług (np. aplikacji powstałych na podstawie danych udostępnionych przez grupę zainteresowanych rolników). Jako przykład zachęt uczestnicy warsztatu podali także dostępność próbek produktów rolniczych stworzonych na podstawie danych przekazywanych za pośrednictwem wspólnicy czy możliwość korzystania z wytworzonej, innowacyjnej infrastruktury technicznej. Możliwość pokrycia kosztów własnych oraz przekonująca oferta zwiększają prawdopodobieństwo na dołączanie nowych członków i rozwój projektu poza pilotaż.

5

PRZYGOTOWANIE INFRASTRUKTURY I PRZYPISANIE ODPOWIEDZIALNOŚCI ZA OBSŁUGĘ

Ostatnim etapem przed uruchomieniem pilotażu jest przygotowanie infrastruktury technicznej (zarówno chmury na *back-endzie*, jak i interfejsów dostępowych na *front-endzie*) i wskazanie które podmioty są odpowiedzialne za obsługę – administrację, pomoc techniczną, partycypowanie w kosztach na pierwszym etapie. Nie wszystkie podmioty będą równie w stanie tworzyć wspólnicę na początku, dlatego korzystne będzie wskazanie liderów technicznych na tym etapie.

6.2. Pilotaż wspólnicy danych zdrowotnych

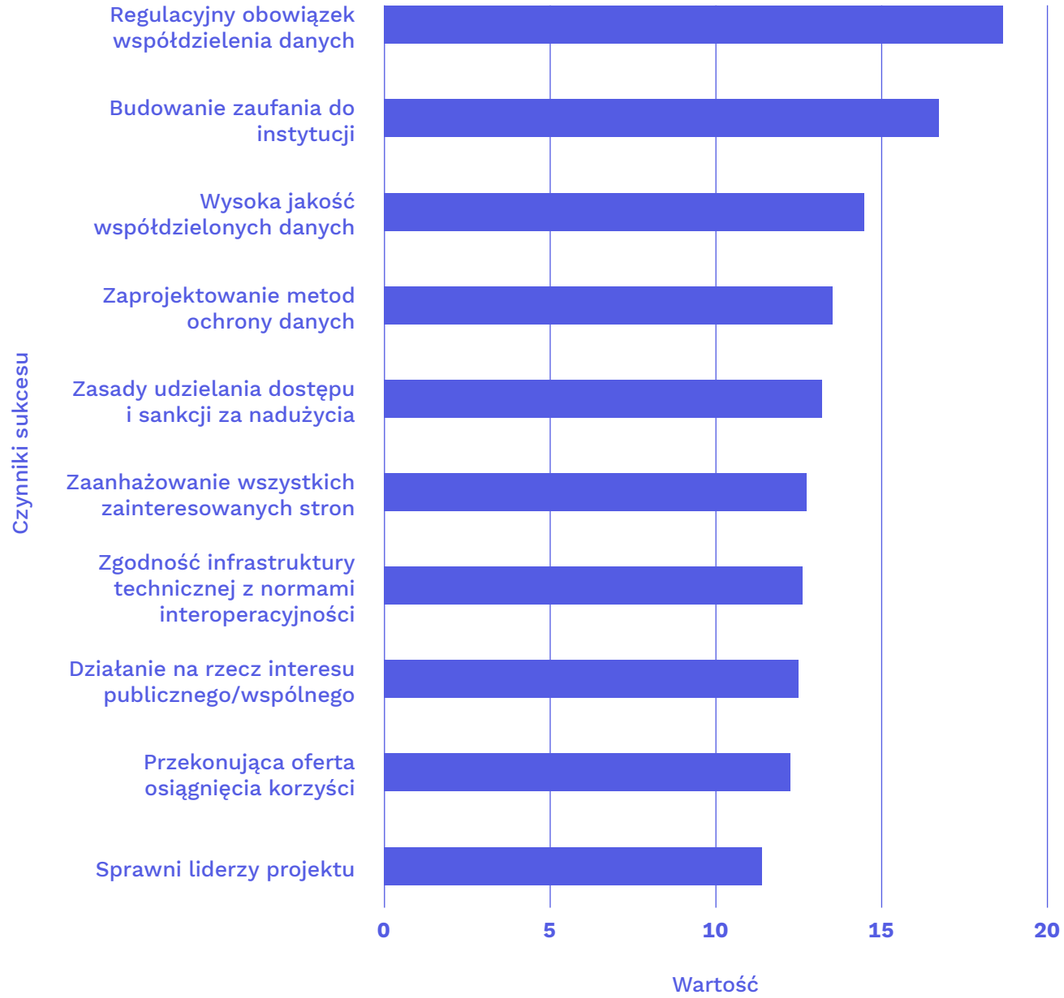
Przyspieszony w ostatnich latach proces cyfryzacji ochrony zdrowia przyczynił się do masowego powstawania zbiorów danych zdrowotnych – zgodnie ze statystykami prowadzonymi przez przez Centrum e-Zdrowia, od 2019 roku liczba zarejestrowanych kont w systemie e-zdrowie wzrosła z ok. 600 tysięcy do ponad 10 milionów (Torchała, 2021). O istotnej roli digitalizacji publicznego sektora medycznego i wprowadzenia odpowiednich zmian świadczy również “Krajowy plan transformacji na lata 2022-2026” opracowany przez Ministerstwo Zdrowia, którego autorzy wskazują konieczność wdrożenia odpowiednich narzędzi wspomagających analizę stanu zdrowia pacjenta, rozwoju algorytmów sztucznej inteligencji oraz budowę centralnego repozytorium danych medycznych, a także dalszej cyfryzacji dokumentacji medycznej i budowanie ekosystemu jej wymiany (Kościelniak, 2021). Jednocześnie, ponad połowa Polaków korzysta z placówek prywatnych, w których również podkreślana jest rola inwestowania w bezpieczeństwo infrastruktury informatycznej oraz gromadzenia danych pacjentów w taki sposób, by były one uporządkowane i możliwe do dla dalszego wykorzystania (Pawlak, 2021). Widoczny jest również wzrost zainteresowania cyfrowymi narzędziami przeznaczonymi m.in. do samodzielnego monitorowania stanu własnego zdrowia, kontrolowania przebiegu leczenia czy przyspieszenia procesu udzielania porady lekarskiej. Zgodnie z raportem „162 mobilne aplikacje zdrowotne” opracowanym przez Ogólnopolski System Ochrony Zdrowia, już kilka lat temu na rynku funkcjonowało ponad 260 tysięcy mobilnych aplikacji zdrowotnych, a liczba ich pobrań sięgała nawet 3 mld (Głos Pacjenta, 2017).

Pomimo coraz powszechniejszego wykorzystywania elektronicznych baz danych dla przechowywania informacji o stanie zdrowia, systemy te (publiczny, prywatny i oraz związany z funkcjonowaniem aplikacji zdrowotnych) nie są przygotowane do korespondowania ze sobą oraz wzajemnego wykorzystywania gromadzonych zasobów. Konieczne jest zatem zapewnienie interoperacyjności silnej między systemami poprzez ustalenie jednolitych standardów gromadzenia danych i wyboru kompatybilnych systemów

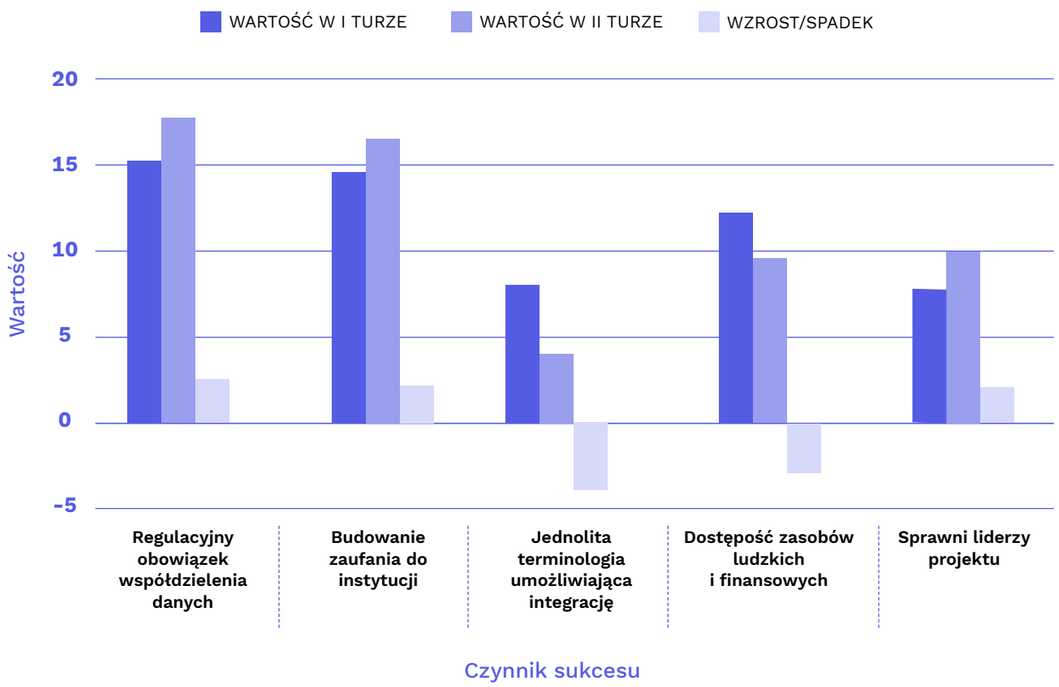
ich przechowywania. Równocześnie, w związku z szczególnie sensytywnym charakterem danych dotyczących zdrowia, nieodzownym elementem tworzenia mechanizmów współdzielenia powinno być stopniowe budowanie zaufania społecznego poprzez zapewnienie współzarządzania zasobami i możliwości decydowania o ich wykorzystaniu w interesie publicznym. Jedynie w ten sposób, a więc poprzez symultaniczne prowadzenie inicjatyw regulacyjnych i społecznych, możliwe będzie efektywne wykorzystanie potencjału danych zdrowotnych dla wspólnego dobra.

W trakcie warsztatów zidentyfikowaliśmy 10 kluczowych czynników sukcesu, które wyznaczają główne wyzwania do rozwiązania w toku pilotażu. Uwagę zwracają też czynniki, które w toku warsztatów znacząco zyskały/straciły na znaczeniu w oczach interesariuszy.

WYKRES 9. 10 kluczowych czynników sukcesu dla wspólnicy danych zdrowotnych



WYKRES 10. 5 czynników, które uległy największej zmianie w toku warsztatów



PILOTAŻ PROPONUJEMY OPRZEĆ O KOLEJNE KROKI:

1

WYZNACZENIE PILOTAŻOWYCH ZBIORÓW DANYCH I STANDARDU WSPÓŁDZIELENIA

Ponieważ jakość danych i ich wolumen ma fundamentalne znaczenie dla powodzenia projektu, tworzący wspólnicę muszą zidentyfikować konkretne zbiory danych, które są dostępne, przetwarzane, oraz mają cechy czyniące je relatywnie prostszymi do współdzielenia niż zbiory wymagające intensywnego oczyszczania lub uzupełniania. Takimi zbiorami w szczególności są: obrazy medyczne, EDM, wyniki morfologii krwi, dane z urządzeń dla cukrzyków, dane z urządzeń mierzących puls. Warto pamiętać, że przyjęcie odpowiedniej ustawy przyjmującej system opt-out i/lub nakazującej współdzielenie B2G może zwiększyć zakres dostępnych zbiorów. Należy przy tym zdefiniować standard współdzielenia tych danych, tak, aby umożliwić interoperacyjność systemów.

2

UTWORZENIE MECHANIZMÓW DEMOKRATYCZNEGO WSPÓŁZARZĄDZANIA DANYMI

Publiczny charakter wspólnic danych wymaga maksymalnej inkluzywności w procesie ich tworzenia. Nieprzychylna reakcja na pomysł udostępniania informacji wynika z braku wcześniejszej praktyki współdzielenia danych i silnego przekonania o możliwej utracie prywatności w sieci. 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. Z tego względu, rola tworzących powinna skupiać się na koordynowaniu procesu zakładania społecznych rad nadzorczych lub innych form demokratycznej kontroli (np. panele obywatelskie wyposażone w narzędzia decyzyjności). To one powinny podejmować decyzje w zakresie funkcjonowania wspólnic – w imię zasady “from decision-maker to decision-taker“ od najniższych szczebli – co pozwoli zbudować zaufanie społeczne, zwiększy transparentność procesu i od początku zaangażuje interesariuzy w dalszy rozwój projektu. Reprezentatywność może zapewnić połączenie kluczowych interesariuszy (Rzecznik Praw Pacjenta, samorządy pracowników ochrony zdrowia, organizacje pacjenckie, organizacje prawnoczłowiecze i gospodarki cyfrowej) z elementem losowej reprezentacji. Rekomendujemy, aby utworzenie takiego demokratycznego nadzoru było krokiem poprzedzającym dalsze wiążące decyzje w pilotażu. Z uwagi również na potencjał płynący z międzynarodowych praktyk współdzielenia danych i uczestnictwa Polski w partnerstwach wielostronnych i organizacjach o zasięgu globalnym, rolą państwa powinno być wspieranie interesariuszy w tworzeniu nowych, transgranicznych powiązań i budowaniu współpracy międzynarodowej w zakresie bezpiecznego i innowacyjnego dzielenia się danymi.

3

WYBÓR FORMUŁY PRAWNEJ, CELÓW I STRATEGII WSPÓLNICY W KONSULTACJACH SPOŁECZNYCH

Rekomendowane formuły prawne to przede wszystkim współzarządzana instytucja publiczna (najprawdopodobniej Centrum eZdrowia) lub współzarządzany ośrodek naukowy (państwowy instytut badawczy bądź uczelnia o wysokim poziomie kompetencji technologicznych). Jednak budowa zaufania do współdzielenia danych powiedzie się, jeżeli wybór formy wspólnicy danych, ustalenie celów, wartości, metod ich osiągania, oraz oczekiwanych rezultatów zostanie wypracowane w dialogu o maksymalnie otwartym charakterze. Choć ten argument może brzmieć powtarzalnie, w przypadku danych wrażliwych nie da się go przecenić. Konsultacje powinny wychodzić od demokratycznego nadzoru wspólnicy, ale uwzględniać także przyszłych uczestników współdzielenia (poza pilotażowymi zbiorami danych), potencjalnych użytkowników danych (sektor nauki, ochrony zdrowia, innowacji) oraz pacjentów. Autorzy rekomendują w tym zakresie nawet przeprowadzenie badań, jak np. replikacja badania DCE dla współdzielenia danych zdrowotnych (Johansson et al., 2021), a następnie kampanię informacyjną aby mieć pewność, że wdrażanie pilotażu zostanie odebrane przez społeczeństwo jako element pozytywnego postępu społecznego. Komunikacja społeczna powinna zawierać jak najwięcej przykładów technicznych poświadczających bezpieczeństwo i odporność technologiczną wspólnic przed możliwymi wyciekami danych. Ponadto, dzięki organizacji warszta-

tów wyjaśniających w jasny i przystępny sposób funkcjonowanie wspólnic i ich model biznesowy, 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

OPRACOWANIE METOD OCHRONY DANYCH, ZASAD DOSTĘPU I MODELU BIZNESOWEGO SŁUŻĄCEGO INTERESOWI PUBLICZNEMU

Istotnym zadaniem w procesie tworzenia wspólnicy będzie jasne określenie możliwości wykorzystania danych polskich użytkowników i użytkowniczek/ pacjentów i pacjentek zgromadzonych w publicznych wspólnicach danych zdrowotnych. Ustalenia wymaga alokacja wartości w tym modelu oraz kwestia komercjalizacji i publicznego udostępniania wyników badań. Kluczowym będzie również określenie, czy z zasobów wspólnicy będą mogły korzystać podmioty spoza Polski, czy Unii Europejskiej, a jeśli tak, to na jakich zasadach. Jednak najważniejszym elementem wydaje się być opracowanie odpowiednich metod ochrony danych. W tym zakresie autorzy rekomendują rozwiązania takie, jak chociażby szyfrowanie homomorficzne, przechowywanie danych syntetycznych, czy wykorzystywanie specjalistycznych systemów ochrony przed wyciekiem danych (*data-loss-prevention*)

(patrz: **5.4.2.**).

5

PRZYGOTOWANIE INFRASTRUKTURY I WYBÓR ORAZ SZKOLENIE LIDERÓW PROJEKTU

Aby zagwarantować pełną operacyjność wspólnicy koniecznym jest zaprojektowanie oraz wdrożenie godnej zaufania infrastruktury technicznej. W tym zakresie rekomendowane jest tworzenie systemów informatycznych wspólnic we współpracy z europejskimi oraz krajowymi dostawcami rozwiązań IT. Zadania z zakresu budowy infrastruktury mogłyby być zlecane w drodze postępowań o zamówienia publiczne – w formie konkursu bądź przetargu ograniczonego. Tryb ten pozwoliłby wyłonić wykonawcę na sprawiedliwych, konkurencyjnych zasadach, dzięki czemu możliwe byłoby zapewnienie technologii najwyższej jakości. Ograniczenie przetargu jedynie do ofert składanych przez podmioty z UE pozwoliłoby natomiast zagwarantować bezpieczeństwo danych oraz (dzięki niezależności tych podmiotów od korporacji z państw trzecich) suwerenność systemów.

Dodatkowo procesowi tworzenia wspólnic powinien towarzyszyć nabór na odpowiednich liderów projektu. Wskazanym jest, aby były to osoby posiadające nie tylko wiedzę techniczną w zakresie infrastruktury, lecz także dysponujące wysoko wykształconymi kompetencjami miękkimi takimi, jak kreatywność, umiejętność zarządzania zespołem, wielozadaniowość. Koordynatorzy Ci mogliby być wyłaniani w drodze otwartego naboru Liderów Innowacji (GovTech) sprofilowanego na poszukiwanie specjalistów z adekwatnym doświadczeniem.

6.3 Wybrane rekomendacje dla państwa

INWESTYCJE W DŁUGOTERMINOWY PROGRAM PODNOSZENIA KWALIFIKACJI I UMIEJĘTNOŚCI CYFROWYCH

Aby w pełni wykorzystywać potencjał wspólnicy konieczne jest wyposażenie jej przyszłych użytkowników w odpowiednie kompetencje cyfrowe. Państwo powinno zainwestować więc w długofalowe działania ukierunkowane na podnoszenie jakości nauczania informatycznego na każdym etapie edukacji. Instytucje szkolnictwa podstawowego powinny rozwijać programy nauczania odpowiadające potrzebom zdigitalizowanego świata. Ponadto dobrym pomysłem byłoby włączenie cyfryzacji w kształcenie zawodowe oraz zapewnienie specjalnych programów szkoleniowych dedykowanych poszczególnym sektorom (np. rolno-spożywczemu; energetycznemu; medycznemu). Programy te mogłyby koncentrować się nie tylko na podstawowej wiedzy i praktycznym rozumieniu narzędzi cyfrowych, ale także na promowaniu innowacyjnych rozwiązań w instytucjach czy prywatnych przedsiębiorstwach.

REGULACJE DLA STANDARYZACJI DANYCH ORAZ ZAPEWNIENIA INTEROPERACYJNOŚCI SILNEJ

Ponieważ wspólnica danych jest całkowicie nowym projektem, należy zbudować świadomość i nowe nawyki w sektorach właściwych dla pilotaży, rozpoczynając od sankcjonowanego obowiązku stosowania ustalonych formatów. Tym samym, zaleca się wprowadzenie obowiązku regulacyjnego umieszczania danych we wspólnicy. Równocześnie wskazanym byłoby promowanie narzędzi RegTech służących do skutecznego egzekwowania obowiązujących przepisów poprzez stosowanie nowych technologii weryfikujących spełniania wymogów regulacyjnych i compliance w poszczególnych branżach. Ponadto konieczne jest ustanowienie odpowiednich standardów. Odpowiedzialny za wyznaczanie formatów właściwych dla danego sektora; certyfikowanie obsługujących lub przystępujących do wspólnicy; egzekwowanie zasad wspólnicy mógłby być podmiot wyznaczony przez rząd występujący jako odrębna instytucja (np. państwowa agencja wykonawcza). Należy rozważyć także wyodrębnienie i zaklasyfikowanie reprezentacji danych charakterystycznych dla danej dziedziny oraz standardów specyficznych dla poszczególnych sektorów (np. zdrowia; przemysłu wytwórczego) (Komisja Europejska, 2019).

WSPÓLNE ZAKUPY CERTYFIKOWANYCH OPROGRAMOWAŃ

Aby wzmocnić bezpieczeństwo danych trafiających do wspólnicy, warto uodpornić systemy na możliwe ataki cybernetyczne, wybierając we wspólnym przetargu najlepsze możliwe oprogramowanie dla poszczególnych instytucji publicznych. Dzięki jednolitej wycenie pożądanych świadczeń,

podmioty korzystające z infrastruktury krytycznej (tj. szpitale, sektor energetyczny, przemysł rolno-spożywczy) będą w stanie przyjąć jedną, spójną politykę zakupową.

USTAWA O PONOWNYM WYKORZYSTANIU DANYCH ZDROWOTNYCH W INTERESIE PUBLICZNYM

Wzorem Finlandii, rekomendujemy opracowanie ustawy umożliwiającej ponowne wykorzystanie danych zdrowotnych do celów rozwoju ochrony zdrowia na zasadach opt-out (domyślne ponowne wykorzystanie). 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[1]. Jednak warto rozważyć poszerzenie katalogu podmiotów, którym można udostępniać dokumentację na podstawie przepisów zawartych w ustawie oraz ustanowić zasady, na podstawie których taki dostęp byłby udzielany z korzyścią dla interesu publicznego. Co ważne, z uwagi na niejasny charakter "celów naukowych", rekomendowanym jest zdefiniowanie tego pojęcia i poszerzenie katalogu możliwości wykorzystywania danych dotyczących zdrowia również na inne uzasadnione cele w interesie publicznym, także dla instytutów prywatnych.

USTANOWIENIE JEDNOLITEJ DEFINICJI DANYCH DOTYCZĄCYCH ZDROWIA

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**
  bierze się pod uwagę kontekst, w którym dane funkcjonują, a więc interesy, warunki i konsekwencje 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, należy wziąć jeszcze pod uwagę kontekst.

W kontekście gromadzenia i przetwarzania danych we wspólnicy rekomendowane jest przyjęcie tzw. podejścia mieszanego, które zapewnia najszerszą ochronę wykorzystywanych danych.

ODPOWIEDNIO DOBRANE ŹRÓDŁA FINANSOWANIA

Niewątpliwie do budowy infrastruktury technicznej oraz utrzymania wyspecjalizowanej kadry zdolnej do obsługi innowacyjnych rozwiązań konieczne są znaczne nakłady finansowe. Proponuje się, aby w ramach funduszy publicznych przyznanych poszczególnym resortom wypracowany został odpowiedni mechanizm umożliwiający pozyskanie finansowania (np. w drodze konkursu, zamówień publicznych, kredytu technologicznego) przeznaczonego na budowę przestrzeni wymiany danych dedykowanych poszczególnym obszarom. Dodatkowo dobrym źródłem pozyskiwania środków na tworzenie wspólnic (np. biznesowych) mógłby być popularny zagranicą crowdfunding polegający na angażowaniu prywatnych inwestorów w dane przedsięwzięcie. Wymagałoby to jednak zbudowania odpowiedniego otoczenia regulacyjnego sprzyjającego crowdfundingowi inwestycyjnemu (Rada Ministrów, 2019).

1 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.

7. 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 i publicznej wartości (Creating Shared/Public 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 wypaczonym modelem kapitalizmu kognitywnego – czyli występującym w dzisiejszej gospodarce nowym reżimem akumulacji, w którym alokacja wartości nie opiera się już na pracy fizycznej i systemie maszynowym, lecz na wyzysku wiedzy oraz kreatywności ludzi pełniących funkcję darmowych pracowników korporacji technologicznych (Zygmuntowski, 2020b). Podobnie należy konsekwentnie stawiać instytucje, pozwalające na porzucenie paradymatu społeczeństwa informacyjnego, jako stanowiącego źródło informacji, na rzecz budowania społeczeństwa wiedzy. W otaczającym nas świecie zarówno działania w sferze prywatnej, jak i profesjonalnej oparte są na współpracy z inteligentnymi, autonomicznymi maszynami. Tym samym, bez zbudowania społeczeństwa uczącego się, zdolnego adaptować się do nowych warunków, niemożliwym będzie wdrażanie innowacyjnych rozwiązań czy przebudowywanie organizacji (zarówno prywatnych, jak i państwowych) (Ministerstwo Cyfryzacji, 2019).

Aby więc 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/C739E1DE-223088FD3D761466DCDA2EFE
- 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, [Dostęp: 08.04.2022], Dostępny w: https://ssrn.com/abstract=3364722
- Bartol, A., Herbst, J., Pierścińska, A., (2021), Wykluczenie społeczno-cyfrowe w Polsce 2021.
- 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/
- Bechtel, M., Buchholz, S., Deloitte (2022), Tech Trends 2022 [Dostęp: 27.05.2022], Dostępny w:https://www2.deloitte.com/content/dam/Deloitte/pt/Documents/tech-trends/tech-trends-2022/DI_Tech-trends-2022.pdf
- Big Data Value Association (2019) Towards a European Data Sharing Space. Enabling data exchange and unlocking AI potential, BDVA Position Paper, Kwiecień 2019
- Bitdefender (2022) Co to jest szyfrowanie danych i kiedy warto je stosować? [online], Bitdefender, [Dostęp: 27.05.2022], Dostępny w: https://bitdefender.pl/co-to-jest-szyfrowanie-danych-i-kiedy-warto-je-stosowac/
- Borowik, M., Maśniak, L., Kroplewski, R., Romaniec, H. (2017) 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...
- Bożykowski, M., Chłoń-Domińczak A., Jasiński, M., Zając, T. (2019) Dane publiczne – nowy impuls do rozwoju Polski, Polski Instytut Ekonomiczny, Policy Paper 8/2019
- Data Collaboratives (2021) Data Collaboratives [online], GovLab, [Dostęp: 08.04.2022], Dostępny w: https://datacollaboratives.org/
- Datatilsynet (2022) Guidance on the use of the cloud, Datatilsynet, Marzec 2022
- Data Trust Initiative (2021) Data trusts: international perspectives on the development of data institutions, DTA, Working Paper 2
- De Groot, J. (2020) What is Data Loss Prevention (DLP)? A Definition of Data Loss Prevention [online], Datainsider: Digital Guardian’s Blog, [Dostęp: 30.05.2022], Dostępny w: https://digitalguardian.com/blog/what-data-loss-prevention-dlp-definition-data-loss-prevention
- 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
- Dymek, D., Komnata, W., Kotulski, L., Federacyjna hurtownia danych w dostępie do informacji poufnej, Akademia Górniczo-Hutnicza w Krakowie, dostęp: http://rocznikikae.sgh.waw.pl/p/roczniki_kae_z33_08.pdf
- 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

- ENISA (2022) Data Protection Engineering [online], ENISA, [Dostęp: 27.05.2022], Dostępny w: https://www.enisa.europa.eu/publications/data-protection-engineering
- 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
- The Global Partnership on Artificial Intelligence (GPAI) (2021) Enabling Data Sharing for Social Benefit Through Data Trusts: Data Trusts in Climate, [Dostęp 27.05.2022], Dostępny w: https://gpai.ai/projects/data-governance/data-trusts-in-climate-interim-report.pdf
- Głos Pacjenta, (2017) Aplikacje mobilne szturmują rynek zdrowia [online], Głos Pacjenta, [Dostęp: 31.05.2022], Dostęp w: https://glospacjenta.pl/aktywnosci/przydatne/305,aplikacje-mobilne-szturmuja-rynek-zdrowia
- 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
- B. Kiełbasa, J. Puchała, Innowacyjność młodych rolników i ich postawy wobec zmian na przykładzie gospodarstw rolnych położonych w regionie rozdrobnionego rolnictwa, „Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu" 2015, t. 17, z. 1, s. 107-111.
- Koloch G., Grobelna K., Zakrzewska-Szlichtyng K., Kamiński B., Kaszyński D. (2017). Intensywność wykorzystania danych w gospodarce a jej rozwój. Analiza diagnostyczna. [Dostęp: 07.06.2022], Dostępny w: https://mc.bip.gov.pl/rok-2017/analiza-diagnostyczna-intesywnosc-wykorzystania-danych-w-gospodarce-a-jejrozwoj.html
- 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 (2021) 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
- Komisja Europejska (2019) High-level Expert Group on AI; Policy and Investment Recommendations for Trustworthy AI, [online], Komisja Europejska, [Dostęp: 27.05.2022], Dostępny w: https://www.europarl.europa.eu/italy/resource/static/files/import/intelligenza_artificiale_30_aprile/ai-hleg_policy-and-investment-recommendations.pdf
- Kościelniak P., (2021) E-zdrowie w planie transformacji na lata 2022-2026, [online], Info. eZdrowie, [dostęp: 31.05.2022], Dostępny w: http://forumezdrowia.pl/info/aktualnosci/e--zdrowie-w-planie-transformacji-na-lata-2022-2026/
- Kroplewski, R. (2020), Protokół z XXXVII posiedzenia Rady do Spraw Cyfryzacji, KPRM, [dostęp: 06.06.2022],Dostępny w: https://www.gov.pl/attachment/af46cdbe-5ead-44cf-a8ca-a-91e9ac052a1
- Kroplewski, R., (2021), W stronę społeczeństwa wiedzy, Przegląd Techniczny Gazeta Inżynierska [Dostęp: 05.05.2022], Dostępny w: http://przeglad-techniczny.pl/artykuly?id=2874
- 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/
- Ministerstwo Cyfryzacji (2016) Program Otwierania Danych Publicznych, Załącznik do uchwały nr 107/2016 Rady Ministrów z dnia 20 września 2016 r.
- Ministerstwo Cyfryzacji, Polityka Rozwoju Sztucznej Inteligencji w Polsce na lata 2019 – 2027; Godna Zaufania Sztuczna Inteligencja autonomia i konkurencja +PL, Ciesielski, M., Flakiewicz,P., Jarzewski, A., Kroszczyńska, E., Lubos, B., Podgórska, A., Pukaluk, M., Pytko, T., Romaniec, H., Wancio, A., Stefaniak, S., Zaczek, A. (2019), [Dostęp: 06.06.2022), Dostępny w: https://www.gov.pl/attachment/0aa51cd5-b934-4bcb-8660-bfecb20ea2a9.
- Ministerstwo Cyfryzacji (2019) Polityka Rozwoju Sztucznej Inteligencji w Polsce na lata 2019 – 2027, Godna Zaufania Sztuczna Inteligencja autonomia i konkurencja +PL [Dostęp 07.06.2022], Dostępny w: https://www.gov.pl/web/cyfryzacja/konsultacje-spoleczne-projektu-polityki-rozwoju-sztucznej-inteligencji-w-polsce-na lata-2019-2027

- Ministry of Economy, Trade and Industry of Japan (2018) Contract Guidelines on Utilization of AI and Data" (on account of amendments to the Unfair Competition Prevention Act of 2018
- Ministerstwo Gospodarki (2013) Strategia innowacyjności i efektywności gospodarki "Dynamiczna Polska 2020", Załącznik do uchwały nr 7 Rady Ministrów z dnia 15 stycznia 2013 r.
- Ministerstwo Rozwoju, Diagnoza do Strategii Produktywności 2030 (2020) [Dostęp: 07.06.2022], Dostępny w: https://www.gov.pl/attachment/65c9d9ab-57e2-44dd-bf09-0cea82426ccf
- 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
- Miniszewski, M. (2021), Dwie dekady rozwoju polskiego rolnictwa. Innowacyjność sektora rolnego w XXI wieku, Kutwa, K. (współpr.), Polski Instytut Ekonomiczny, Warszawa.
- 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], Dostępny w: https://edukacjamedialna.edu.pl/lekcje/cyfrowy-slad/
- OECD (2019) Recommendation of the Council on Artificial Intelligence [online], OECD Legal Instruments, [Dostęp: 7.06.2022], Dostępny w: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
- OVH (2018) Chmura prywatna ochroni dane wrażliwe [online], virtual-it.pl, [Dostęp: 27.05.2022], Dostępny w: https://www.virtual-it.pl/8436-chmura-prywatna-ochroni-dane-wrazliwe.html
- 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/
- Pawlak, M. 2021. Już prawie połowa Polaków leczy się prywatnie [online], Rzeczpospolita, [Dostęp: 31.05.2022], Dostępny w: https://pieniadze.rp.pl/ubezpieczenia-zycia/art-18940831-juz-prawie-polowa-polakow-leczy-sie-prywatnie
- 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.
- Schneider, G., Health Data Pools under European Policy and Data Protection Law: Research as a New Efficiency Defence?, 11 (2020) JIPITEC 49 para 1.

- 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
- Torchała, K. 2021. Internetowe konto pacjenta posiada już ponad 10 mln osób [online], Bankier.pl, [Dostęp: 31.05.2022], Dostępny w: https://www.bankier.pl/wiadomosc/Internetowe-Konto-Pacjenta-posiada-juz-ponad-10-mln-osob-8150520.html
- Uchwała nr 196 Rady Ministrów z dnia 28 grudnia 2020 r. w sprawie ustanowienia "Polityki dla rozwoju sztucznej inteligencji w Polsce od roku 2020"
- 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., Iwanowski D. (2021). Cyfrowy monopol Nadużycia, których dopuszczają się największe korporacje technologiczne. Instrat Policy Paper 02/2021.
- Wawrzyniak, B., Zygmuntowski, J. J., Lamański, F. (2020) Polska suwerenna cyfrowo. Regulacje na rzecz sprawiedliwej i konkurencyjnej gospodarki cyfrowej. Instrat Policy Paper 06/2020
- Wojciechowski, M. (1998) Zalety i wady architektury rozproszonej wykorzystującej migawki tylko do odczytu, Materiały IV konferencji PLOUG, Zakopane
- 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
- Velotio Technologies, 2019, [Dostęp: 27.05.2022], Dostępny w: https://medium.com/velotio-perspectives/a-beginners-guide-to-edge-computing-6cfea853aa11
- 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
- VMware (2017) Can private cloud be cheaper than public cloud? 41% said yes, and the survey reveals how. VMware, 451 Research, Advisory
- 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., 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/

Załącznik do raportu

Lista wszystkich uczestników warsztatów

01. Przedstawiciel/ka Open Future
02. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
03. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
04. Przedstawiciel/ka Fundacji Instrat
05. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
06. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
07. Przedstawiciel/ka QuantLabs
08. Przedstawiciel/ka Centrum e-Zdrowia; NFZ
09. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
10. Przedstawiciel/ka Departamentu Innowacji KOWR
11. Przedstawiciel/ka Polskiego Instytutu Ekonomicznego
12. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
13. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
14. Przedstawiciel/ka Jutromedical
15. Przedstawiciel/ka Fundacji Instrat
16. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
17. Przedstawiciel/ka Akademii Leona Koźmińskiego
18. Przedstawiciel/ka Polskiego Instytutu Ekonomicznego
19. Przedstawiciel/ka Urzędu m.st. Warszawy
20. Przedstawiciel/ka Departamentu Innowacji KOWR
21. Przedstawiciel/ka HTA
22. Przedstawiciel/ka COT Łukasiewicz
23. Przedstawiciel/ka Fundacji Instrat
24. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
25. Przedstawiciel/ka Krajowego Ośrodka Wsparcia Rolnictwa
26. Przedstawiciel/ka Aida Diagnostics
27. Przedstawiciel/ka Ministerstwa Zdrowia
28. Przedstawiciel/ka alxd
29. Przedstawiciel/ka Zhiva
30. Przedstawiciel/ka Centrum e-Zdrowia
31. Przedstawiciel/ka Ministerstwa Rozwoju i Technologii
32. Przedstawiciel/ka KPRM
33. Przedstawiciel/ka Kancelarii Prezesa Rady Ministrów
34. Przedstawiciel/ka Kancelarii Prezesa Rady Ministrów
35. Przedstawiciel/ka Kancelarii Prezesa Rady Ministrów
36. Przedstawiciel/ka Fundacji Instrat

Prelegenci

37. Przedstawiciel/ka Politechniki Wrocławskiej
38. Przedstawiciel/ka Findata
39. Przedstawiciel/ka Open Data Institute
40. Przedstawiciel/ka OVHCloud