



Warszawa, dnia 29 września 2016 r.

RZECZPOSPOLITA POLSKA
MINISTER CYFRYZACJI

Anna Streżyńska

BM-WSKN.053.36.2016

Panie i Panowie
Ministrowie

Według rozdzielnika

Szanowni Państwo,

uprzejmie informuję, że do Ministerstwa Cyfryzacji w dniu 10 września br. wpłynęła petycja p. Karola Breguły dotycząca między innymi podjęcia działań na rzecz opracowania standardu publikowania danych finansowych przez administrację publiczną. Mając na względzie sformułowane w niej oczekiwania, którym w istotnym zakresie nie sposób odmówić słuszności, pragnę zachęcić Państwa do udostępniania wykazów umów w otwartych formatach na stronach Biuletynu informacji Publicznej albo na stronach internetowych kierowanych przez Państwa ministerstw i instytucji podległych lub nadzorowanych oraz w Centralnym Repozytorium Informacji Publicznej (danepubliczne.gov.pl) przy wykorzystaniu interfejsu programistycznego – API.

Wiele instytucji publicznych publikuje wykazy umów na stronach Biuletynu informacji Publicznej. Odpowiadając na oczekiwania społeczne, wykazy te powinny być publikowane w otwartych formatach ułatwiających ponowne wykorzystywanie tego rodzaju danych. Pragnę zauważyć, że publikacja danych wyłącznie w formatach PDF nie tylko utrudnia ponowne wykorzystywanie danych, lecz także może być źródłem przypadkowych błędów powstałych przy konwertowaniu przez użytkowników plików do postaci edytowalnej.

Udostępnianie wykazów umów w otwartych formatach wpisuje się w realizację przyjętego w dniu 20 września br. przez Radę Ministrów Programu Otwierania Danych Publicznych. Załącznik nr 1 do Programu wskazuje standard udostępniania danych na portalu danepubliczne.gov.pl. Dokument prezentuje wytyczne dotyczące przygotowania danych, w tym zalecenia dotyczące plików danych, organizacji i lokalizacji źródeł danych, standaryzacji i formatów plików danych, interfejsu programistycznego (API) oraz wymagania dotyczące jakości danych.

Uprzejmie informuję, że Ministerstwo Cyfryzacji publikuje wykaz zawieranych umów na portalu danepubliczne.gov.pl w formatach XLSX i CSV, na stronie podmiotowej BIP w formatach XLS

i PDF oraz na swojej stronie internetowej w formacie HTML. W załączeniu przesyłam *Instrukcję przygotowania danych publicznych w formie arkusza danych*, zawierającą prostą metodę przygotowania plików danych w otwartych formatach w celu ich udostępnienia oraz zalecenia, co do struktury wykazu umów.

Z poważaniem,
Anna Streżyńska
Minister Cyfryzacji

/-podpisano elektronicznie/

Otrzymują:

1. Wiceprezes Rady Ministrów – Minister Kultury i Dziedzictwa Narodowego;
2. Wiceprezes Rady Ministrów – Minister Nauki i Szkolnictwa Wyższego;
3. Wiceprezes Rady Ministrów – Minister Rozwoju, Minister Finansów;
4. Szef Kancelarii Prezesa Rady Ministrów – Minister Członek Rady Ministrów;
5. Minister Edukacji Narodowej;
6. Minister Energii;
8. Minister Gospodarki Morskiej i Żeglugi Śródlądowej;
9. Minister Infrastruktury i Budownictwa;
10. Minister Obrony Narodowej;
11. Minister Rodziny, Pracy i Polityki Społecznej;
12. Minister Rolnictwa i Rozwoju Wsi;
13. Minister Skarbu Państwa;
14. Minister Sportu i Turystyki;
15. Minister Sprawiedliwości;
16. Minister Spraw Wewnętrznych i Administracji
17. Minister Spraw Zagranicznych;
18. Minister Środowiska;
19. Minister Zdrowia;
20. Prezes Rządowego Centrum Legislacji.

Instrukcja przygotowania danych publicznych w formie arkusza danych

Instrukcja tworzenia otwartych zasobów w popularnych formatach arkusza danych, możliwych do prostej konwersji na pliki w formacie CSV lub pokrewnym – tekstowym z separatorami.

Przygotowanie arkusza ustrukturyzowanych danych w formacie XLS lub XLSX

1. Arkusz musi posiadać jeden wiersz nagłówka, który zawiera opisy wszystkich kolumn z danymi.
2. Każda kolumna tabeli, poza wierszem nagłówka, musi zawierać dane jednego rodzaju, tzn. liczbę, tekst lub ciąg cyfrowo-tekstowy o określonej strukturze i znaczeniu, np. numer telefonu, kod pocztowy.
3. Każda kolumna tabeli powinna zawierać dane dotyczące pojedynczej wielkości (elementarnej), np. numer kodu, a nie numer kodu i miejscowość; ułatwia to sortowanie i łączenie danych z wielu tabel.
4. Zapis w każdym polu w kolumnie musi być zgodny z założonym formatem i modelem – maską, właściwą dla danego typu zawartości.
5. Nie należy stosować formatowania tekstu i niektórych znaków interpunkcyjnych właściwych dla formatowania zaawansowanego, które w pewnych kontekstach mogą imitować operatory arytmetyczne lub zmieniać wartość liczby, np. łącznik (dywiz) może prowadzić do utraty danych generując w arkuszu wynikowym błąd formuły lub zmieniać wartość liczby na ujemną.
6. Jeżeli nie można powyższych warunków spełnić ze względu na złożoność danych zawartych w tabeli, należy tabelę podzielić w sposób, który umożliwi ich spełnienie.
7. Zaleca się umieszczenie w przygotowanym arkuszu danych o charakterze referencyjnym, np. nr REGON w rekordzie dot. podmiotu; ułatwi to w przyszłości powiązanie danych w różnych zasobach i od różnych dostawców.

Niedopuszczalne jest:

- scalanie komórek zarówno w ramach wiersza, jak i kolumny – zaburza to strukturę arkusza danych,
- umieszczanie treści dotyczących jednego rekordu w kilku wierszach, np. kolejnych danych adresu (adresat, ulica, miejscowość itp.),
- wprowadzanie dodatkowych tytułów, opisów, przypisów i komentarzy poza obszarem danych tabeli,

- używanie formuł agregujących dane, np. funkcji filtrowania danych lub funkcji SUMA; wynik sumowania musi być wpisany jako liczba,
- używanie formuł pobierających dane spoza źródła danych, które będą niedostępne po przeniesieniu arkusza na inną platformę.

Wymagania dla plików danych w formacie XLS lub XLSX, aby mogły być prawidłowo wyświetlane w podglądzie na podstronie zbioru danych na portalu DanePubliczne.gov.pl

Wielkość pliku nie powinna być większa niż 100 KB, gdyż powyżej tej objętości oprogramowanie generujące podgląd zasobu może działać nieprawidłowo lub podgląd będzie niedostępny.

Ponadto arkusz danych, aby mógł być prawidłowo wyświetlany w oknie zasobu (zakładka Eksplorator Danych), musi spełniać następujące warunki właściwe dla standardu CSV, który jest podstawą automatycznego utworzenia przez narzędzie systemowe, widoku publikowanych danych:

- musi być to pojedynczy arkusz, nie skoroszyt (zbiór kilku arkuszy w jednym pliku),
- nie mogą występować puste wiersze powyżej obszaru danych (nad tabelą),
- nie mogą występować puste kolumny z lewej strony obszaru danych,
- nie może zawierać wiersza tytułu nad tabelą; tytuł tabeli może w tej sytuacji umieszczony być w polu „Nazwa zasobu” lub „Opis” (kilka przydatnych informacji o danych) na podstronie Edycji zasobu,
- nie może zawierać formuł pobierania danych z zewnętrznego zasobu; mechanizm ochrony danych Excela zachowuje ostatnio zebrane dane, lecz ich kopiowanie może zakończyć się ich zatarciem w arkuszu docelowym (wykryty zostaje nieprawidłowy argument funkcji).

Rekomendacje dotyczące struktury wykazu umów

Wykaz umów powinien zawierać co najmniej niżej wymienione kolumny:

- (i) numer umowy,
- (ii) data zawarcia umowy,
- (iii) podmiot z którym zawarto umowę,
- (iv) przedmiot umowy,
- (v) termin obowiązywania umowy,
- (vi) wartość umowy.

Od 1 stycznia 2017 r. wykaz ponadto powinien zawierać:

- (vii) komórkę Urzędu z której działalnością wiąże się zawarcie umowy,
- (viii) numer NIP/REGON/KRS podmiotu z którym zawarto umowę.

W przypadku przygotowania publikacji zbiorów do udostępnienia na portalu DanePubliczne.gov.pl wartości pól winny być elementarne (atomowe, niepodzielne).