



Otwieranie danych

Podręcznik dobrych praktyk



Spis treści

1. **Co to są dane i dlaczego warto je wykorzystywać** ————— 4
2. **Co to znaczy, że dane są otwarte** ————— 9
3. **Otoczenie prawne i sposoby publikacji danych** ————— 15
 - 3.1 Dostęp do informacji publicznej ————— 16
 - 3.2 Ponowne wykorzystywanie informacji sektora publicznego ————— 18
 - 3.3 Ponowne wykorzystywanie informacji sektora publicznego a ochrona danych osobowych ————— 20
4. **W jaki sposób udostępniać dane publiczne** ————— 22
 - 4.1 Portale otwartych danych ————— 23
 - 4.2 Wizualizacje danych ————— 26
5. **Jak zorganizować pracę w urzędzie** ————— 28
 - 5.1 Opracowanie i wdrażanie procesu otwierania danych. Rola osoby kierującej urzędem i lidera/liderki zespołu. ————— 30
 - 5.2 Powołanie zespołu ————— 33
 - 5.3 Opracowanie polityki otwartości ————— 34
6. **Schemat wdrażania polityki otwartości. Otwieranie zbiorów danych** ————— 39

Jeśli zastanawiacie się nad otwieraniem danych w swoim urzędzie albo nawet już to robicie i szukacie sposobów na usprawnienie tych działań [rozdz. 6] - koniecznie przeczytajcie ten podręcznik.

Dzielimy się w nim swoimi doświadczeniami w otwieraniu danych publicznych, jak również prezentujemy spostrzeżenia samorządów, z którymi rozmawialiśmy.

Otwieranie danych to proces, który nie ma swojego wyraźnego początku ani końca, a wiele zależy od czynników takich jak regulacje prawne, otoczenie instytucjonalne, poziom współpracy między instytucjami publicznymi, czy ogólne nastawienie urzędników i urzędniczek do jawności życia publicznego. Podręcznik opisuje podstawowe ramy tego procesu, odsyłając do odpowiednich regulacji prawnych, wskazując na pożądane ramy instytucjonalne i przedstawiając praktyczne scenariusze otwierania danych w urzędach.

Podręcznik napisaliśmy dla kierujących i pracujących w urzędach administracji rządowej i samorządowej, przedstawicieli i przedstawicielek organizacji pozarządowych oraz wszystkich, którzy na podstawie danych tworzonych i przetwarzanych w instytucjach publicznych budują swoją wiedzę o działaniach administracji. Skupiliśmy się na omówieniu najważniejszych przepisów, mających wpływ na otwieranie i wykorzystywanie danych publicznych. Pokazaliśmy modele współpracy wewnątrz i pozainstytucjonalnej, które najlepiej sprawdzają się w tym kontekście. Dowiedziecie się również, w jaki sposób praktycznie realizować proces otwierania danych. Jeśli szukasz wskazówek technicznych, znajdziesz je w standardach otwartości danych: [technicznym](#), [API \[Application Programming Interfaces\]](#) i [bezpieczeństwa](#).



1.

Co to są dane
i dlaczego warto je
wykorzystywać?



Zacznijmy od tego, czym w ogóle są **dane**. Słownik języka polskiego odwołuje się do dwóch znaczeń.

Dane to:

1. «fakty, liczby, na których można się oprzeć w wywodach»
2. «informacje przetwarzane przez komputer»

W rządowym [Programie otwierania danych publicznych](#) (dalej: Program) znajdziecie z kolei definicję danych publicznych, czyli takich, które są w posiadaniu urzędów, niezależnie od tego, kto je wytworzył.



Dane publiczne - liczby i pojedyncze wydarzenia lub obiekty na możliwie najniższym poziomie agregacji, które nie zostały poddane przez administrację publiczną przetworzeniu do postaci raportów, wykresów itp. oraz nie został im nadany odpowiedni kontekst lub interpretacja.

Trudno jest wymienić wszystkie przykłady danych, którymi urzędy posługują się w swojej codziennej pracy. Są to m.in. informacje o dochodach, wydatkach, edukacji, mieszkalnictwie, poziomie zanieczyszczenia powietrza, rejestrach lokalnych przedsiębiorstw, oświacie (w tym wynikach egzaminów), bibliotekach, atrakcjach turystycznych, pozwoleniach na budowę, hotelach, systemach wodnych, portach lotniczych i wiele innych.

Rekomendacje

Dokonajcie analizy najbardziej popularnych zasobów wykorzystywanych przez podmioty zewnętrzne.

Obywatele mają prawo wiedzieć o każdej działalności publicznej, a wykorzystywanie danych może pozytywnie wpłynąć na poprawę jakości ich życia.

W Hiszpanii dużą popularnością cieszą się np. [dane o stacjach benzynowych i oferowanych przez nie cenach paliw](#), a w Polsce z powodzeniem od lat rozwijają się [serwisy przetwarzające dane komunikacyjne](#). Popularne są również serwisy wykorzystujące dane publikowane online, jak np. zawarte w [Krajowym Rejestrze Sądowym](#). To doskonale obrazuje, że z ogólnodostępnych danych organizacje i firmy potrafią zbudować użyteczne produkty.

Dane publicznie mają wiele zastosowań, a ich otwarcie pozwala realizować cele takie jak:

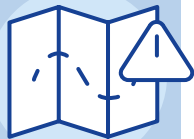
Poprawa funkcjonowania instytucji publicznych i ułatwienie pracy urzędników i urzędniczek

W Programie napisaliśmy: „administracja publiczna nie zawsze dostrzega, że sama jest głównym beneficjentem otwierania danych. Uwalnianie danych publicznych wpływa na poprawę wymiany wiedzy i informacji pomiędzy jednostkami administracji oraz w kontaktach z obywatelami. Dane mogą być wykorzystywane przez inne podmioty publiczne do tworzenia innowacyjnych usług dla obywateli”. Jak pokazało [Ministerstwo Finansów, otwieranie danych może służyć nawet walce z przestępczością skarbową](#). Przykładów wykorzystywania danych w administracji jest wiele, poczynając od skomplikowanych baz danych jak PESEL czy CEPIK po mniej złożone, jak np. ewidencja wniosków o udostępnienie informacji publicznej wpływających do urzędu.

Na otwartości danych zyskują również autorzy polityk, dokumentów strategicznych i projektów aktów prawnych. Ogromnym mankamentem Ocen Skutków Regulacji pozostaje jednak brak odniesień do konkretnych danych czy analiz powstałych na ich podstawie. Dzięki dostępowi do danych ta sytuacja może ulec zmianie.



Zwiększenie zaangażowania obywateli w sprawy publiczne



Udostępnianie danych może poprawić wiedzę o działaniach władzy ustawodawczej i wykonawczej, a tym samym zapewniać bardziej przejrzysty dostęp do prawa. Na podstawie [danych pobieranych z Sejmu](#) można śledzić przebieg debaty publicznej oraz [informacje o swoich reprezentantach](#). Trudno się zgodzić ze słowami autorów [Krajowej mapy zagrożeń bezpieczeństwa](#), że „poczucie bezpieczeństwa należy do podstawowych potrzeb każdego człowieka i ma decydujący wpływ na funkcjonowanie każdej społeczności. Dlatego ważne jest tworzenie narzędzi, które pozwolą na rzetelne i czytelne zidentyfikowanie i przedstawienie, w tym społecznościom lokalnym, skali i rodzaju zagrożeń oraz instytucji współodpowiedzialnych za zapewnienie bezpieczeństwa i porządku publicznego”. Dane budują również wiedzę o lokalnych i ogólnokrajowych problemach i zjawiskach, np. ekologicznych lub społecznych.

Budowanie środowiska innowacyjnego



Otwarte dane to wsparcie dla rozwoju biznesu i start-upów, a także rozwijanie rynku pracy. Według danych [Komisji Europejskiej](#) rynek usług i produktów wykorzystujących otwarte dane ma wzrosnąć do wartości 75,7 mld EUR w 2020 r. Przewiduje się również, że do 2020 roku liczba osób pracujących w sektorach opartych na otwartych danych wzrośnie do około 100 tys. Intensywny rozwój innowacji technologicznych możemy obserwować na razie za granicą, gdzie powstają m.in. serwisy oparte na danych dotyczących [rynku pracy](#) czy wspierające [działania straży pożarnej](#). W Polsce najbardziej popularne pozostają serwisy prezentujące rozkłady jazdy czy informacje na temat partnerów biznesowych.

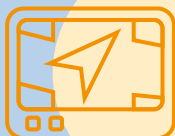
Rozwiązywanie problemów społecznych i gospodarczych

Na proces otwierania danych musimy patrzeć całościowo, jako na element realizacji strategii państwa. Kwestie otwierania danych i ułatwienia ich ponownego wykorzystywania znalazły się m.in. w [Strategii Sprawne Państwo 2020 oraz Strategii Innowacyjności i Efektywności Gospodarki „Dynamiczna Polska 2020”](#).

W oparciu o dane publiczne powstają strategie rozwiązywania problemów na poziomie lokalnym i ogólnokrajowym. Oprócz wspomnianych wcześniej kwestii bezpieczeństwa i zdrowia można wyróżnić te kategorie danych, które pozwalają na prowadzenie racjonalnej [polityki oświatowej](#), [promocję sportu](#) czy [rozwój gospodarki](#). Same rozwiązania technologiczne nie spowodują, że znikną problemy typowe dla danego obszaru. Zdaniem [OECD](#), [...] wyzwaniem nie jest wprowadzenie cyfrowych narzędzi do administracji. Wyzwaniem jest ich użycie do jej modernizacji”. I właśnie szerokie korzystanie z danych publicznych, zwłaszcza otwartych danych, przez administrację to jeden z warunków poprawy jej efektywności. **Dlaczego to takie ważne?** Aby działania państwa były skuteczne, muszą być na każdym etapie: od planowania przez monitorowanie po ewaluację zderzane z faktami, czyli danymi o natężeniu danego zjawiska i zmianach, jakie zaszły. Co więcej, otwarty dostęp do takich informacji zwiększa transparentność działań administracji i umożliwia zaangażowanie się środowisk naukowych, pozarządowych, eksperckich we współpracę przy rozwiązywaniu problemów społeczno-ekonomicznych i kształtowaniu polityk publicznych.

Sprawdzacie w aplikacji rozkład jazdy, korzystacie z nawigacji?

To właśnie są efekty otwartych danych. Otwarte dane to większa przejrzystość działań państwa i lepsza jakość polityk publicznych. Otwarte dane pomagają rozwijać biznes i ułatwiają pracę naukowcom.



2. Co to znaczy, że dane są otwarte?



Dane, aby efektywnie służyły do realizacji wskazanych wyżej celów, powinny być:

1. przygotowane w sposób przyjazny dla użytkownika/ użytkownicy - kluczowy jest komunikatywny, zrozumiały dla odbiorcy język;
2. spełniające wszystkie filary otwartości, w których jest odniesienie do kwestii prawnych, udostępnione w formacie przeznaczonym do odczytu maszynowego¹, i w otwartych formatach² do ponownego wykorzystywania³.

Bardziej szczegółowe cechy otwartości danych publicznych można **odnaleźć w standardach otwartości danych: [technicznym](#), [API](#) i [bezpieczeństwa](#)** oraz w Programie, który odwołując się do **[standardów międzynarodowych](#)**, w tym opracowanych przez **[Komisję Europejską](#)**, określa, że są to dane:

1. **dostępne** - udostępnione bez żadnych ograniczeń szerokiego gronu użytkowników/użytkowniczek (w szczególności: obywatele, firmy, uczelnie, instytucje) do dowolnych celów;
2. **upublicznione w wersji źródłowej** - dostępne w oryginalnej i niezmienionej formie, nie w postaci analiz, podsumowań, skrótów czy streszczeń, tak aby możliwe było np. łączenie danych z różnych źródeł;
3. **kompletne** - udostępnione w całości;
4. **aktualne** - udostępnione na tyle szybko, aby zachować wartość tych danych;

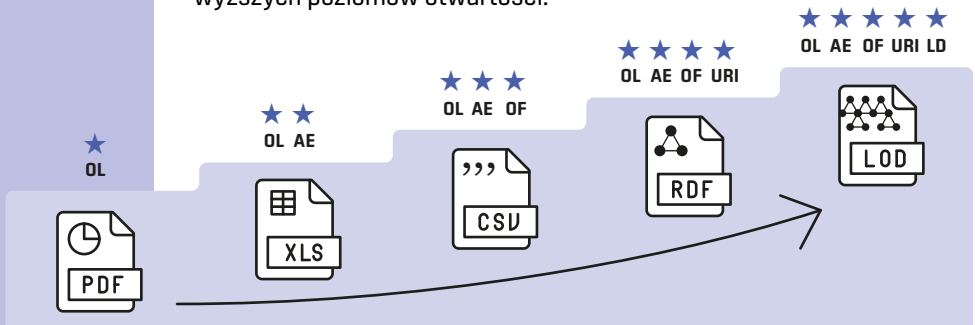
¹ **Format przeznaczony do odczytu maszynowego** - format pliku zorganizowany tak, aby aplikacje komputerowe mogły łatwo zidentyfikować, rozpoznać i uzyskać określone dane, w tym poszczególne stwierdzenia faktów i ich wewnętrzną strukturę.

² **Format otwarty** - format pliku, który nie jest powiązany z platformą oraz jest udostępniany obywatelom bez żadnych ograniczeń, które utrudniałyby ponowne wykorzystywanie dokumentów.

³ O ponownym wykorzystaniu jego definicji i zasadach znajdziecie więcej informacji w dalszej części podręcznika.

5. **odczytywalne maszynowo** - udostępnione w formatach przeznaczonych do odczytu maszynowego i formacie otwartym. Przykładem takich formatów jest CSV, XML, arkusz kalkulacyjny. Zazwyczaj trudne do maszynowego odczytu są formaty PDF, HTML czy pliki tekstowe, które stają się użyteczne do ponownego wykorzystywania dopiero po przełożeniu ich na jeden z formatów otwartych;
6. **udostępnione niedyskryminująco** - dostępne dla każdego bez konieczności rejestracji, weryfikacji tożsamości poprzez podawanie hasła, loginu czy podpisywania jakichkolwiek umów;
7. **dostępne bez ograniczeń prawnych** - dane nie są przedmiotem praw autorskich, patentów, znaków towarowych lub tajemnicy handlowej i mogą być wykorzystywane w dowolnych celach bez konieczności ubiegania się o jakąkolwiek zgodę na ich używanie;
8. **niezastrzeżone** - dostępne w formacie powszechnie stosowanym, który nie jest kontrolowany przez żaden podmiot.

Chcąc uchwycić różnicę między danymi, które są otwarte, a tymi, do których dostęp - a więc również i możliwość ich wykorzystywania - jest utrudniony, należy odwołać się do **systemu 5 poziomów otwartości**. System został opracowany przez prekursora Internetu, Tima Berners-Lee. Poniżej pokazujemy, że samo zamieszczenie pliku z informacjami w Internecie jest jedynie spełnieniem pierwszego z pięciu poziomów otwartości. Warto umożliwić mieszkańcom, przedstawicielom biznesu, studentom czy organizacjom pozarządowym korzystanie z kolejnych, wyższych poziomów otwartości.



gdzie poszczególne gwiazdki oznaczają:

- ★ Udostępnienie danych w sieci Web (w dowolnym formacie) na warunkach otwartej licencji
- ★★ Udostępnienie danych w formie ustrukturyzowanej (np. arkusz kalkulacyjny zamiast zeskanowanego obrazu tabeli)
- ★★★ Używanie formatów otwartych (np. CSV zamiast arkusza kalkulacyjnego)
- ★★★★ Używanie URI do oznaczania zasobów, aby możliwe było ich wyszukiwanie
- ★★★★★ Łączenie danych, aby zapewnić kontekst

★ Udostępnienie danych w formie tabeli w pliku PDF

TABL. 4. NAJDŁUŻSZE RZEKI
THE LONGEST RIVERS

Rzeki Rivers	Recypient ^a Recipient ^a	Długość w km Length in km		Powierzchnia dorzecza w km ² Drainage basin area in km ²	
		ogółem total	w tym w Polsce of which in Poland	ogółem total	w tym w Polsce of which in Poland
Wisła	Morze Bałtyckie Baltic Sea	1022	1022	193960 ^b	168868 ^c
Odra		840	726	119074	106043
Warta	Odra	795	795	54520	54520
Bug	Narew	774	590	38712	19239 ^c
Narew	Wisła	499	443	74527	53846
San	Wisła	458	457	16877	14426
Noteć	Warta	391	391	17302	17302
Wieprz	Wisła	349	349	10497	10497
Pillica	Wisła	333	333	9258	9258
Bóbr	Odra	279	276	5874	5830

• jako plik tekstowy ★★★

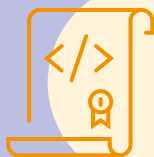
```
Nazwa rzeki;Recipient;Długość w km ogółem;Długość w km w Polsce;Powierzchnia dorzecza w km2 ogółem;Powierzchnia dorzecza w km2 w Polsce
Wisła;Morze Bałtyckie;1022;1022;193960;168868
Odra ;Morze Bałtyckie;848;726;119074;106043
Warta ;Odra;795;795;54520;54520
Bug;Narew;774;590;38712;19239
Narew;Wisła;499;443;74527;53846
San;Wisła;458;457;16877;14426
Noteć;Warta;391;391;17302;17302
Wieprz;Wisła;349;349;10497;10497
Pillica;Wisła;333;333;9258;9258
Bóbr;Odra;279;276;5874;5830
```

Powyższe przykłady łączy podobny sposób zapisu liczb. **To kluczowa sprawa w przygotowaniu się do otwierania danych.**

Rekomendacje

Zapisujcie dane zgodnie z jednolitym standardem. Dzięki temu komputer będzie „wiedział”, która dana odnosi się do liczby, która do daty, a która stanowi nazwę danej kategorii. Umożliwi to porównywanie, analizę lub łączenie danych ze sobą, a nawet tworzenie na ich podstawie wizualizacji, produktów i usług.

Dane publiczne zwykle nie są tworzone przez jedną osobę, a standardy zapisu danych rzadko są wdrażane w instytucjach współpracujących przy przygotowaniu i udostępnianiu danych, więc dużo pracy wymaga „czyszczenie” danych i dostosowywanie ich do różnych potrzeb. Dlatego, **rekomendujemy przygotowywanie danych zgodnie z określonym standardem już w momencie ich zbierania.**



Ministerstwo Cyfryzacji przedstawiło [szczegółowe standardy techniczne otwartości danych, uwzględniające różnice wynikające z 5 poziomów otwartości](#). Przed rozpoczęciem procesu otwierania danych zachęcamy do zapoznania się z całym dokumentem.

3. Otoczenie prawne i sposoby publikacji danych



Udostępnienie danych publicznych budzi od wielu lat wątpliwości i powoduje, że urzędy - na wszelki wypadek - wolą ich nie rozpowszechniać. Postaramy się rozwiązać te obawy poniżej. Bardziej szczegółowe informacje o sposobach radzenia sobie np. z anonimizacją otwartych danych odnajdziecie w dwóch opracowaniach przygotowanych przez Ministerstwo Cyfryzacji: [standardzie bezpieczeństwa](#) oraz [standardzie prawnym](#).

3.1 Dostęp do informacji publicznej

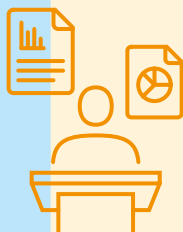
Dostęp do danych publicznych odbywa się na zasadach określonych w [ustawie z dnia 6 września 2001 r. o dostępie do informacji publicznej](#).

Zgodnie z definicją informacji publicznej przedstawioną przez Naczelny Sąd Administracyjny w jednym z orzeczeń już w 2002 r.⁴ i często przywoływaną przez sądy administracyjne:

Informacją publiczną jest „każda wiadomość wytworzona przez szeroko rozumiane władze publiczne oraz osoby pełniące funkcje publiczne, a także inne podmioty, które tę władzę realizują, bądź gospodarują mieniem komunalnym lub majątkiem Skarbu Państwa, w zakresie swych kompetencji. Taki charakter będzie miała również **wiadomość niewytworzona** przez podmioty publiczne, lecz **odnosząca się do tych podmiotów**. Stąd informacją publiczną jest każda wiadomość wytworzona lub odnosząca się do władz publicznych, a także wytworzona lub odnosząca się do innych podmiotów wykonujących funkcje publiczne w zakresie wykonywania przez nie zadań władzy publicznej i gospodarowania mieniem komunalnym lub majątkiem Skarbu Państwa”.

Przed udostępnieniem danych publicznych sprawdźcie jednak, czy nie zachodzą, wymienione w art. 5 ust. 1 i 2 ustawy o dostępie

⁴ Wyrok z 30 października 2002 r. II SA 181/02 (niepubl)



do informacji publicznej, przesłanki ograniczające przekazanie ich każdej zainteresowanej osobie. Zgodnie z tym przepisem, można odmówić udostępnienia danych publicznych jeśli stanowią one **tajemnicę przedsiębiorcy**, narusza to **prawo do ochrony prywatności** lub odnoszą się do **innych przewidzianych prawem tajemnic**. Warto odnotować, że zgodnie z art. 61 ust. 3 Konstytucji Rzeczypospolitej Polskiej, owe tajemnice muszą być wprost wymienione w innych przepisach rangi ustawowej. Jako przykłady takich tajemnic można wskazać **tajemnicę skarbową** dotyczącą danych identyfikujących indywidualnego podatnika, **tajemnicę statystyczną** chroniącą dane osób, od których zebrano dane, czy wreszcie **informacje niejawne**, których udostępnienie może zagrozić bezpieczeństwu publicznemu. W tym ostatnim przypadku sprawa udostępnienia danych lub niemożności ich przekazania wnioskodawcy wydaje się stosunkowo prosta. Jeśli informacje zostały uznane za niejawne, będą opatrzone przez upoważnioną do tego osobę odpowiednimi klauzulami tajności.

Również w przypadku tajemnicy przedsiębiorcy, to na nim, zgodnie z art. 11 ust. 4 **ustawy z dnia 16 kwietnia 1993 r. o zwalczaniu nieuczciwej konkurencji**, ciąży obowiązek należytego zastrzeżenia określonych danych jako tajemnicy przedsiębiorcy. Należy jednak pamiętać, że **nie może tego czynić zupełnie dowolnie**.

Więcej problemów pojawia się na styku prawa do informacji publicznej i ochrony prawa do prywatności. Są to często sprawy indywidualne, wymagające przeanalizowania każdego przypadku. W głośniejszej sprawie ujawnienia danych osób, które nie wykonują funkcji publicznych, ale zawarły z urzędem umowy na wykonanie usług lub dostarczenie towarów, **Sąd Najwyższy** wyraźnie wskazał, że ta forma relacji obywatela z urzędem przesądza o tym, że jego prawo do zachowania w tajemnicy danych osobowych podlega ograniczeniu.

Pamiętajmy, że z ochrony prywatności nie będą również korzystały osoby, których dane podawane są w kontekście faktu, że **reprezentują osoby prawne lub przepisy prawa wprost wskazują, że ich dane są podawane do publicznej wiadomości** - np. **dane osób, którym umorzono obowiązek wpłaty podatków lokalnych**.



3.2 Ponowne wykorzystywanie informacji sektora publicznego

Otwarte dane mogą stanowić surowiec do budowy aplikacji, serwisów czy usług. Ramy prawne dla ponownego wykorzystywania wyznacza ustawa z dnia 25 lutego 2016 r. o ponownym wykorzystywaniu informacji sektora publicznego. Ustawa ta wdraża regulacje dyrektywy 2003/98/WE Parlamentu Europejskiego i Rady z dnia 17 listopada 2003 r. w sprawie ponownego wykorzystywania informacji sektora publicznego zmienionej dyrektywą 2013/37/UE Parlamentu Europejskiego i Rady z dnia 26 czerwca 2013 r.



Informacja sektora publicznego to każda treść lub jej część, niezależnie od sposobu utrwalenia, w szczególności w postaci papierowej, elektronicznej, dźwiękowej, wizualnej lub audiowizualnej, będąca w posiadaniu podmiotów zobowiązanych (czyli instytucji publicznych i niektórych spółek z udziałem mienia publicznego).

W praktyce informacja sektora publicznego opiera się na tym samym zakresie przedmiotowym co definicja informacji publicznej, z tą różnicą, że może zawierać jeszcze inne dane, takie jak meteorologiczne czy materiały muzealne.



Ponowne wykorzystywanie - wykorzystywanie przez osoby fizyczne, osoby prawne i jednostki organizacyjne, nieposiadające osobowości prawnej, informacji sektora publicznego w celach komercyjnych lub niekomercyjnych, innych niż pierwotny cel, dla którego informacja została wytworzona.

Tę definicję najlepiej przedstawić na przykładzie kulinarnym. Jeżeli efektem ponownego wykorzystywania informacji publicznej ma być wypiek szarlotki, czyli np. serwis udostępniający obywatelom dane transportowe, to potrzebne są produkty takie jak jabłka, mąka

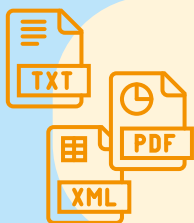
i cukier, tj. dane dotyczące tras, rozkładów jazdy, rozmieszczenia przystanków. Zadbajcie o to, żeby składniki były dobrej jakości i świeże. Bez spełnienia tych warunków żaden wypiek się nie uda. Składniki te trzeba ze sobą odpowiednio połączyć, by cieszyć się deserem - sprawnie działającym serwisem transportowym.

Rekomendacje

**Zadbajcie o to, żeby dane były dobrej jakości i aktualne.
Bez spełnienia tych warunków żaden serwis się nie uda.**

W przypadku udostępniania lub przekazywania danych do ponownego wykorzystywania przy użyciu systemów teleinformatycznych trzeba pamiętać o obowiązku stosowania formatów danych wymienionych w rozporządzeniu 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 [Dz. U. z 2017, poz. 2247]. Istotne jest, jeśli to tylko możliwe, aby stosować otwarte formaty przeznaczone do odczytu maszynowego wraz z metadanymi.

Formaty do danych zawierających dokumenty tekstowe, tekstowo-graficzne lub multimedialne: TXT, RTF, PDF, XPS, ODT, ODS, ODP, DOC, XLS, PPT, DOCX, XLSX, PPTX, CSV. Do definiowania układu informacji polegającego na określeniu elementów informacyjnych oraz powiązań między nimi stosuje się m.in. XML




Więcej na temat ponownego wykorzystywania informacji sektora publicznego, w tym m.in. dopuszczalnych **w warunków ponownego wykorzystywania**, procedury **udostępnienia informacji na wniosek**, kwestii **naliczania opłat**, a także **stosowania ewentualnych licencji** możecie dowiedzieć się z [publikacji Ministerstwa Cyfryzacji „Ponowne wykorzystywanie informacji sektora publicznego”](#).

3.3. Ponowne wykorzystywanie informacji sektora publicznego a ochrona danych osobowych

Informacje sektora publicznego, które są przekazywane każdemu zainteresowanemu do ponownego wykorzystywania, mogą zawierać dane osobowe. Wówczas ponowne wykorzystywanie tego rodzaju informacji lub ich części będzie wykluczone lub ograniczone. Zawsze jednak, gdy zagrożona jest ochrona prywatności i ochrona danych osobowych, trzeba zastosować zrównoważone podejście, mając na uwadze, że przepisy o ochronie osobowych oraz o ponownym wykorzystywaniu są równorzędne.

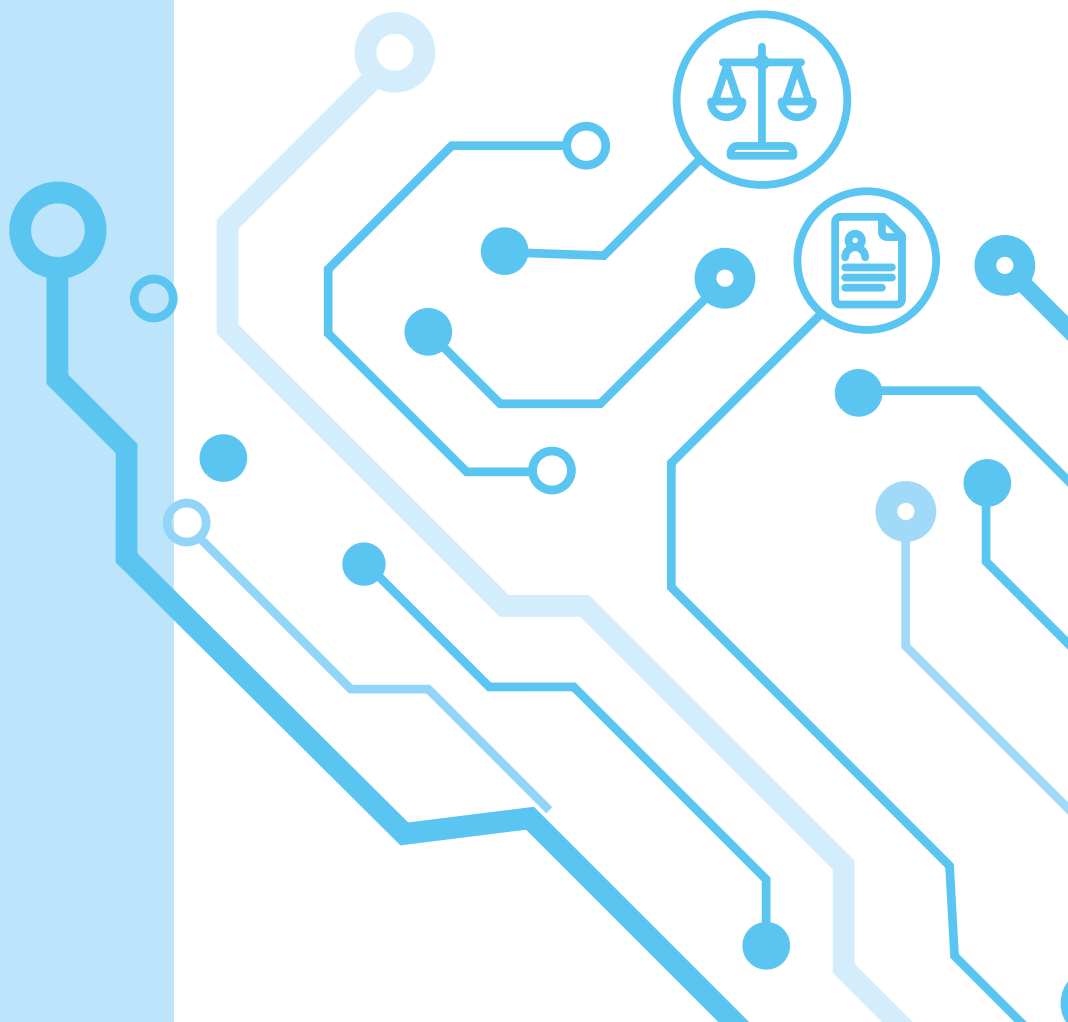
RODO kontynuuje w tym zakresie [opinię Grupy Roboczej art. 29 \(niezależnego zespołu przedstawicieli organów nadzorujących ochronę danych osobowych w państwach członkowskich UE\)](#), w której stwierdzono m.in., że:



Dalsze wykorzystanie dla większej przejrzystości już ujawnionych danych osobowych (na przykład ponowna publikacja danych przez prasę lub dalsze rozpowszechnienie pierwotnie opublikowanego zbioru w bardziej innowacyjny oraz przyjazny użytkownikowi/ użytkownicze sposób przez organizacje pozarządowe) może także być pożądane. To czy taka ponowna publikacja oraz ponowne wykorzystanie jest możliwe, będzie także zależeć od rezultatu testu równowagi, przy którym powinno się brać pod uwagę, między innymi, charakter informacji oraz skutek ponownej publikacji oraz ponownego wykorzystania dla osób.

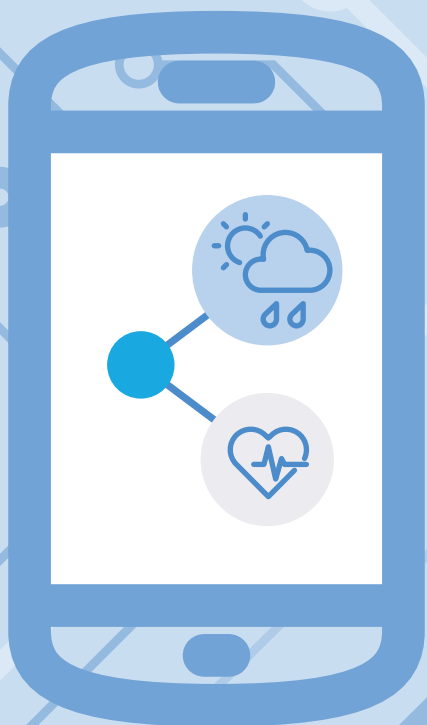
W związku z tym RODO nie wprowadza, co do zasady, ograniczenia ponownego wykorzystywania informacji sektora publicznego, w kontekście danych osobowych, które nie dotyczą wyłącznie sfery życia prywatnego. Należy jednak dochować wszelkich starań, aby ocenić, czy w danym przypadku dopuszczalna jest zmiana celu, dla którego dane były pobierane.

Podobnie jak w przypadku udostępnienia danych na podstawie przepisów o dostępie do informacji publicznej, najlepiej, w razie jakichkolwiek wątpliwości, skonsultować się z osobą odpowiedzialną za politykę ochrony danych osobowych w urzędzie.



4.

W jaki sposób
udostępniać dane
publiczne?



4.1 Portale otwartych danych

Rekomendacje

Publikujcie dane online. Zagwarantuje to wykorzystanie potencjału otwartych danych.

W zakresie trybu udostępniania danych na wniosek odsyłamy do przywołanego w rozdz. III.2 tekstu dotyczącego ponownego wykorzystywania informacji sektora publicznego. Należy tylko wspomnieć o - przewidzianym w art. 21 ust. 2 ustawy o ponownym wykorzystywaniu informacji sektora publicznego - wniosku cyklicznym, czyli realizacji wniosku poprzez umożliwienie przez okres nie dłuższy niż 12 miesięcy, ponownego wykorzystywania, w sposób stały i bezpośredni w czasie rzeczywistym, informacji sektora publicznego gromadzonych i przechowywanych w systemie teleinformatycznym podmiotu zobowiązanego.

Niektóre urzędy, takie jak w [UM Gdynia](#) lub [UM Warszawa](#) tworzą własne **portale otwartych danych** zawierające dane wytworzone przez urząd i jego jednostki lub odnoszące się do realizowanych przez nie zadań. Portale otwartych danych ułatwiają - szczególnie na poziomie miast - odszukanie konkretnych informacji, gdyż z założenia stanowią repozytorium większości zestawów danych, które są możliwe do ponownego wykorzystania w danym urzędzie lub w jego jednostkach organizacyjnych.

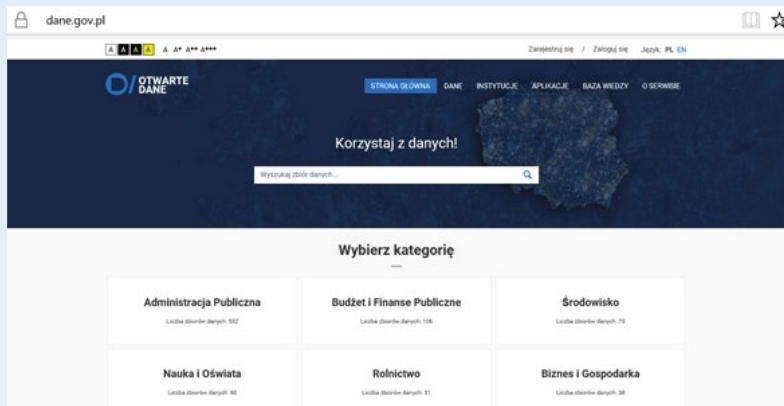
Stworzenie i utrzymanie takiego portalu wiąże się jednak ze znaczącymi kosztami utrzymania własnego zespołu rozwijającego portal lub ryzykiem uzależnienia od komercyjnego dostawcy usługi. Wyzwaniem pozostaje również modernizacja i uzupełnianie danych w przypadkach, gdy w urzędzie nie funkcjonuje międzywydziałowy zespół zajmujący się otwartymi danymi. Bez spójnej wizji i stałego zasilania danymi portal traci na atrakcyjności, a w efekcie maleje motywacja urzędników do przygotowywania danych i ich publikacji.

Przykładem imponującego portalu jest [portal z danymi miejskimi w Londynie](#). Wykorzystuje wprawdzie nieskomplikowane narzędzia [CKAN i WordPress], ale jest efektem pracy kilkunastu osób, do których obowiązków należy również zachęcanie do **korzystania z portalu** czy współpraca z organizacjami zajmującymi się wykorzystywaniem danych publicznych.

Rozwiązaniem nie jest też publikowanie danych w Biuletynach Informacji Publicznej. Nie są one zazwyczaj przystosowane do pełnienia funkcji repozytoriów danych, zwłaszcza jeśli chcemy [udostępniać dane za pośrednictwem API](#).


Rekomendacje

Zanim zaczniecie tworzyć własne repozytoria danych, rozważcie możliwość publikacji danych w Centralnym Repozytorium Informacji Publicznej [CRIP], czyli na portalu [dane.gov.pl](#).



Część instytucji działających w ramach administracji centralnej ma już obowiązek przekazywania określonych grup danych do CRIP. Ustawa o dostępie do informacji publicznej, która reguluje sposób prowadzenia CRIP, przewiduje również możliwość publikowania innych zasobów, np. przekazanych przez jednostki samorządu terytorialnego, co już ma zresztą miejsce.

Bardzo istotnym elementem towarzyszącym udostępnianiu danych na portalach otwartych danych, w tym CRIP, jest rzetelny opis informacji, które zostają opublikowane, czyli **metadane**.



Metadane to dane o danych - ustrukturyzowane informacje stosowane do opisu zasobów informacji lub obiektów informacji, dostarczające szczegółowych danych, dotyczących atrybutów zasobów lub obiektów informacji, w celu ułatwienia ich znalezienia, identyfikacji, a także zarządzania tymi zasobami.

Rozporządzenie Rady Ministrów z dnia 12 marca 2014 r. w sprawie Centralnego Repozytorium Informacji Publicznej określa, jakie informacje muszą znaleźć się w opisie zbiorów udostępnionych na tym portalu. Ten minimalny zestaw powinien również być stosowany na każdym portalu otwartych danych.

- 1. Na metadane opisujące zasób informacyjny wprowadzane przez dostawcę składają się:**
 - a) tytuł - nazwa zasobu informacyjnego, która umożliwia jego identyfikację;
 - b) słowa kluczowe - zestaw słów lub wyrażeń zwięźle opisujących zasób informacyjny;
 - c) opis - streszczenie lub krótkie określenie zawartości zasobu informacyjnego;
 - d) częstotliwość aktualizacji - informacja o częstotliwości aktualizacji zasobu informacyjnego lub metadanych;
- 2. Dostawca wprowadza także adres URL strony internetowej lub jej podstrony, w ramach której udostępniono zasób informacyjny.**

W przypadku publikowania danych na portalach otwartych danych, **zalecamy również dodanie krótkiego opisu wyjaśniającego sposób zbierania danych** tj. jaką metodologię przyjęto oraz czy pobrano pierwotne dane [karty wskaźników], a jeśli tak, to z jakiego źródła.

4.2 Wizualizacje danych

Przedstawianie otwartych danych na wykresach czy w formie atrakcyjnych wizualizacji wykracza poza zakres otwierania danych - jest już etapem ich wykorzystywania.

Wizualizacja danych może jednak znacznie ułatwiać analizę danych. Niekiedy wystarczy prezentacja danych na mapie, dzięki której bez trudu zauważymy miejsca o największym natężeniu danego zjawiska. Tak funkcjonuje np. wspomniana wyżej [Krajowa mapa zagrożeń bezpieczeństwa](#), [mapa zabytków](#) czy [mapa zagrożeń drogowych](#). Inne dane, szczególnie liczbowe, dobrze prezentują się na wykresach. Dzięki temu łatwiej jest wyobrazić sobie skalę jakiegoś zjawiska albo prześledzić zmiany w czasie. Można to robić w taki sposób, w jaki Ministerstwo Finansów prezentuje [dane o wykonaniu budżetu państwa](#), czy Gdańsk przedstawiający [dynamikę różnych zjawisk w ostatnich kilkunastu latach](#).

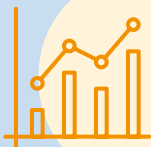
Pamiętajcie jednak, że typ wizualizacji musi być dobrany do charakteru prezentowanych danych. Jeśli pokazujemy dane o wydatkach miasta w kolejnych latach, zastosujemy wykres liniowy. Nie sprawdzi się on jednak, gdy będziemy chcieli zestawić na jednym wykresie najważniejsze kategorie wydatków w danym roku. Więcej o zasadach i narzędziach do wizualizacji danych przeczytacie w publikacji [„Mów do rzeczy”](#) Fundacji Pracownia Badań i Innowacji Społecznych „Stocznia”, a o prezentacji danych na mapach w [publikacji GUS, omawiającej standardy kartograficznej prezentacji danych statystycznych](#).

Przedstawienie danych w formie graficznej może także ułatwić dotarcie do szerszego kręgu odbiorców, zwłaszcza tych nieprzygotowanych do analizy danych w arkuszach kalkulacyjnych



czy plikach csv. Przykładem może być wizualizacja [budżetu Gdańska](#), dzięki której każdy podatnik może lepiej zrozumieć, na co miasto wydaje najwięcej i w jaki sposób sam „dokłada” się do finansów miasta.

Na rynku istnieje wiele produktów, umożliwiających zwizualizowanie w łatwy sposób dużych zasobów danych, poczynając od najbardziej popularnych arkuszy kalkulacyjnych, poprzez dedykowane komercyjne i [niekomercyjne](#) rozwiązania sprawdzające się zarówno w środowisku wewnętrznym, jak i publikacjach w Internecie.



Często dopiero na etapie tworzenia wizualizacji możemy wykryć błędy, powstałe podczas zbierania czy przygotowywania danych. Dlatego warto wizualizować dane nawet wyłącznie dla celów wewnętrznym i upewnienia się, że udostępniane dane są poprawne.



5. Jak zorganizować pracę w urzędzie



Sukces w otwieraniu danych publicznych jest możliwy wyłącznie dzięki kompleksowemu podejściu. Dlatego warto tworzyć strategię, polityki otwierania danych na poziomie każdego z urzędów, czerpiąc ze sprawdzonych wzorców z innych instytucji.

Rekomendacje

Zaprojektujcie proces otwierania danych w swoim urzędzie w taki sposób, żeby w pełni wykorzystać potencjał danej jednostki.

Silosowość polskiej administracji dotyka też obszaru danych publicznych. Brak współpracy między wydziałami, nakładanie dodatkowych obowiązków, czy niewystarczająca wiedza i utrudniona komunikacja wewnątrz urzędu powodują, że w miejsce spójnej polityki otwierania danych publicznych pojawiają się rozwiązania wyspowe i incydentalne, które nie realizują w efekcie celów otwierania danych publicznych.

Główne problemy w procesie otwierania danych publicznych:



- brak spójnej polityki otwierania danych na poziomie całego urzędu;
- sztywna, hierarchiczna struktura urzędu;
- brak wiedzy nt. regulacji prawnych, zarówno dot. obszaru legislacyjnego, jak i zapisów umownych;
- brak wiedzy i zrozumienia wśród urzędników na temat wagi, roli otwierania danych;
- brak osoby koordynującej i inspirującej proces otwierania danych;
- brak interdyscyplinarnego zespołu zajmującego się procesem otwierania danych.

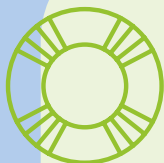
5.1. Opracowanie i wdrażanie procesu otwierania danych. Rola osoby kierującej urzędem i lidera/liderki zespołu.

Odpowiedzią na przywołane wyżej problemy jest wdrożenie wewnętrznych polityk otwartości powołujących jednocześnie horyzontalne struktury zarządzania procesem otwierania danych publicznych. Przykłady takich rozwiązań możemy odnaleźć w niektórych polskich miastach. Polityki otwartości zostały wprowadzone m.in. w [Gdańsku](#), [Słupsku](#), [Łodzi](#) i [Starachowicach](#).

5 kroków skutecznej realizacji polityki otwartości

1. Decyzja na szczeblu politycznym inicjująca proces otwierania danych.
2. Wybór lidera/liderki i określenie jego/jej zadań.
3. Powołanie zespołu oraz zobowiązanie innych pracowników do współpracy.
4. Opracowanie polityki otwartości [określenie katalogów danych, harmonogramu udostępniania, standardów, osób odpowiedzialnych za realizację].
5. Wdrażanie i ewaluacja polityki otwartości.

Niezbędnym elementem wprowadzenia polityki otwartości w urzędzie jest zdobycie poparcia wśród ścisłego grona decydentów **na czele z osobą kierującą daną jednostką**. Ta osoba musi wziąć na siebie odpowiedzialność przedstawienia rozwiązań, które w sposób kompleksowy regulują obszar otwartych danych, a jednocześnie wpływają na zmianę systemu zarządzania pracownikami.





Każdy zespół potrzebuje lidera/liderki. Liderem koordynującym pracę zespołu ds. otwierania danych, powinna być osoba pracująca w jednostce koordynującej pracę całej instytucji, zainteresowana tematyką otwierania danych, rozumiejąca potrzeby mieszkańców, organizacji pozarządowych i biznesu.

Osoba, od której zależy powodzenie procesu otwierania danych w urzędzie, wdraża politykę w urzędzie i koordynuje pracę zespołu. Jej zaangażowanie i umiejętność mobilizacji zespołu warunkuje powodzenie procesu.

Podajemy poniżej **kilka wskazówek**, które mogą w tym pomóc⁵.

1. **Pokażcie swoją wizję, z jasnymi przykładami korzyści, jakie przyniosą otwarte dane.** Przedstawcie problem i jak na jego rozwiązanie mogą wpłynąć otwarte dane - nie tylko z jednego zbioru danych, ale na skutek łączenia danych znajdujących się w różnych wydziałach lub instytucjach. Pomoże to innym osobom pracującym w urzędzie identyfikować się z procesem otwierania danych.
2. **Zbudujcie zrozumienie dla otwierania danych wśród osób kierujących departamentami i wydziałami.** To nie one będą bezpośrednio odpowiadać za proces, ale ich akceptacja jest kluczowa dla motywacji pracowników i pracowniczek. Namówcie ich na krótkie szkolenie, zaprezentujcie potencjalne korzyści dla ich jednostki i całego urzędu.
3. **Zidentyfikujcie osoby, które podzielają wizję otwierania danych publicznych.** O ile poparcie kierownictwa jest istotne dla zapewnienia stabilności procesu, to wciągnięcie „reformatorów” pozwoli na innowacyjność i ciągły rozwój.
4. **Ustalcie kanały komunikacji poświęcone zbieraniu uwag i opinii od osób,** które nie będą na co dzień zaangażowane

⁵ Opracowane na podstawie [The Open Data Institute. Open Data In Government: How To Bring About Change](#). W kontekście zarządzania procesem otwierania danych w urzędach, polecamy również, The Open Data Institute. [Lessons On Supporting Public Sector Open Data Leadership](#).



w proces otwierania danych, ale biorą udział w ich wytwarzaniu, opracowywaniu, czy korzystają z nich w zakresie swoich obowiązków. Dzięki temu zbudujecie poparcie dla polityki otwartości wśród większej grupy.

5. **Pokażcie, że możliwe są też małe sukcesy.** Ustalcie, że otworzycie na początek małe zestawy danych albo przeprowadzicie pilotaż, w jaki sposób można je wykorzystać z pożytkiem dla urzędu. Pamiętajcie jednak, że muszą wpisywać się w całościową i długoterminową politykę otwartości. Łączenie szybkich zwycięstw z długoterminowymi celami może pomóc utrzymać dynamikę zmian.
6. **Bądźcie elastyczni. Reagujcie na pomysły i potrzeby różnych działów i innych zespołów.** Wyraźcie swoją długoterminową wizję otwartych danych, ale uważajcie, aby od samego początku nie zostać zablokowanym jedynym planem wdrożenia polityki otwartych danych. Bycie elastycznym pomaga w dostosowywaniu się do zmieniających się narzędzi technologicznych, potrzeb społeczeństwa i zmian w funkcjonowaniu administracji.
7. **Upewnijcie się, że w ramach Zespołu ds. otwartych danych są osoby odpowiedzialne za zarządzanie zmianą** - nie tylko na poziomie konkretnych działań, takich jak zbudowanie portalu, publikacji zbiorów danych i animacji ich ponownego wykorzystywania. Osoby wchodzące w skład zespołu powinny szukać synergii i uwzględniać w zarządzaniu zmianą również inne czynniki, pozornie niezwiązane z otwieraniem danych.
8. **Zdobycie zewnętrzne wsparcie dla swoich działań w środowisku biznesowym, pozarządowym czy akademickim,** aby napędzać stały popyt na otwarte dane. Wsparcie zewnętrzne może pomóc w utrzymaniu politycznej woli dla procesu otwierania danych i być źródłem ciągłego uczenia się oraz dialogu.
9. **Postarajcie się, aby urzędnicy i urzędniczki mieli możliwość edukacji.** Może to być udział w szkoleniach, konferencjach, czy nawet chwilowe oddelegowanie do pracy z analitykiem/ analityczką danych lub osobą zajmującą się technologiami. Pomoże to w dalszym ulepszaniu polityki otwartych danych.



5.2. Powołanie zespołu

Wdrażane polityki otwartości powinny wprowadzać **horyzontalną strukturę zarządzania** procesem otwierania danych. Do tego niezbędne jest powołanie zespołu osób, których umiejętności i wiedza pozwolą na pełną realizację potencjału otwierania danych.

Przy powoływaniu zespołu pamiętajcie, aby zaprosić do niego **osoby odpowiadające za:**

- **rozwój strategii opartych na zbieraniu i wykorzystywaniu danych** (np. strategii rozwiązywania problemów społecznych),
- **wdrażanie narzędzi technologicznych,**
- **ochronę danych osobowych i kwestie prawne,**
- **obróbkę dużej ilości danych** (np. w zakresie rachunkowości).

W Gdańsku w skład zespołu wchodzi m.in.: Administrator Bezpieczeństwa Informacji, pracownik Wydziału Promocji i Komunikacji Społecznej, Dyrektor Biura Informatyki i koordynator z ramienia Kancelarii Prezydenta. W Starachowicach w skład zespołu powołano dodatkowo referenta prawnego oraz inspektora z Wydziału Finansowego.

Zespół powinien realizować co najmniej następujące zadania:

1. Przygotowanie, koordynowanie, monitorowanie i ewaluacja strategii otwierania danych.
2. Wypracowanie standardów otwierania danych (lub wybór gotowych standardów np. Ministerstwa Cyfryzacji) oraz ich wdrożenie w urzędzie.
3. Inwentaryzację i analizę posiadanych danych publicznych
4. Przygotowanie i testowanie środowiska informatycznego, które będzie służyło do wykorzystywania danych (np. API).
5. Opracowywanie harmonogramów udostępniania danych publicznych.



6. Współpracę z organizacjami pozarządowymi, przedstawicielami sektora biznesowego oraz innymi instytucjami, zainteresowanymi otwieraniem i wykorzystywaniem danych.

Niezwykle istotnym elementem pracy zespołu powinno być regularne (nie rzadziej niż raz do roku) **sporządzanie stanu dostępności danych publicznych za ubiegły okres**. Raport powinien zawierać również plan działań na kolejny okres mający na celu ulepszenie dostępu do danych publicznych on-line.

Kluczem do efektywnego procesu otwierania danych publicznych jest wdrożenie pracowników do planowanych zmian, biorąc pod uwagę ich doświadczenie w codziennej pracy z danymi i to, że będą mieli najprawdopodobniej więcej obowiązków związanych z realizacją celów polityki.

5.3 Opracowanie polityki otwartości

Podobnie jak i w przypadku innych dokumentów opracowywanych przez instytucje publiczne zalecamy, aby polityka otwartości została przyjęta po możliwie szerokich konsultacjach zarówno z osobami pracującymi w urzędzie, jak i mieszkańcami.

Polityka otwartych danych nie jest adresowana wyłącznie do osób, które bezpośrednio odpowiadają za jej wdrożenie. Z uwagi na jej kompleksowy charakter, adresatami polityki są również pozostałe osoby pracujące w urzędzie oraz osoby i instytucje, które mogą potencjalnie korzystać z danych publicznych. Polityka otwartości ma także charakter edukacyjny. W związku z tym powinna zawierać następujące elementy⁶:

1. **Definicja otwartych danych** i cele otwierania danych wraz z uzasadnieniem, dlaczego są ważne dla urzędu i jakie powody stoją za potrzebą wdrażania polityki.

⁶ Opracowano na podstawie: <https://theodi.org/article/how-to-write-a-good-open-data-policy/>.



2. **Ogólna deklaracja zasad**, które powinny przyświecać publikowaniu i ponownemu wykorzystaniu otwartych danych [m.in. określenie zasady publikowania danych online].
3. **Zakres kategorii danych** gromadzonych przez urząd i to, czy można je udostępnić w celu swobodnego ich ponownego wykorzystywania.
4. **Odniesienie do wszelkich stosownych aktów prawnych, polityk lub innych wytycznych**, które mają również zastosowanie w relacjach z obywatelami i innymi podmiotami zainteresowanymi wykorzystywaniem danych, w tym dotyczących kwestii ochrony prywatności.

Istotnym elementem gwarantującym harmonizację procesu otwierania danych oprócz powołania zespołu jest **zobowiązanie osób kierujących wszystkimi wydziałami danej instytucji do współpracy z zespołem i wzięcia aktywnego udziału w procesie**. Kluczowe jest ponadto zgromadzenie wiedzy na temat danych, które znajdują się w urzędzie.

Fragment zarządzenia Nr 306/14 Prezydenta Miasta Gdańska z dnia 12 marca 2014 roku.

1. 1. Zobowiązuje się dyrektorów komórek organizacyjnych do stworzenia katalogu Otwartych Danych.
2. 2. Katalog winien zawierać:
 - a) spis zbiorów danych publicznych tworzonych przez komórkę organizacyjną,
 - b) dane w ramach zbioru winny być pogrupowane na dane, które są już dostępne i dane, które planuje się udostępnić.
3. Spisy zbiorów i zawartych w nich danych winny być przez komórki organizacyjne tworzone pod nadzorem i we współpracy z Zespołem.
4. Zobowiązuje się dyrektorów komórek organizacyjnych do stałej współpracy z Zespołem w realizacji i rozwoju strategii polityki otwartości poprzez:

- a) delegowanie pracowników do współpracy z Zespołem,
- b) dokonywanie analiz posiadanych danych publicznych,
- c) przygotowywanie odpowiednich danych do publikacji,
- d) stopniowe publikowanie nowych zbiorów danych w terminach uzgodnionych z Zespołem.

Pamiętajcie, by **sposób prezentacji informacji odnoszących się do katalogów danych był czytelny**. Warto podzielić je na kategorie (np. zdrowie, finanse, demografia), oznaczyć słowami kluczowymi, jak również wskazać, które z nich mogą być pobierane przez API.

Dobra polityka otwartości powinna również zawierać **harmonogram jej wdrażania** i wskazywać wyraźnie **obowiązki sprawozdawcze** z jej wykonania.

Przy tworzeniu harmonogramu otwierania danych warto wziąć pod uwagę następujące zasady⁷:

1. **Określcie priorytety.** Zaczniście od otwierania danych mających dalekosiężne znaczenie dla obywateli i instytucji: dane potrzebne do opracowania aktu prawnego, strategii, polityk czy ułatwiających wdrożenie nowych regulacji w urzędzie. Następnie otwórzcie te dane, które cieszą się dużym zainteresowaniem obywateli, organizacji i biznesu. Nie ograniczajcie się wyłącznie do tego, co chcą osoby wchodzące w interakcję z urzędem. We współpracy z pozostałymi osobami pracującymi w urzędzie ustalcie, jakie dane pozwolą na poprawę funkcjonowania instytucji, czy poprawę jakości życia mieszkańców.
2. **Zwróćcie uwagę na koszty.** Czasami koszt otwierania danych jest wysoki z uwagi na ich skomplikowaną strukturę. Pamiętajcie o tym, tworząc harmonogram. Być może trzeba będzie pozyskać dodatkowe fundusze, co oznacza, że warto

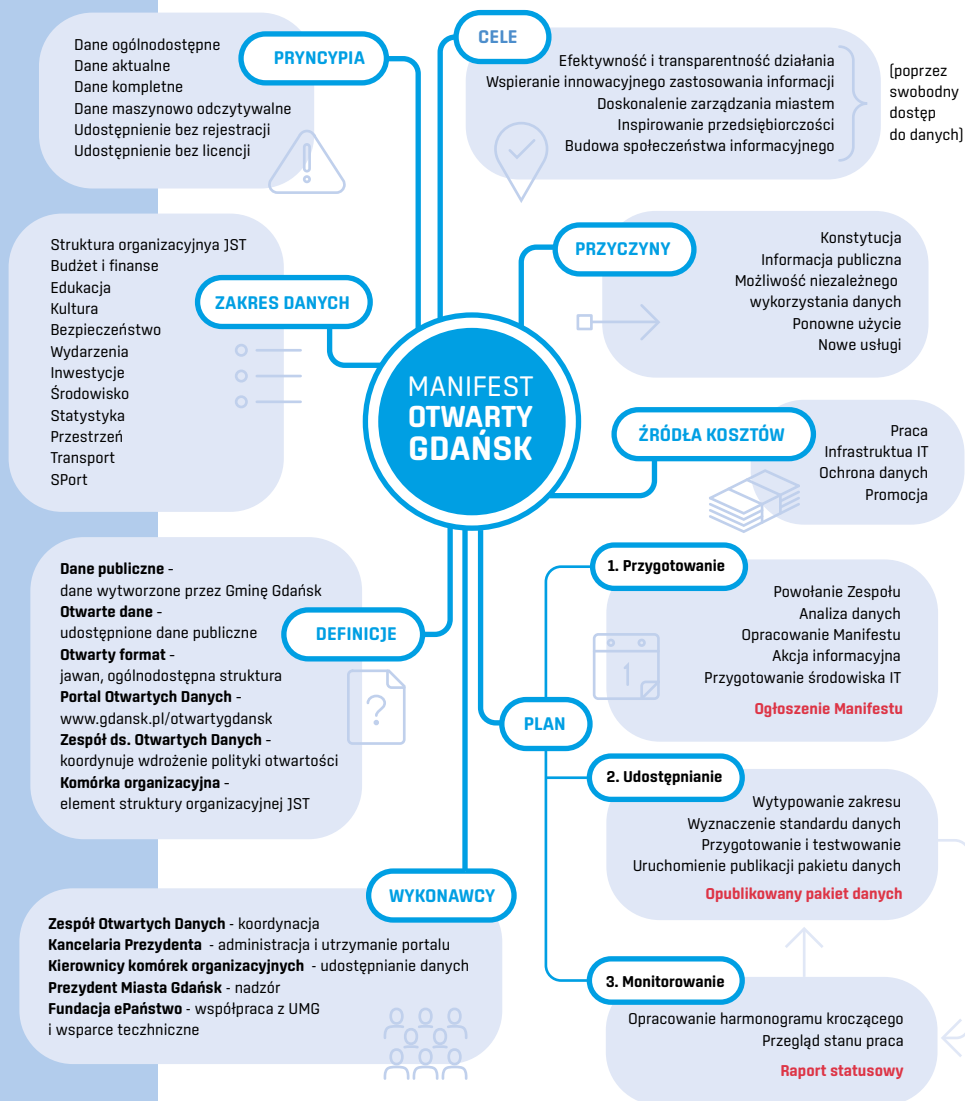
⁷ Opracowano na podstawie <https://sunlightfoundation.com/opendataguidelines/>.

odłożyć otwieranie tych danych w czasie, by później osiągnąć lepszy efekt. Dlatego warto zacząć od tych danych, które już są publikowane i ich dostosowanie do standardów otwartości nie jest skomplikowane.

Polityka musi zawierać precyzyjne określenie, **kto i w jaki sposób będzie dokonywał przeglądu udostępnionych danych**. Taki obowiązek powinien należeć do zespołu, który we współpracy z osobami kierującymi wydziałami i innymi jednostkami ustali szczegółowe terminy sprawdzenia poprawności [aktualności i zgodności ze standardami] danych, biorąc również pod uwagę częstotliwość ich pobierania, sposoby wykorzystywania, zmieniające się otoczenie technologiczne, czy informacje zwrotne od użytkowników/użytkowniczek. **Rolą lidera/liderki jest bieżące monitorowanie wdrażania polityki otwartości i przestrzegania realizacji harmonogramu**. Dobrą praktyką jest, aby polityka zawierała również **określenie standardów danych**. Zarówno odnoszących się do etapu ich tworzenia, dalszego opracowywania jak i publikowania. Warto dążyć do tego, aby **standardy były również spójne w ramach wielu urzędów, jeśli dotyczą tych samych danych**. W zagwarantowaniu spójności danych pomogą rekomendacje zawarte w [standardzie technicznym opracowanym przez Ministerstwo Cyfryzacji](#).

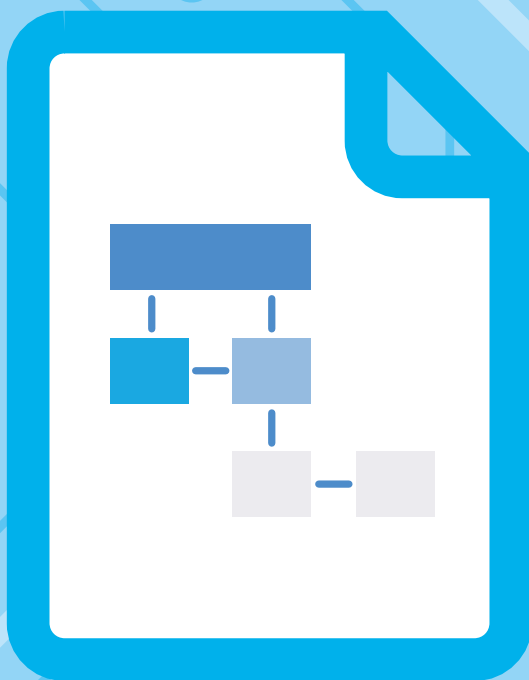


Schemat polityki otwartości na podstawie przykładu miasta Gdańska⁸



⁸ Załącznik Nr 2 do Zarządzenia Nr 306/14 Prezydenta Miasta Gdańsk z dnia 12 marca 2014 r. Dostępne: http://otwartygdansk.pl/wp-content/uploads/2016/07/PMG_2014_6_306_zal013.pdf.

6. Schemat wdrażania polityki otwartości. Otwieranie zbiorów danych



We wcześniejszej części podręcznika przedstawiliśmy informacje dotyczące tworzenia otoczenia instytucjonalnego i strategicznego dla procesu otwierania danych. Poniżej prezentujemy schemat otwierania konkretnych zbiorów danych.

1. Zidentyfikujcie zasoby danych

Naprawdę nie ma potrzeby udostępniania wszystkich publicznych danych na raz. Zaczynajcie od tego, co jest najłatwiejsze do udostępnienia, czyli od danych, które spełniają pewne standardy lub łatwo je będzie dostosować (patrz **rozdział 2. Co to znaczy, że dane są otwarte**). W tym momencie warto zorganizować szkolenie dla osób pracujących w urzędzie, aby uwspólnić pojęcia, zdiagnozować wewnętrzne potrzeby samych urzędników i urzędniczek oraz upewnić się, jakie dane mogą być udostępniane. Skorzystajcie również z Bazy wiedzy dostępnej na portalu dane.gov.pl.

Proponujemy taktyki działania, które można traktować jako następujące po sobie elementy procesu lub jak oddzielne działania:

1.1 Zinventaryzujcie bazy danych we wszystkich jednostkach

Są różne style prowadzenia baz danych, związane z codzienną pracą urzędu. Trzeba zrobić wśród nich kwerendę, aby dowiedzieć się, jakie informacje są zgromadzone. Często same wydziały czy departamenty nie wiedzą, jakie bazy danych posiadają inne jednostki. Dobrym przykładem jest [inwentaryzacja przeprowadzona w związku z realizacją Programu](#). Dowiedcie się z niej także, jakie dane są już otwarte.

1.2 Przeanalizujcie wnioski i zapytania o dostęp do informacji publicznej i ponowne wykorzystywanie informacji sektora publicznego

Do waszego urzędu na pewno wpływają wnioski obywateli o udostępnienie informacji. Szczególnie tryb dostępu do informacji

publicznej stał się dość popularny i pozwala obywatelom otrzymać informacje publiczne, które nie są dostępne w przyjaznych dla nich formatach. Powtarzające się pytania obywateli sygnalizują, jakie informacje warto udostępnić w bardziej przyjaznej formie. Można także przeanalizować inne kanały, którymi obywatele zgłaszają prośby o dane, informacje, pliki czy usługi. Na podstawie powyższej analizy, wybierzcie kilka najważniejszych baz, do otwarcia w pierwszej kolejności.

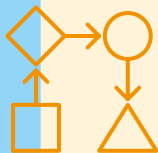
1.3 Sprawdźcie uwarunkowania prawne

Sprawdźcie przepisy prawa, regulujące rejestry z danymi, które zidentyfikowaliście. Jeśli nabywacie lub udostępniacie utwór pamiętajcie o odpowiednim zabezpieczeniu praw autorskich. Dobrą praktyką jest jego udostępnianie na wolnej licencji. Więcej informacji o zagadnieniach prawnych znajdziecie w standardzie prawnym.

2. Zapytajcie otoczenie o potrzeby

2.1 Zidentyfikujcie ważnych interesariuszy

Można przeznaczyć spore nakłady finansowe i czasowe na udostępnienie pewnych baz danych, na które brak jest społecznego zapotrzebowania. Może to powodować opór i wątpliwości wewnątrz urzędu przed udostępnianiem innych danych. Dlatego warto stworzyć sobie okazję do poznania potrzeb otoczenia urzędu i poszukać sprzymierzeńców projektu. Kto najczęściej prosi o dane czy informacje? Zaczniacie od partnerów urzędu, którzy najczęściej wnioskuje o informacje lub z którymi wasz urząd już współpracuje np. w formie zlecania zadań. Także takie grupy jak biznes, organizacje społeczne czy naukowcy posługują się w swojej pracy danymi, dlatego doskonale wiedzą, jakich danych potrzebują i jakich im brakuje. Warto być obecnym na spotkaniach branżowych, aby obserwować, w jakich sprawach jest aktualnie zapotrzebowanie na informacje, analizy czy zbiory danych.



Co może przynieść współpraca interdyscyplinarna i wymiana danych między jednostkami administracji publicznej?

Z inspiracji prof. Jonathana Shepherd, który zauważył, że $\frac{3}{4}$ przyjmowanych w szpitalach ofiar przemocy nie zgłasza zdarzenia na policję, miasto Cardiff w Walii stworzyło system wymiany zanonimizowanych informacji między szpitalnymi izbami przyjęć a władzą lokalną i policją. W szpitalu ofiarom takich zdarzeń zadawano 4 krótkie pytania, a odpowiedziami, po zanonimizowaniu, dzielono się z policją. Służby wraz z przedstawicielami władz analizowały zbiorczo przypadki i wprowadzały zmiany w miejscach, gdzie najczęściej dochodziło do przestępstw. Doprowadziło to do oszczędności rządu 6,6 mln \$ i zmniejszyło poziom przestępstw z udziałem przemocy o 42%. Model obowiązuje w różnych formach w całej Wielkiej Brytanii, przenoszony jest także do Stanów Zjednoczonych, Australii, Holandii [wszystkie szpitale stosują ten model] i RPA.

Źródło: [Wales cuts reported violence by 42% in its capital with data sharing model.](#), apolitical.co; [Cutting crime: surgeon's research that helps stop street violence](#), theguardian.com.

2.2 Przeprowadźcie wywiady indywidualne z najważniejszymi interesariuszami

Przeprowadzenie indywidualnych rozmów z najważniejszymi podmiotami, które są zainteresowane danymi, pozwoli uzyskać szczegółowe informacje na temat ich potrzeb w zakresie otwartych danych oraz potencjalnych skutków i korzyści związanych z zaspokojeniem tych potrzeb. Jeśli przeprowadzicie wywiady tylko ze specyficzną grupą, dostaniecie niepełny obraz. Warto zdywersyfikować badaną grupę, aby w danym temacie poznać różnorodne potrzeby. Wykorzystajcie tę wiedzę do wygenerowania jednego lub więcej przypadków użycia otwartych danych.



Jak to zrobiono w Nowym Jorku?

Określono, że podczas wywiadów indywidualnych zostanie przebadana taka sama liczba respondentów korzystających z danych publicznych urzędu z częstotliwością:

- **niską:** nigdy nie korzystali i tylko słyszeli o danych publicznych;
- **średnią:** przeglądali wasze dane i/lub prosili o nie, ściągnęli od 1 do 5 zbiorów;
- **wysoką:** posiadają obszerną wiedzę z zakresu waszych zbiorów danych i często z nich korzystają;

Źródło: [„Understanding the Users of Open Data: Research Findings”](#)

Dla przykładu w Nowym Jorku w wyniku wywiadów powstały poniższe typy użytkowników/użytkowniczek Otwartych Danych.

Poznaj użytkowników danych z Nowego Jorku

Obecni użytkownicy danych

Potencjalni użytkownicy danych



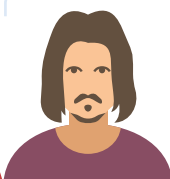
Stefan
skrupulatny
planista



Roman
lokalny lider



Barbara
wpływową
przedsiębiorczyni



Marek
kompetentny
odkrywca



Zofia
zabiegana
obserwatorka



Maria
aktywistka
społeczna

Źródło: [„Understanding the Users of Open Data: Research Findings”](#)

2.3 Wybierzcie szczegółowy temat

Na podstawie poprzednich kroków wybierzcie jeden przypadek potrzeby użycia danych ze specyficznej sfery np. na podstawie zidentyfikowanych potrzeb jednego z ważnych partnerów urzędu. Weźcie pod uwagę własne zasoby i możliwości oraz potrzeby otoczenia. **Na początek warto wybrać coś prostszego, aby zniwelować ryzyka i przećwiczyć sam cykl otwierania zbiorów.** Kluczowe jest skupienie się i ograniczenie się do jednego przypadku, jednej potrzeby na raz. Warto utrzymać bliski kontakt z daną organizacją czy podmiotem, którego potrzeby wybrało się do realizacji. Ich uwagi i informacje zwrotne są niezbędne do otwarcia, odpowiadającego na potrzeby, zbioru danych.

Współpraca z biznesem: Ranking sanitarny restauracji w San Francisco

Restaurants Nightlife Home Services Write a Review Events Talk Collections Log In

Zero Zero > Health Scores

Zero Zero

December 7, 2017 — Routine Inspection

Violations

- Inadequate and inaccessible handwashing facilities [date violation corrected: 12/7/2017]
- Unclean or degraded floors walls or ceilings

Inspections

Date	Inspection Type	Violations	Score
December 7, 2017	Routine	2	94
September 1, 2016	Routine	3	90

Health Score

94

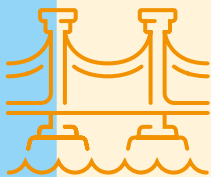
out of 100

About Health Scores

We collect public inspection data directly from your local health department. Due to the local health department's inspection schedule as well as the time it takes to pass that information on to us, it is possible that we may not display the most recent inspection data.

Please report any health complaints about this business such as potential food borne illnesses or any unreasonable delay and data inaccuracies via one of the methods below:

- [Email](#)
- [Website](#)



Wydział Zdrowia Publicznego miasta San Francisco udostępnił dane z kontroli sanitarnych w restauracjach. Portal yelp.com wykorzystał te dane do stworzenia dodatkowego kryterium oceny restauracji, aby mieszkańcy mogli poznać poziom higieny w danej restauracji. Dodatkowo, kontrolerzy korzystają z nowego sposobu wyboru miejsc do kontroli sanitarnej. Otrzymują specjalne raporty z portalu, które zwracają uwagę na restauracje, których klienci narzekają na zatrucia lub niski poziom higieny. Badanie Harvard Business Review wykazało, że publikacja ocen z kontroli sanitarnej sprawia, że nisko oceniane restauracje mają tendencję do zwiększania poziomu higieny i uzyskują lepsze wyniki w kolejnych kontrolach.

Więcej: [Cities take to Yelp to tell their residents where not to eat](http://apolitical.co), apolitical.co.

3. Sprawdzajcie i standaryzujcie dane

Zobaczcie, jakie dane macie w bazie, którą chcecie otworzyć. Sprawdźcie czy nie ma w niej błędów, czy wszystkie pola mają prawidłowe wartości, czy nie ma „dziur” w zestawieniach, literówek i innych błędów obniżających użyteczność całej bazy. Ustalcie jeden standard dla swoich danych [jeden sposób zapisu daty, adresu, lokalizacji, wartości]. By zapewnić spójność standardów, korzystajcie ze standardów [Ministerstwa Cyfryzacji](#), również w zakresie [API](#). Warto także wykorzystać informację zwrotną od potencjalnych użytkowników/użytkowniczek zbioru, który otwieracie lub skonsultować użyteczność bazy z ekspertami i ekspertkami. Upewnijcie się, że dostarczacie dane w formacie odpowiednim dla partnerów, którzy przedstawili swoje potrzeby.

4. Podajcie kontekst - metadane

Nie publikujcie samej bazy danych. Napiszcie, co w niej się znajduje, dodajcie zrozumiałe słowa kluczowe (raczej: „samochód” niż „pojazd”), dodajcie warunki wykorzystania tych danych, zakres danych, oraz ich aktualność (data utworzenia, data ostatniej modyfikacji, częstotliwość aktualizacji). Takie informacje pomogą użytkownikom/użytkowniczkom zorientować się, co dokładnie przedstawia zbiór. Skorzystajcie z [wytycznych Ministerstwa Cyfryzacji](#), również w zakresie [API](#).

Warto pokazać także efekty diagnozy potrzeb odnośnie danej bazy danych, zebranej z otoczenia urzędu. Pozwoli to pokazać kontekst użycia danych i może zwiększyć wykorzystanie bazy przez mniej zaawansowanych użytkowników/użytkowniczki. Zaprosicie użytkowników/użytkowniczki danych do dzielenia się rzeczywistymi historiami użycia danych w jakimś celu.

5. Wybierzcie otwarty format publikacji

Na przykład CSV albo XML - pamiętajcie, że dane muszą być maszynowo odczytywalne, więc skany czy tworzone „ręcznie” tabelki w dokumencie tekstowym nie są najlepszym rozwiązaniem. Skorzystajcie z [wytycznych Ministerstwa Cyfryzacji](#), również w zakresie [API](#).

6. Opublikujcie dane

Na fali entuzjazmu, niektóre urzędy miejskie w Polsce tworzyły własne, dedykowane portale otwartych danych. Jednakże, doświadczenia niektórych samorządów pokazują, że utrzymanie własnych serwisów jest czasochłonne i kosztowne. **Wyjściem z sytuacji dla wszystkich jednostek publicznych może być umieszczanie danych na dane.gov.pl.**

Możecie także umieścić zbiory danych na swoim BIPie, o ile posiada on odpowiednie parametry techniczne i umożliwia łatwe wyszukanie danych. Ważne jest, aby nie ograniczać dostępu do strony i aby była ona czytelna dla użytkowników/użytkowniczek.

7. Upewnijcie się, że dane są dostępne

Aby zapewnić aktualność danych dla obiorców korzystających z maszynowego dostępu do danych, najlepiej dołączyć do udostępnionych danych [API \[Application Programming Interfaces\]](#), które spowoduje, że kiedy tylko zmienicie swoje dane, od razu nastąpi ich aktualizacja u tych, którzy z nich korzystają. Jest to bardzo istotne przy danych często zmieniających się w czasie lub po prostu udostępnianych w czasie rzeczywistym [aktualne zanieczyszczenie powietrza, informacje o stanie wód, komunikaty o utrudnieniach w ruchu itp.]. Skorzystajcie ze standardów opracowanych przez Ministerstwo Cyfryzacji w zakresie [API](#).

8. Poinformujcie świat

Dane mają znaczenie, jeśli są używane. Wasze otoczenie, to bliższe jak i dalsze, powinno dowiedzieć się o nowych zbiorach danych. Jeśli na podstawie wcześniejszych kroków nawiązaliście relacje z najbardziej zainteresowanymi podmiotami, to właśnie oni są „tubą”, która powie światu, że mamy nowe, ciekawe zbiory danych.

Warto skontaktować się z blogerami technologicznymi, uczelniami, kołami naukowymi i organizacjami pozarządowymi. Poinformujcie, że macie dane i że mogą ich używać. Takie szerokie podejście, może zapewnić dodatkową informację zwrotną. Wsłuchanie się w te głosy może skutkować propozycjami ciekawych rozwiązań, które zwiększą użyteczność danych.

Jeśli chcecie dodatkowo wypromować otwarte dane lub zidentyfikowaliście ciekawe potrzeby z otoczenia urzędu, do których

zaspokojenia nie wystarczy samo otworzenie zbioru danych, możecie zorganizować wydarzenie typu hackathon / datathon.

Jest to rodzaj maratonu programistycznego lub analitycznego, odbywającego się najczęściej w weekend. Podczas takiego wydarzenia, pasjonaci i pasjonatki o różnych specjalizacjach, starają się wypracować innowacyjne sposoby rozwiązania jakiegoś problemu. Dzięki temu mogą dostarczyć wiele ciekawych rozwiązań bazujących na danych publicznych z różnych źródeł oraz na oprogramowaniu [open source](#). Takie oprogramowanie daje możliwość bezpłatnego dostosowania narzędzia technologicznego wytworzonego już przez kogoś innego (np. <https://www.opentreeemap.org/>).

Tego typu wydarzenia warto organizować w partnerstwie branżowym z innymi instytucjami publicznymi i biznesem czy organizacjami pozarządowymi. Jeśli nagrodą będzie wdrożenie narzędzia w waszym urzędzie lub wsparcie w niezależnym wdrożeniu, poczucie wpływu na rzeczywistość zwycięskiego zespołu może być wystarczającą motywacją.

Ważnym elementem hackathonu/datathonu jest zdefiniowanie zadań/wyzwań, które nie mogą być zbyt ogólne. W tym miejscu bardzo przydadzą się szczegółowe potrzeby, które zostały zebrane z otoczenia na początkowym etapie procesu otwierania waszych baz.



Muzeum Narodowe w Warszawie zorganizowało w maju 2018 roku: HackArt/ Hackathon MNW, aby poszukać nowych sposobów na wykorzystanie ich zbiorów (traktowanych jako zbiory w domenie

publicznej]]. Główną nagrodą było m.in. wdrożenie najlepszych projektów z pomocą firm partnerskich. Aby uczestnicy nie tworzyli rozwiązań nieistniejących problemów, organizatorzy przedstawili 5 wyzwań, które są związane z codzienną pracą muzeum np.: Jak pomóc zainteresowanym eksplorować kolekcje zarówno te wirtualne, jak i znajdujące się na terenie budynku Muzeum?

W czerwcu 2018 dwa zespoły dołączyły do społeczności [Koduj dla Polski](#) i rozwijają projekt przy wsparciu pracownika Muzeum. Razem zmierzają do wdrożenia np. chatbota w stylu Stańczyka, który odpowiada na proste pytania dotyczące pracy muzeum i odciąża pracowników.

Więcej: <https://hackathon.mnw.art.pl/>

9. Aktualizujcie dane

Dane otwarte wymagają aktualizacji, która umożliwi ich ponowne wykorzystywanie. Ustalcie częstotliwość aktualizacji, poinformujcie dokładnie o tym i trzymajcie się planu.

10. Wróćcie do punktu pierwszego

Poszczególne kroki warto przejść dla każdego otwieranego zbioru danych. Uchroni to przed otwieraniem zbiorów danych o niskiej użyteczności. Zarówno jeśli chodzi o techniczne aspekty, jak i potrzeby społecznego otoczenia waszego urzędu.



Podręcznik powstał we współpracy pracowników Departamentu Otwartych Danych i Rozwoju Kompetencji Ministerstwa Cyfryzacji w składzie: Agata Miazga, Aneta Cesarek, Anna Gos, Anna Bramska, Magdalena Dąbrowska, Karolina Kościerzyńska, Sylwia Pichlak-Pawlak, Dominik Sybilski, Iwona Szelenbaum, Jacek Żuchowicz, z Fundacją ePaństwo po warsztacie, który odbył się 11 czerwca 2018 r. w Gdańskim Centrum Informatycznym, zorganizowanym przez Ministerstwo Cyfryzacji, Miasto Gdańsk i Fundację ePaństwo, z udziałem przedstawicieli Miasta Gdyni. Warsztat służył wymianie doświadczeń pomiędzy administracją samorządową i rządową w zakresie otwierania danych publicznych. Poprzedzony był wywiadami indywidualnymi z praktykami otwierania danych w Gdańsku. Treść podręcznika była ponadto konsultowana z Centrum Cyfrowym.

Podręcznik został opracowany w ramach projektu: **„Otwarte dane - dostęp standard, edukacja”**.

Jeśli skorzystaliście z porad zawartych w podręczniku, potrzebujecie dodatkowych informacji lub jeśli chcecie podzielić swoją opinią na jego temat napiszcie do nas na adres: otwartedane@mc.gov.pl.

