

API do Dziedzinowych Baz Wiedzy



28 lutego 2022 r. Główny Urząd Statystyczny udostępnił na portalu api.stat.gov.pl **API do Dziedzinowych Baz Wiedzy (API DBW)**.

API DBW wzbogaca dotychczasowy katalog produktów GUS wyposażonych w interfejs API (API REGON, API TERYT, API BDL, API SDG, API STRATEG), które są dostępne na portalu api.stat.gov.pl.

Co udostępniono w API DBW?

API DBW pozwala na powszechny, otwarty i wygodny dostęp do danych i metadanych pochodzących ze zmodernizowanych i uzupełnionych zakresowo **Dziedzinowych Baz Wiedzy**. Poprzez API DBW publikowane są oficjalne dane statystyczne w długich szeregach czasowych, na szczegółowym poziomie przekrojów klasyfikacyjnych i nomenklaturowych. W DBW publikowane są dane krótkookresowe (publikowane co miesiąc lub co kwartał) oraz dane roczne. Zostały one pogrupowane w 31 obszarów tematycznych, zawierających informacje z różnych dziedzin życia społeczno-gospodarczego oraz o środowisku. W zmodernizowanych DBW znajduje się obecnie ponad 1200 zmiennych (ponad 626 mln rekordów), ale w kolejnych okresach przewiduje się sukcesywne zwiększanie publikowanego zakresu.

Opublikowane API DBW jest pierwszym rezultatem GUS wypracowanym w projekcie **Otwarte Dane plus**, w ramach którego w drugim kwartale br. zostanie uruchomiona nowa witryna do zmodernizowanych Dziedzinowych Baz Wiedzy. Do podstawowych założeń witryny należą m.in. intuicyjny interfejs użytkownika (wersja polska i angielska), możliwość zestawiania danych z różnych obszarów oraz łatwy dostęp do metadanych. Do każdej zmiennej, oprócz definicji, dołączono m.in. informacje na temat źródła danych, gestora, czy długości dostępnego szeregu czasowego. Zakres prezentowanych informacji będzie spójny z tym udostępnianym poprzez API, tj. bogatszy w stosunku do obecnego DBW. Wśród dodanych obszarów znalazły się Finanse publiczne, Gospodarka morska i śródlądowa, Gospodarka społeczna, Kultura, Świadczenia i pomoc społeczna, Turystyka i sport oraz Wymiar sprawiedliwości. Ponadto, w przypadku części z dotychczas prezentowanych w DBW tematów znacząco rozszerzono zakres informacyjny.

Czym jest API i jakie korzyści przynosi użytkownikom?

Informacje ogólne

API, czyli interfejs programowania aplikacji (ang. *Application Programming Interface*) to zestaw reguł opisujących, w jaki sposób systemy lub aplikacje komunikują się między sobą.

Interfejs ten daje możliwość szybkiego i wygodnego pobierania dużych zestawów danych. Dzięki temu, że dane można w łatwy sposób wprowadzić do własnego środowiska informatycznego, poprawia się efektywność ich przetwarzania.

Dane udostępniane przez API DBW mogą być wykorzystywane przez wszystkie zainteresowane strony, w tym:

- przedsiębiorstwa tworzące innowacyjne rozwiązania programistyczne – produkty i usługi, takie jak: aplikacje mobilne i internetowe, programy komputerowe, widżety itp.,
- użytkowników zainteresowanych danymi statystycznymi z zakresu prezentowanego w zmodernizowanych Dziedzinowych Bazach Wiedzy:

- bezpośrednio: tworząc własne zestawienia, raporty, analizy i wizualizacje (wymagany jest dostęp do narzędzia, które będzie obsługiwać import danych zewnętrznych, np. arkusz kalkulacyjny lub program do analizy danych, gdyż API jest interfejsem programistycznym, służącym wyłącznie do przeglądania i pobierania danych)
- albo pośrednio – korzystając z produktów i usług powstałych na bazie API.

Dodatkową zaletą pobierania danych przez API jest możliwość ich automatycznego aktualizowania. Aplikacje i serwisy korzystające z API pobierają dane bez konieczności ingerencji człowieka.

Informacje szczegółowe

Dane ze zmodernizowanych Dziedzinowych Baz Wiedzy udostępniane są poprzez REST-owe API w formacie JSON. Dodatkowym zaimplementowanym formatem jest CSV. Do budowy API DBW wykorzystano platformę NET.Core, a tworzenie dokumentacji API wspierał Swagger UI.

Witryna interaktywnej dokumentacji API DBW (wersja polska i angielska) podaje wskazówki i zasady przeglądania oraz pobierania do dalszego przetwarzania danych i metadanych. Pozwala ona na wypróbowanie zapytań bezpośrednio w przeglądarce.

Ze względów bezpieczeństwa dostęp do danych za pomocą API jest ograniczony. Użytkownicy potrzebujący większych limitów danych niż użytkownik anonimowy powinni zarejestrować się w celu uzyskania indywidualnego klucza API.

API do Dziedzinowych Baz Wiedzy udostępnia zasoby w ramach 4 sekcji:

- Dziedziny DBW - możliwość pobrania listy wszystkich obszarów tematycznych DBW w postaci hierarchii oraz listy zmiennych z danego obszaru tematycznego;
- Słowniki - możliwość pobrania listy pozycji ze słownika: Data, Okresy, Typ informacji, Brak wartości, Tajność oraz Flaga, czyli objaśnienia do wartości;
- Zmienne - możliwość pobrania listy pozycji ze słownika zmiennych, listy przekrojów dla określonej zmiennej, listy wymiarów oraz pozycji wymiarów dla określonego przekroju, a także pobrania danych dla określonej zmiennej w określonym przekroju i okresie za dany rok oraz pobrania pełnego zestawu metadanych dla określonej zmiennej;
- Aktualna wersja API.

Jak korzystać z API DBW bez umiejętności programowania? Podstawowe wskazówki

Krok 1. Tworzenie URL lub pobranie pliku (JSON albo CSV)

W interaktywnej dokumentacji API DBW istnieje możliwość wygenerowania linku (URL) do interesującego zasobu danych w DBW. W tym celu przy opisie wybranego zasobu należy skorzystać z przycisku *Wypróbuj* (można wtedy zmienić dostępne parametry), a następnie przycisku *Wykonaj*. Efektem poprawnego wykonania zapytania jest wygenerowanie adresu URL. W niektórych zasobach można również skorzystać z filtrowania i sortowania danych.

Przykład linku pozwalającego na pobranie **listy zmiennych** (area/area-variable) **z obszaru tematycznego Migracje ludności** (id-obszaru=17), dla którego obszarem nadrzędnym jest Demografia, w języku polskim (lang=pl)

<https://api-dbw.stat.gov.pl/api/1.0.0/area/area-variable?id-obszaru=17&lang=pl>

Przykład linku pozwalającego na pobranie **danych dla zmiennej Placówki gastronomiczne** (id-zmienna=102) w **przekroju Polska, województwa, Rodzaj placówki gastronomicznej** (id-przekroj=182) dla roku 2020; dane roczne, 50 wyników na stronie, w języku polskim

<https://api-dbw.stat.gov.pl/api/1.0.0/variable/variable-data-section?id-zmienna=102&id-przekroj=182&id-rok=2020&id-okres=282&ile-na-stronie=50&numer-strony=0&lang=pl>

Przykład linku pozwalającego na pobranie **danych dla zmiennej Pracujący w gospodarce narodowej** (id-zmienna=397) w **przekroju Polska; Sekcje PKD 2007 + Przemysł (zgrupowane BCDE); Płeć; Sektor własności; Status zatrudnienia pracujących** (id-przekroj=317) dla roku 2020; dane roczne, 50 wyników na stronie, w język angielskim

<https://api-dbw.stat.gov.pl/api/1.0.0/variable/variable-data-section?id-zmienna=397&id-przekroj=317&id-rok=2020&id-okres=282&ile-na-stronie=50&numer-strony=0&lang=en>

Numer Id obszaru, Id zmiennej oraz Id przekroju można odnaleźć wykonując wcześniej dodatkowe zapytania o: listę wszystkich obszarów, listę wszystkich zmiennych (listę pozycji ze słownika Zmienne) oraz listę przekrojów dla określonej zmiennej.

Alternatywnie, w dokumentacji jest możliwość pobrania pliku JSON lub CSV.

Krok 2. Import danych poprzez arkusz kalkulacyjny lub program/narzędzie do analizy danych

W celu wyświetlenia danych z API należy skorzystać z importu danych z sieci Web (należy wkleić URL) lub podobnej funkcji w programie, ewentualnie można zaimportować pobrany wcześniej plik JSON lub CSV.

Przydatną opcją przeglądania danych z API są **przeglądarki tekstów w formacie JSON** (JSON Viewer). Są one dostępne jako (zazwyczaj) bezpłatne serwisy internetowe lub dodatki do przeglądarek. Jest to rozwiązanie do prostszych zapytań, bardzo wygodne do szybkiego przejrzania lub odszukania danych (np. odnalezienia numeru Id, przeglądu listy zmiennych itp.). Nie jest to narzędzie do analizy danych. W odpowiednie miejsce w serwisie należy wkleić wygenerowany URL lub załadować plik JSON.

API do Dziedzinowych Baz Wiedzy jest efektem działań nad doskonaleniem dostępności danych statystycznych. Jego opublikowanie wpływa na zwiększenie poziomu otwartości danych statystycznych i możliwości ich powtórnego wykorzystania. API zostało przygotowane zgodnie z obowiązującymi standardami otwartości danych.

Prace nad API DBW były prowadzone w ramach partnerskiego projektu Otwarte dane plus i współfinansowane z Programu Operacyjnego Polska Cyfrowa.

#StatystykaPrzyszłościJestJużTu