

# Analizy Programu Odpowiedzialne Państwo

Warszawa, wrzesień 2014

analiza

Fundacji Batorego

policy brief

Batory Foundation

Łukasz Jachowicz

## Czekając na otwarte rządy: otwartość danych publicznych

### Rekomendacje:

- Jednostki publiczne są w posiadaniu danych, których udostępnienie zwiększy przejrzystość sfery publicznej, doprowadzi do stworzenia nowoczesnych rozwiązań bazujących na danych oraz do zwiększenia wartości całej gospodarki (dodatkowe 40 miliardów euro rocznie w całej UE).
- W Polsce brakuje całościowego, spójnego podejścia do otwierania danych publicznych.
- Konieczne są zmiany prawne, których celem jest wprowadzenie otwartości danych publicznych jako reguły, a nie jako wyjątku.
- Do sprawnego otwierania danych w celu ich ponownego wykorzystania konieczna jest edukacja pracowników instytucji publicznych zarówno w zakresie prawa otwartości, jak i możliwych sposobów późniejszego wykorzystania danych.
- Konieczna jest zmiana praktyki przy projektowaniu publicznych systemów informatycznych, by umożliwiały one wykorzystanie upublicznianych danych z pominięciem interfejsu WWW dostarczonego przez administrację publiczną.

### Wstęp

Od kilku lat jednym z głównych tematów związanych ze stykiem sfery publicznej i prywatnej oraz nowoczesnych technologii jest otwieranie dostępu do danych publicznych. Rozwój techniki, w tym popularyzacja urządzeń mobilnych oraz systemów mogących przetwarzać ogromne ilości informacji, dały obywatelom możliwość wykorzystywania danych, w sposób, który wcześniej dostępny był jedynie wielkim instytucjom. Otwieranie danych wpływa pozytywnie zarówno na przejrzystość działań państwa, jak i na gospodarkę, pozwalając



FUNDACJA  
BATOREGO

IM. STEFANA

kontakt: [www.batory.org.pl](http://www.batory.org.pl)  
[www.otwartyrzad.org.pl](http://www.otwartyrzad.org.pl)

obywatelom na analizowanie działań administracji, a przedsiębiorcom – na tworzenie nowych usług.

Wagę otwartości danych publicznych oraz interoperacyjności systemów informacji publicznej podkreśla coraz więcej państw. Partnerstwo na rzecz Otwartych Rządów – międzynarodowa inicjatywa promująca zasady przejrzystości, otwartości i rozliczalności<sup>1</sup> rządu skupiająca obecnie 64 kraje, w swym dokumencie programowym<sup>2</sup> wskazuje na znaczenie zastosowania nowych technologii dla zwiększenia dostępności danych publicznych oraz wymiany dobrych praktyk w zakresie otwierania danych. Unia Europejska przyjęła *Dyrektywę w sprawie ponownego wykorzystywania informacji sektora publicznego*<sup>3</sup>.

W Polsce dyskusja o udostępnianiu danych publicznych jest wciąż zaledwie elementem szerszej debaty o dostępie do informacji publicznej i o otwartych zasobach (głównie dziełach kultury finansowanych z publicznych środków). Z tego powodu rozmowy o danych skupiają się na tym, co udostępniać, z pominięciem namysłu nad tym, w jaki sposób to robić. Efektem tego są utrudnienia w ponownym wykorzystaniu upublicznionych już informacji oraz brak całościowego i usystematyzowanego podejścia do otwartości.

Pierwszym zwiastunem zmiany wydaje się być nowopowstałe Centralne Repozytorium Informacji Publicznej, które – po usunięciu błędów utrudniających ponowne wykorzystywanie zawartych w nim danych – ma szansę stać się centralnym punktem dostępu do danych generowanych przez instytucje publiczne. Niniejsza analiza ma na celu zasygnalizowanie głównych problemów z otwartością danych w Polsce oraz zaproponowanie kierunków zmian, do których powinna dążyć teoria i praktyka otwierania danych.

## Dlaczego wspierać otwartość danych publicznych?

Zapewnienie dostępu do danych generowanych i znajdujących się w zasobach instytucji publicznych jest jednym z podstawowych sposobów zapewnienia przejrzystości działań administracji publicznej. Pierwotną motywacją do otwierania danych było zwiększenie kontroli nad jednostkami finansowanymi z funduszy publicznych, czyli podatków obywateli.

W praktyce prawidłowe otwieranie danych, umożliwiające wyszukanie i ponowne wykorzystanie informacji, usprawnia też działanie samych instytucji publicznych, ułatwiając wewnętrzny obieg informacji oraz likwidując konieczność generowania podobnych zbiorów przez inne jednostki.

Otwarcie dostępu do danych będących w posiadaniu instytucji publicznych umożliwia dodatkowo rozwijanie innowacyjnych przedsięwzięć związanych z nowoczesnymi technologiami. Możliwość wykorzystania szerokich zasobów informacyjnych pozwala rodzimym przedsiębiorstwom na konkurowanie z największymi producentami rozwiązań przetwarzających dane, dając im na starcie te same szanse. Jak wskazuje Komisja Europejska, otwarcie dostępu do danych będących w posiadaniu instytucji publicznych doprowadzi do wzrostu wartości unijnej gospodarki o 40 miliardów euro rocznie.

Wśród najciekawszych danych, które warto w pierwszej kolejności upublicznić w sposób umożliwiający ich ponowne przetwarzanie, można wyszczególnić:

- informacje o lokalach komunalnych do wynajęcia
- rejestry wydatków instytucji publicznych
- informacje o zamówieniach publicznych
- rejestry typu CEPIK, KRS, ewidencja gospodarcza, dane o spółkach akcyjnych
- dane pogodowe
- dane geodezyjne
- dane ze stacji badających jakość powietrza
- projekty aktów prawnych: prawo centralne i miejscowe

<sup>1</sup> Ang. *accountability*.

<sup>2</sup> Open Government Declaration, <http://www.opengovpartnership.org/about/open-government-declaration>, nieoficjalnie tłumaczenie dostępne na stronie [http://informacjapubliczna.org.pl/wwwdane/files/deklaracja\\_na\\_rzecz\\_otwartych\\_rzadow\\_mkwu.pdf](http://informacjapubliczna.org.pl/wwwdane/files/deklaracja_na_rzecz_otwartych_rzadow_mkwu.pdf) [dostęp: 01.09.2014 r.].

<sup>3</sup> Dyrektywa 2003/98/WE Parlamentu Europejskiego i Rady z dnia 17 listopada 2003 r. w sprawie ponownego wykorzystywania informacji sektora publicznego.

- dane dotyczące przestępczości w poszczególnych obszarach
- dane statystyczne w formie źródłowej (zanonimizowane)
- informacje o wynikach kontroli jakości w restauracjach, sklepach, na stacjach benzynowych
- informacje o prowadzonych i planowanych remontach/objazdach
- aktualizowane na bieżąco dane na temat opóźnień komunikacji publicznej lub lokalizacji poszczególnych pojazdów
- informacje o stopniu zajętości parkingów publicznych w poszczególnych obszarach miasta (na podstawie danych z parkomatów)
- informacje o lokalizacji wypożyczalni rowerowych, miejskich szaletów, posterunków policji, biletomatów, parkomatów, przystanków, miejsc historycznych, żłobków, przedszkoli, szkół, szpitali, punktów opieki, hoteli, bibliotek, hotspotów itp.
- rozkłady jazdy komunikacji publicznej

### Otwartość danych publicznych w zapisach prawa i w praktyce

Obecnie nie ma w Polsce aktu prawnego, który całościowo regulowałby wszystkie kwestie związane z otwartością danych. Zapisy, które częściowo regulują upublicznianie danych oraz ich wymianę między nowo budowanymi systemami informatycznymi, są dostępne w kilku aktach prawnych.

Historycznie pierwszym aktem jest ustawa o informatyzacji podmiotów realizujących zadania publiczne (z 2005 roku), nakazująca m.in. udostępnianie zestawień stosowanych struktur dokumentów, formatów danych oraz protokołów komunikacyjnych oraz szyfrujących – co jest konieczne dla sprawnego przetwarzania danych publikowanych w systemach publicznych. Ustawę doprecyzowuje rozporządzenie zakazujące dyskryminowania jakiegokolwiek rozwiązania technologicznego przy budowaniu systemów wymieniających informacje z innymi<sup>4</sup>. Rozporządzenie nakłada

obowiązek udostępniania publikowanych zasobów informacyjnych w co najmniej jednym formacie określonym w załączniku, wymieniającym głównie otwarte, dobrze zdefiniowane formaty danych.

W 2014 r. weszły w życie dwa rozporządzenia związane bezpośrednio z otwartością danych publicznych: pierwsze, tworzące Centralne Repozytorium Informacji Publicznej (CRIP)<sup>5</sup> i drugie, wskazujące zbiory danych, jakie znajdują się w Repozytorium<sup>6</sup>. Rozporządzenia realizują zapisy znowelizowanej Ustawy o dostępie do informacji publicznej, nakazującej publikowanie w CRIP „informacji publicznych o szczególnym znaczeniu dla rozwoju innowacyjności w państwie i rozwoju społeczeństwa informacyjnego, które ze względu na sposób przechowywania i udostępniania pozwalają na ich ponowne wykorzystywanie w sposób użyteczny i efektywny”. Jest to pierwszy w polskim prawie zapis ustawowy kładący nacisk na taki sposób publikowania danych, by można je było z łatwością ponownie wykorzystywać.

Poza przytoczonymi ustawami, zapisy mogące ułatwić dostęp do konkretnych danych można znaleźć w innych aktach prawnych, dotyczą one jednak wyłącznie fragmentarycznych zasobów i nie stanowią elementu spójnej polityki państwa w tym zakresie.

Z brakiem spójnego podejścia prawnego łączy się brak spójnej praktyki w otwieraniu i udostępnianiu danych publicznych. Aby dane można było uznać za otwarte i bezproblemowo wykorzystywać je w celu analizowania, łączenia z innymi zbiorami oraz tworzenia na ich podstawie nowych rozwiązań informatycznych, muszą spełniać kilka warunków. Muszą być kompletne; upubliczniane w wersji źródłowej; aktualne; publikowane w formatach umożliwiających ich ponowne użycie, niezależnie od wykorzystywanego oprogramowania; przetwarzalne maszynowo; udostępnione niedyskryminująco, czyli dostępne bez koniecz-

<sup>4</sup> 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.

<sup>5</sup> Rozporządzenie Rady Ministrów z dnia 12 marca 2014 r. w sprawie Centralnego Repozytorium Informacji Publicznej.

<sup>6</sup> Rozporządzenie Ministra Administracji i Cyfryzacji z dnia 26 marca 2014 r. w sprawie zasobu informacyjnego przeznaczonego do udostępniania w Centralnym Repozytorium Informacji Publicznej.

ności rejestracji czy podpisywania specjalnych umów; dostępne bez ograniczeń licencyjnych; opublikowane w sposób nieutrudniający ich ponownego wykorzystania (na przykład przez rozrzucenie danych z tego samego zbioru na wielu stronach internetowych)<sup>7</sup>. Do tej pory zaledwie kilka miast (np. Gdańsk, Poznań, Wrocław) zdecydowało się na udostępnienie w sieci niewielkiej części gromadzonych przez siebie informacji: dokładnych informacji o transporcie miejskim, danych geodezyjnych czy informacji o położeniu parkomatów.

Teoretycznie obywatele mają prawo dostępu do nieopublikowanych w sieci danych na podstawie ustawy o dostępie do informacji publicznej – mogą złożyć indywidualne wnioski, w których można wskazać formę, w jakiej instytucja publiczna powinna udostępnić poszukiwane informacje. W praktyce istnieją jednak pewne przeszkody:

- brak katalogu danych znajdujących się w posiadaniu danej jednostki, przez co utrudnione jest dokładne wskazanie rodzaju informacji, jakich potrzebuje użytkownik;
- niedostosowanie ustawy o dostępie do informacji publicznej w części dotyczącej ponownego wykorzystywania do wymogów nowoczesnych systemów IT – brak w niej zdefiniowanego trybu żądania „w trybie ciągłym” (czyli stałego dostępu do najnowszej wersji zbioru danych);
- niechęć do udostępniania przez instytucje publiczne danych, które mogą być wykorzystywane również w celach komercyjnych;
- ograniczenia prawne lub niechęć do udostępniania danych, które są wykorzystywane przez daną instytucję w celach komercyjnych (sprzedawane) lub niekomercyjnych (wykorzystywane do budowy własnego systemu oferującego, przykładowo, rozkłady jazdy);
- domniemane naruszenie ustawy o ochronie danych osobowych<sup>8</sup>;
- domniemane lub rzeczywiste naruszenie praw autorskich strony trzeciej;

<sup>7</sup> Za: *8 principles of open data*, [https://public.resource.org/8\\_principles.html](https://public.resource.org/8_principles.html) [dostęp dn. 1.08.2014 r.].

<sup>8</sup> Dane osobowe nie zawsze są chronione – np. lista obecności uczestników zebrania wiejskiego, zgodnie z wyrokiem NSA, sygn. I OSK 2022/12, może zostać udostępniona, dotyczy też np. list osób, które podpisały umowy z jednostkami samorządu terytorialnego (wyrok SN z 8.11.2012, I CSK 190/12).

- objęcie danych ustawą szczegółową, kompleksowo regulującą dany rodzaj informacji, wyłączając go z zakresu ustawy o dostępie do informacji publicznej;
- brak zrozumienia potrzeb osób/podmiotów wykorzystujących dane i publikowanie ich w formatach nieprzetwarzalnych maszynowo, mimo posiadania technicznych możliwości publikowania ich w otwartych formatach;
- niedostępność u wytwarzającego danych w formie spełniającej warunki otwartych danych: brak formy przetwarzalnej maszynowo, brak możliwości łatwego eksportu danych do otwartego formatu plików, brak umiejętności eksportu danych do otwartego formatu;
- podpisywanie przez instytucje publiczne umów o ponowne wykorzystywanie danych zakładające wyłączność jednej firmy na ich wykorzystanie (tego typu umowy są sprzeczne z treścią dyrektywy re-use i ustawy o dostępie do informacji publicznej);
- nieposiadanie przez instytucję publiczną aktualizowanych zbiorów danych.

Również nowo powstały system CRIP, dostępny pod adresem <https://danepubliczne.gov.pl/>, nie rozwiązuje tych problemów. W obecnej formie jest on raczej katalogiem części danych rozrzuconych po stronach internetowych poszczególnych urzędów niż centralnym repozytorium danych. Aby spełnił pokładane w nim nadzieje, musi zostać rozbudowany o możliwość prezentowania poszczególnych zbiorów danych, udostępniania ich najnowszej wersji przez interfejs programistyczny API (a nie tylko na stronie WWW), wreszcie – jednostki publiczne umieszczające dane za jego pośrednictwem muszą je udostępniać w otwartych, dobrze opisanych formatach plików, nie zaś jako dokumenty PDF czy egzotyczne pakiety 7z.

Niezależnie od CRIP, różne instytucje centralne pracują nad ułatwieniem dostępu przez sieć do części swoich zbiorów. Najczęściej tworzone są jednak systemy służące do wyszukiwania i przeglądania danych, w sztuczny sposób utrudniające ich automatyczne przetwarzania. Wśród najistotniejszych systemów ułatwiających dostęp do danych publiczny, lecz niepozwalających na ich łatwe ponowne wykorzystanie, warto wyróżnić:

- [BezpiecznyAutobus.gov.pl](http://BezpiecznyAutobus.gov.pl) oraz [HistoriaPojazdu.gov.pl](http://HistoriaPojazdu.gov.pl) pozwalające na sprawdzenie w bazie danych szczegółów dotyczących pojazdów;
- KRS online pozwalający na sprawdzenie aktualnego wpisu do rejestru przedsiębiorców lub stowarzyszeń;
- Centralna Baza Danych Ksiąg Wieczystych pozwalająca na sprawdzenie księgi wieczystej o znanym numerze;
- Centralna Ewidencja i Informacja o Działalności Gospodarczej umożliwiająca m.in. przeglądanie danych z rejestrów działalności gospodarczej.

Żaden z tych systemów nie udostępnia informacji poprzez publicznie dostępny interfejs programistyczny API. Uniemożliwia to napisanie zewnętrznego oprogramowania korzystającego z danych zawartych w państwowych bazach danych, na których oparte są wskazane wyżej portale. Dodatkowo dostęp przez przeglądarkę jest ograniczony mechanizmem *captcha*, co utrudnia dostęp zarówno osobom słabo widzącym, jak i automatom próbującym korzystać z bazy bez pośrednictwa nieistniejącego API.

## Postulowane zmiany

Pełne wykorzystanie potencjału rozwojowego i gospodarczego tkwiącego w danych przygotowywanych przez jednostki publiczne wymaga zmian zarówno mentalnościowych, jak i prawnych.

W związku z tym rekomendujemy:

- Prowadzenie szerokiej edukacji technicznej opartej na zrozumieniu specyfiki otwartych danych oraz edukowanie pracowników instytucji publicznych, by zmienić kulturę urzędniczą.
- Stworzenie osobnego aktu prawnego – Ustawy o otwartych danych publicznych – regulującego kwestie otwartości danych.
- Dokonanie przeglądu istniejącej legislacji związanej z danymi i rejestrami publicznymi w celu dostosowania jej do wymogów nowej Ustawy o otwartych danych publicznych.

- Wprowadzenie obowiązku publikowania danych w sieci w sposób spełniający warunki otwartości omówione wyżej w tym tekście;
- Wprowadzenie obowiązku publikowania w sieci danych i informacji wytworzonych lub zaktualizowanych po dniu wejścia nowej Ustawy w życie oraz publikowania danych archiwalnych, co do których został złożony wniosek o dostęp do informacji publicznej; pełna lista wyjątków od tej zasady powinna być wskazana w Ustawie o otwartych danych.
- Przeprowadzenie w ciągu 2 lat od wejścia nowej Ustawy w życie pełnej ewidencji zbiorów danych znajdujących się w jednostkach publicznych i udostępnienie jej na stronach BIP oraz w systemie CRIP.
- Przekształcenie CRIP z katalogu danych w system umożliwiający realny do nich dostęp poprzez odpowiedni interfejs programistyczny API.
- Uregulowanie kwestii otwartości surowych danych zbieranych przez państwowe instytucje badawcze (np. Instytut Meteorologii i Gospodarki Wodnej) oraz jednostki finansowane z funduszy publicznych (np. gminne przedsiębiorstwa transportowe).
- Uregulowanie kwestii dostępu do danych mogących zawierać chronione dane osobowe samodzielnie lub w połączeniu z innymi rejestrami; z założenia dane powinny być – po dokonaniu agregacji i anonimizacji – udostępniane, nie zaś zamykane przed użytkownikami.

**Łukasz Jachowicz** – od blisko 15 lat zajmuje się kwestiami otwartości oprogramowania i danych w kontekście edukacji, administracji publicznej i biznesu. Autor i współautor publikacji poświęconych przyszłości zarządzania internetem, modelom tworzenia aplikacji dla administracji publicznej w trybie konkursowym oraz otwartości rządu. Związany z Polską Grupą Użytkowników Linuxa oraz Fundacją Wolnego i Otwartego Oprogramowania. Dyrektor ds. polityki publicznej w MSLGROUP Poland. Kontakt: [lukasz@jachowicz.com](mailto:lukasz@jachowicz.com) oraz przez twitter: @ljachowicz.