

Dane (nie)dostępne

Fundacja Instrat apeluje do poszczególnych resortów o działania na rzecz dostępności danych o energetyce i górnictwie, gospodarce wodnej, mieszkalnictwie, zdrowiu i dochodach obywateli

Pomimo że Polska zajmuje miejsca na podium w rankingach takich jak Open Data Maturity czy OURdata Index 2023, nadal występują u nas problemy z dostępem do danych z obszarów kluczowych dla społeczeństwa. Wysoka pozycja naszego kraju nie powinna usypiać czujności rządzących, bo stawka jest wysoka - brak dostępu do danych to miliardowe straty dla obywateli, gospodarki i samego państwa.

Dobre wyniki Polski w obszarze otwartych danych to niewątpliwie efekt konsekwentnie rozwijanej przez Ministerstwo Cyfryzacji polityki oraz silnej współpracy wielu instytucji na różnych szczeblach administracji. Co oczywiste, wiodąca pozycja w międzynarodowych rankingach cieszy, jednak nie jest równoznaczna z tym, że nie występują u nas problemy z dostępem do kompletnych informacji będących w posiadaniu urzędów administracji. Co więcej, poza otwartymi danymi, istnieje szereg informacji będących w posiadaniu instytucji publicznych, które choć uznawane są za poufne, mogłyby zostać w bezpieczny sposób udostępnione.

Te kluczowe dla życia społeczno-gospodarczego dane powinny być dostępne powszechnie i nieodpłatnie. Publikowane są jednak w ograniczonym zakresie, charakteryzują się niską jakością bądź formatami o niskim poziomie otwartości.

W niniejszym dokumencie opisaliśmy (tylko!) pięć przykładów mówiących o realnych problemach wynikających z braku dostępności danych. Skupiliśmy naszą uwagę na konkretnych instytucjach publicznych i sporządziliśmy listę obszarów, w których dostrzegamy ograniczone ponowne wykorzystanie informacji. W nadchodzących miesiącach eksperci_tki Instratu prześlą także Ministrom i Pełnomocnikom_czkom ds. otwartości danych pisma z bardziej szczegółowymi rekomendacjami dot. konkretnych obszarów.

(1) Transformacja energetyczna potrzebuje paliwa w postaci... danych

W dobie kryzysu energetycznego dotykającego nas wszystkich i błyskawicznego rozwoju energetyki obywatelskiej, dane energetyczne dawno wyszły poza domenę specjalistów. Społeczeństwo powinno mieć dostęp do informacji o składnikach cenotwórczych, realnych kosztach polityk czy rozwoju technologii.

Jak wskazaliśmy w liście otwartym z 2020 roku, statystyka dotycząca sektora energetycznego i górnictwa prowadzona jest przez podmioty, które celowo nie dzielą się danymi zgodnie ze standardami otwartości. List podpisało ponad 60 ekspertów_ek, w tym byłych ministrów.

Poprzez opublikowanie naszego poprzedniego apelu zwróciliśmy uwagę opinii publicznej i decydentów na problem związany z ograniczonym dostępem do danych z sektora

energetycznego. W wyniku naszej inicjatywy istotne zbiory statystyk zostały udostępnione. Jednak pomimo poprawy sytuacji, ciężko mówić o pełnej transparentności tego sektora.

Przede wszystkim formaty udostępnianych danych znacząco utrudniają prace analityczne i wyciąganie wniosków. Przykładowo, dane dotyczące tego samego zjawiska, pochodzące z różnych instytucji, miewają rozbieżności sięgające kilkudziesięciu procent, a metodologie stojące za procesem ich zbierania nie są publicznie dostępne lub w ogóle nie powstały.

Publikowanie plików z tabelami w formacie PDF jest nadal regularną praktyką. Uwzględnienie potrzeb automatycznego przetwarzania danych na portalach publicznych to rzadkość, a problem ten, mimo że jest dyskutowany od lat, nadal nie został rozwiązany. Ogromne zamieszanie wywołują przebudowy systemów takich podmiotów jak Polskie Sieci Elektroenergetyczne, w wyniku których niespodziewanie zmienia się sposób dostępu do danych, ich format i rozdzielczość.

(2) Dane o gospodarce wodnej potrzebne jak woda

Podobny problem występuje w przypadku danych dotyczących wód w Polsce, co unaoczniała katastrofa ekologiczna na Odrze. Jak wskazaliśmy w raporcie Mętne wody z 2023 r. *“dostępne zestawy danych są rozdzielone między różne agencje rządowe i ich regionalne oddziały. Utrudnione procedury pozyskiwania informacji oraz sposób ich prezentacji sprawiają, że trudno je analizować i monitorować sytuację w gospodarce wodnej”*. Podobne wnioski formułował NIK w swoim raporcie wskazując na brak systemu zarządzania informacją o wodzie w Polsce.

Informacje dotyczące poboru wody i zrzutu ścieków są istotne nie tylko dla obywateli, organizacji ekologicznych, samorządów, lecz przede wszystkim samych instytucji państwowych, które dopuściły do wystąpienia niebezpiecznych stężeń chlorków i siarczanów, a następnie zbyt opieszale zareagowały na kryzys na Odrze w 2022 r.

Gdyby takie działania monitoringowe jak te prowadzone przez Główny Instytut Górnictwa były oparte na dynamicznym ostrzeganiu i otwartych danych, to takim katastrofom dałoby się przeciwdziałać lub nawet ich uniknąć. Niestety, bez otwartych danych o pozwoleniach na korzystanie ze środowiska oraz o faktycznych poborach wód czy zrzutach substancji, niezwłoczne reagowanie w sytuacjach kryzysowych nie jest możliwe. Monitoring środowiska i uzyskiwane z niego dane już teraz są szczególnie istotne w kontekście adaptacji do zmian klimatu, a w przyszłości będą tylko zyskiwać na znaczeniu.

(3) Rekordowe ceny mieszkań i rekordowe marże

Innym przykładem obrazującym to, jak dostęp do danych i posiadanie pełnych informacji może bezpośrednio wpływać na otaczającą nas rzeczywistość, są dane dot. mieszkalnictwa. W dobie kryzysu mieszkaniowego w Polsce rząd, zamiast zadbać o większą transparentność rynku i zapobiec sztuczному zawyżaniu cen, proponuje programy, które jedynie potęgują problem.

Częściowym rozwiązaniem źródeł problemu rekordowych cen i marż w sektorze nieruchomości mogłaby być publicznie dostępna baza danych o transakcyjnych cenach mieszkań. Rejestr pozwoliłby sprawdzić, jak w danym okresie kształtowały się realne średnie

ceny nieruchomości w danej dzielnicy. To zbudowałoby przewagę negocjacyjną co do zasady słabszych kupujących i ułatwiłoby im podjęcie trafnej decyzji ws. mieszkania, ograniczyło działania spekulacyjne oraz możliwość windowania cen przez sprzedających.

W lipcu ub.r. na stronie Ministerstwa Rozwoju i Technologii znalazła się oficjalna informacja o tym, że przepisy, dzięki którym powstać ma Portal Cen Mieszkań zostały przyjęte przez Radę Ministrów. W ostatnim czasie pojawiły się jednak głosy, że Portal został uwzględniony w projekcie ustawy “Kredyt na Start”, a następnie z niej wyłączony, w celu uniknięcia opóźnienia jej wdrożenia.

(4) Słaby przepływ danych dot. zdrowia to niższej jakości usługi medyczne, wyższe koszty dla służby zdrowia i hamowanie innowacji

Rosnąca popularność przechowywania danych w formie elektronicznej nie sprawia, że systemy publiczne i prywatne są dobrze przygotowane do korespondowania ze sobą oraz wzajemnego wykorzystywania gromadzonych przez siebie zasobów (brak interoperacyjności silnej). Informacje o sektorze zdrowia i opiece medycznej publikowane na portalu dane.gov.pl w większości pozyskiwane są od Narodowego Funduszu Zdrowia. Równocześnie, rekordowo niewielu pacjentów korzysta wyłącznie z usług oferowanych przez NFZ (zgodnie z CBOS jedynie 24% badanych). Tym samym, publicznie dostępne zbiory są niedostatecznie reprezentatywne i przedstawiają jedynie wycinek obrazu polskiego systemu opieki zdrowotnej. To natomiast utrudnia prowadzenie przez ekspertów spoza administracji analiz w sferze medycznej.

Nadziei na lepszy przepływ danych pomiędzy różnymi podmiotami można upatrywać we wspólnej europejskiej przestrzeni danych dotyczących zdrowia (EHDS). O ile jednak unijne przepisy powinny stopniowo wchodzić w życie już od jesieni br., o tyle w Polsce prace w tym zakresie nie zostały jeszcze podjęte, a Ministerstwo Zdrowia wskazuje, że wdrażanie unijnego prawa może zająć od dwóch do nawet dziesięciu lat.

Implementację EHDS może utrudniać niski poziom ucyfrowienia dokumentacji medycznej, niejednolite standardy i problemy z interoperacyjnością danych w ochronie zdrowia. Co więcej, sytuację komplikuje wspomniane już rozproszenie informacji o zdrowiu pacjenta pomiędzy podmiotami publicznymi i prywatnymi, a tym samym brak możliwości prześledzenia kompletnej historii leczenia pacjenta.

(5) Czy nierówności dochodowe w Polsce są niskie czy wysokie?

Podstawowym źródłem danych o dochodach obywateli wykorzystywanym w bieżącej debacie publicznej są badania ankietowe GUS. W literaturze naukowej wskazuje się jednak, że tego typu badania nie odzwierciedlają poprawnie struktury dochodów w gospodarce. Przede wszystkim pomijają istotną część dochodów osiągniętych przez najzamożniejszych obywateli.

W konsekwencji przez wiele lat w debacie publicznej w Polsce ignorowano problem nierówności dochodowych - z badań ankietowych wynikało, że są one umiarkowane na tle innych państw oraz nie zmieniają się znacząco. Badania naukowe opublikowane w ostatnich

latach wskazują jednak, że pogląd ten był błędny - nierówności dochodowe w Polsce są wyjątkowo wysokie (jedne z najwyższych w Unii Europejskiej), a ich trend w XXI wieku był rosnący.

Poprawne określenie struktury dochodów w gospodarce wymaga połączenia badań ankietowych z danymi ze źródeł administracyjnych - deklaracji składkowych i podatkowych. Od 2019 roku Ministerstwo Finansów opracowuje zbiory danych, w których znajdują się ogólne informacje w tym obszarze na podstawie danych z deklaracji PIT i danych Zakładu Ubezpieczeń Społecznych. Publikacja opisująca wnioski z tego zbioru ukazuje się jednak nieregularnie i z dużym opóźnieniem, z kolei dane są w niej przedstawiane w wartościach zagregowanych, ograniczając możliwości wykorzystania ich do dalszych analiz np. zmian w systemie podatkowym czy dla celów naukowych. Ministerstwo Finansów nie udostępnia danych jednostkowych na potrzeby badań lub robi to w sposób selektywny na niejasnych zasadach.

Brak dostępu do zanonimizowanych jednostkowych danych z deklaracji składkowych i podatkowych ogranicza potencjał badań nad polityką publiczną w Polsce. Rzetelna ocena wpływu różnego rodzaju planowanych lub działających rozwiązań na dochody i wydatki publiczne, dochody obywateli, nierówności dochodowe czy ubóstwo może zostać dokonana rzetelnie wyłącznie przez Ministerstwo Finansów. To utrudnia tworzenie nowych idei w przestrzeni publicznej, na przykład przez organizacje pozarządowe czy partie polityczne. Obliczeń dokonywanych przez MF nie da się także poddać jakościowej weryfikacji przez instytucje zewnętrzne. Należy podkreślić, że udostępnianie tego typu zanonimizowanych danych jest standardem w wielu rozwiniętych demokracjach.

Apele do poszczególnych ministerstw

Im więcej danych, tym lepsze usługi publiczne, większy wzrost gospodarczy i szybszy rozwój innowacji, a co za tym idzie dobrobyt społeczny. Dlatego jako eksperci zaangażowani w działania na rzecz zasilania debaty i polityk publicznych rzetelnymi analizami, apelujemy do poszczególnych resortów o podjęcie konkretnych kroków celem poprawy dostępności, jakości i standardów udostępniania danych.

1. Do Ministry Klimatu i Środowiska Pauliny Hennig-Kloski, Ministry Przemysłu Marzeny Czarneckiej, Ministra Aktywów Państwowych Jakuba Jaworowskiego oraz Ministra Rozwoju i Technologii Krzysztofa Paszyka apelujemy o:

- **harmonizację państwowych rejestrów i zapewnienie spójności raportowanych danych** pomiędzy wieloma instytucjami współodpowiedzialnymi za udostępnienie podobnych informacji (np. wydobywaniu węgla, emisji zanieczyszczeń do powietrza i zużyciu wody, efektywności energetycznej budynków);
- **reformę modelu prowadzenia statystyki publicznej** dot. gospodarki paliwowo-materiałowej (energetyka i górnictwo) i wprowadzenie standardów otwartości danych oraz likwidację barier w zakresie dostępności danych (odpłatność za dostęp do treści - sprzedaż przez sklepy internetowe ARE i ARP Katowice, formaty plików blokujące maszynowy odczyt danych, np. PDF);
- **opracowanie planu udostępniania danych i modelu zarządzania** (*data governance*) skierowanego na regularne zwiększanie zestawu otwieranych danych w oparciu o zapotrzebowanie administracji rządowej i samorządowej, społeczeństwa obywatelskiego, sektora prywatnego i środowiska eksperckiego;
- **uruchomienie interaktywnego narzędzia do monitoringu transformacji energetyczno-klimatycznej** we wszystkich sektorach polityki klimatycznej (w tym górnictwie) w oparciu o informacje z wielu źródeł i instytucji oraz w oparciu o otwarte licencje, np. w oparciu o rządowy portal dane.gov.pl;
- **mobilizowanie państwowego sektora energetycznego i wydobywczego do dobrowolnej publikacji danych o istotnej społecznej wartości**, np. informacji o zanieczyszczeniu i wykorzystaniu środowiska, zatrudnieniu i oddziaływaniu lokalnym.

2. Do Ministry Klimatu i Środowiska Pauliny Hennig-Kloski o udostępnianie informacji zawartych w pozwoleniach wodno-prawnych i pozwoleniach zintegrowanych oraz o faktycznym poborze wody/zrzucie substancji

Dane o dozwolonym oraz faktycznym korzystaniu ze środowiska, w tym z wód powierzchniowych są niezbędne do rzetelnej oceny wpływu na środowisko wodne, w tym możliwość wystąpienia katastrofy ekologicznej takiej jak miała miejsce w Odrze w 2022 r. Do grona interesariuszy w dziedzinie ochrony wód należą nie tylko instytucje państwowe jak np. Państwowe Gospodarstwo Wodne Wody Polskie, ale też przedsiębiorcy prowadzący działalność np. turystyczną w okolicach wód, a także osoby fizyczne np. wędkarze. Naturalny ekosystem również powinien mieć faktyczną i skuteczną ochronę prawną realizowaną również

przy pomocy osób i organizacji działających w jego obronie. Dane o korzystaniu ze środowiska wodnego powinny być otwarte ze względu na rolę wód powierzchniowych jako dobra wspólnego.

3. Do Ministra Rozwoju i Technologii Krzysztofa Paszyka o kontynuację prac nad Portalem Cen Mieszkań

Dane dotyczące transakcyjnych cen nieruchomości są kluczowe dla przejrzystości rynku nieruchomości - pozwalają prowadzić analizy na jego temat i formułować rekomendacje, które w dalszej kolejności przekładają się na lepsze polityki mieszkaniowe. Dlatego, aby uniknąć pogłębiania się kryzysu mieszkaniowego apelujemy o niezwłoczne uruchomienie Portalu Cen Mieszkań (PCM) - ogólnopolskiego rejestru cen transakcyjnych mieszkań oraz domów. Dzięki większej transparentności rynku mieszkaniowego możliwe będzie ograniczenie windowania cen przez sprzedających oraz wynajmujących, a także weryfikowanie zasadności programów takich jak „Bezpieczny kredyt 2%” czy kredyt „Mieszkanie na Start”.

4. Do Ministry Zdrowia Izabeli Leszczynej o wspieranie rozwoju interoperacyjności silnej pomiędzy placówkami oraz budowanie otoczenia prawnego i instytucjonalnego dla wspólnej przestrzeni danych dot. zdrowia

Interoperacyjność danych ma kluczowe znaczenie dla powszechnie dostępnego i efektywnego kosztowo systemu opieki zdrowotnej (pierwotne wykorzystanie) oraz rozwoju innowacji w sektorze zdrowia (wykorzystanie wtórne). Konieczne jest więc zapewnienie przepływu danych pomiędzy poszczególnymi placówkami medycznymi (zarówno publicznymi, jak i prywatnymi) oraz instytucjami naukowo-badawczymi.

Cel ten można osiągnąć poprzez sprawne wdrożenie europejskiej wspólnej przestrzeni danych dotyczących zdrowia. Aby jednak EHDS odniosło sukces niezbędna jest: pełna digitalizacja dokumentacji medycznej (obecnie dokumentacja ta służy głównie lekarzom i często występuje w formie pisemnej); objęcie obligatoryjnymi jednolitymi standardami wszystkich podmiotów medycznych i wprowadzenie ram dla interoperacyjności silnej (wymiany danych pomiędzy jednostkami sektora publicznego i prywatnego); jak najszybsze przygotowanie otoczenia instytucjonalnego dla EHDS poprzez wskazanie godnego zaufania organu ds. dostępu do danych dot. zdrowia (*Health Data Access Body*), który odpowiadać będzie za procesowanie zapytań i udzielanie zgód dostępu do danych.

5. Do Ministra Finansów o udostępnienie zanonimizowanych danych jednostkowych z połączonej bazy danych deklaracji PIT i danych ZUS na potrzeby badań naukowych

Nasz apel w tym zakresie obejmuje trzy elementy:

- stworzenie data-hub, w ramach którego będą przygotowywane, zanonimizowane i przechowywane dane z deklaracji podatkowych (w kolejnym kroku także składkowych) - w sposób zapewniający bezpieczeństwo tych danych.

- Ustalenie procedury dostępu do tych danych do celów badawczych. Procedura powinna obejmować szkolenie oraz być obwarowana klauzulami chroniącymi dane osobowe.
- Regularne publikowanie (z maksymalnie rocznym opóźnieniem) podstawowych statystyk o strukturze dochodów obywateli oraz płaconych przez nich podatkach i składkach na podstawie posiadanych przez Ministerstwo Finansów danych administracyjnych.

Co resorty i podległe im instytucje mogą zmienić już dziś?

Jednolite standardy dzielenia się danymi

- zadbać o lepszą dostępność i jakość danych we wskazanych powyżej obszarach;
- zapewniać dostęp do wszystkich publikowanych tabel w formie edytowalnych plików (XLSX, CSV);
- zwrócić uwagę na kwestie standardów danych i dobrych praktyk, np. zapewnianie zgodności wstecznej zbiorów, utrzymanie starych serwisów w okresach przejściowych;
- uzupełniać zbiory o metadane, informować o dacie zamieszczenia, aktualizacji, źródłach, harmonogramie utrzymania.

Budowanie potencjału administracji publicznej

- podejmować działania mające na celu podnoszenie świadomości pracowników administracji publicznej oraz ich kompetencji cyfrowych;
- wzmacniać współpracę pomiędzy pełnomocnikami ds. otwartości danych w ministerstwach, Kancelarii Prezesa Rady Ministrów oraz Głównym Urzędzie Statystycznym, a także promować współpracę pełnomocników z podmiotami zewnętrznymi zainteresowanymi ponownym wykorzystywaniem danych publicznych (think-tankami, ośrodkami naukowo-badawczymi, niezależnymi ekspertami);
- zapewnić instytucjom publicznym środki (np. z programu FERC) oraz mandat na rozwój i współpracę na froncie udostępniania danych.

Komunikacja i współpraca

- oczekiwać od pełnomocników ds. otwartości danych lepszego komunikowania ich działań szerszej publiczności;
- konsultować potrzeby użytkowników przy rozwijaniu usług i serwisów związanych z udostępnianiem danych;
- zapowiadać zmiany w sposobie udostępniania danych przez różne instytucje z odpowiednim wyprzedzeniem, umieszczając czytelne komunikaty w widocznych miejscach.

Do wiadomości

1. Anna Gos, Dyrektorka Departamentu Zarządzania Danymi MC, Pełnomocniczka Ministra Cyfryzacji ds. otwartości danych
2. Agencja Rynku Energii
3. Agencja Rozwoju Przemysłu, Oddział w Katowicach
4. Główny Urząd Nadzoru Budowlanego
5. Główny Instytut Górnictwa
6. Polskie Sieci Elektroenergetyczne
7. Polskie Towarzystwo Przesyłu i Rozdziału Energii Elektrycznej
8. Polskie Towarzystwo Elektrociepłowni Zawodowych
9. Towarzystwo Gospodarcze - Polskie Elektrownie
10. Zakład Ubezpieczeń Społecznych
11. Ministerstwo Cyfryzacji
12. Ministerstwo Funduszy i Polityki Regionalnej
13. Centrum Projektów Polska Cyfrowa
14. Instytut Finansów Publicznych

Kontakt

Blanka Wawrzyniak, blanka.wawrzyniak@instrat.pl, 668 487 653

Michał Bukowiński, michal.bukowinski@instrat.pl, 504 643 695