

Analiza części dostępnych w serwisie danepubliczne.gov.pl baz danych

Przed Wami krótka analiza części baz danych, które znaleźć możecie na www.danepubliczne.gov.pl. Przygotowując ją przeprowadziliśmy demo (Python), które pozwala określić, gdzie dane i w jakim formacie można znaleźć, co się w nich znajduje, gdzie znajdziecie przykładowy kod (demo) oraz jakie zmapowaliśmy ewentualne problemy podczas ich wykorzystywania.

Wszystkie demo znajdziecie na GitHubie: <https://github.com/EE/hackaton-examples>

Gdybyście z jakiegoś powodu nie mogli dostać się do bazy / pojawiłby się z nią jakiś problem - dajcie znać Organizatorom (szare identyfikatory), a dostarczymy Wam backup na dysku.

Powodzenia!

Zespół EE

Baza	Technologia	Dokumentacja	Dane w bazie	Metody wyszukiwania	Backup EE	Demo aplikacji korzystającej z danych	Potencjalne problemy
TERYT	SOAP	https://danepubliczne.gov.pl/data/set/ed74a82c-bdb5-4afc-8a36-0317e4b00ce5/resource/6b600174-3340-48cc-a211-2adc6ecc0afd/download/instrukcjatechnicznaslugiterytws1.zip	struktura podziału terytorialnego: regiony, podregiony, województwa, powiaty, miejscowości, ulice	Jednostki w innej jednostce (np powiaty w województwie). Nazwy ulic.	-	https://github.com/EE/hackaton-examples/tree/master/teryt drzewko struktury organizacyjnej, statystyki nazw ulic	Niejasno opisany sposób autoryzacji (trzeba przekazać użytkownika i hasło w nagłówkach SOAP). Bez rejestracji i weryfikacji możliwe jest korzystanie z serwera testowego. Serwer testowy zwraca rzeczywiste dane.
REGON	SOAP/XML	https://danepubliczne.gov.pl/data/set/2b71f322-7930-4cc1-a5af-d49ecb7d1a4d/resource/35168d71-4785-44f3-b0a3-b602aa25c379/download/regoninstrukcjatechnicznabir1dlapodmiotowkomercyjnychv010.zip	dane o działalności gospodarczych: REGON, NIP, KRS, obszary działalności, osoby	REGON, NIP, KRS	-	https://github.com/EE/hackaton-examples/tree/master/regon listowanie danych o labEE	Odpowiedzi od serwera jako HTTP multipart. Odpowiedzi od serwera jako napis XML. Bez rejestracji i weryfikacji możliwe jest korzystanie z serwera testowego. Serwer testowy zwraca rzeczywiste dane.
CEIDG	SOAP/XML	dostępna po zalogowaniu do serwisu https://datastore.ceidg.gov.pl		Obszar na którym zarejestrowana jest działalność.	-	https://github.com/EE/hackaton-examples/tree/master/ceidg listowanie e-maili działalności w Warszawie	Konieczna rejestracja i zaakceptowanie konta. Problemy z logowaniem za pomocą certyfikatu pod Linuxem. Dane zwracane jako napis XML.
SIIS	REST/JSON	https://siis-api.itl.waw.pl/doc/	Dane o punktach dostępu do sieci szerokopasmowej (współrzędne). Dane o właścicielach końcówek. (Dane z lat [2012-2015].)	po dowolnej kolumnie	Zgłoś się do organizatorów	https://github.com/EE/hackaton-examples/tree/master/siis narzędzie do pobierania całej bazy, rysowanie węzłów sieci na mapie	-
SIRS	REST/JSON	https://sirs-api.itl.waw.pl/doc/	Częściowe dane o sieci szerokopasmowej: węzły (wraz z współrzędnymi), połączenia. (Graf jest	po dowolnej kolumnie	Zgłoś się do organizatorów	https://github.com/EE/hackaton-examples/tree/master/sirs narzędzie do pobierania całej	-

			niespójny.)			bazy, analiza grafu sieci	
BN	REST/JSON	https://data.bn.org.pl/docs	Rekordy bibliograficzne. Rekordy wzorcowe. Troche danych jest w formacie MARC.	rodzaj, data utworzenia, data modyfikacji	Zgłoś się do organizatorów	https://github.com/EE/hackaton-examples/tree/master/bn narzędzie do pobierania całej bazy, rysowanie wykresu aktywności dodawania nowych wpisów względem czasu	-
POLONA	REST/JSON	https://polona.pl/api/	Dane wizualizowane na stronie polona.pl: rekordy bibliograficzne, sugestie wyszukiwania, kontrola różnych zdarzeń na stronie.		-	-	Część danych dostępna tylko dla zalogowanych użytkowników.
UM-Gdańsk: dane ze stacji meteorologicznych	REST/JSON	http://pomiar.gdm.pl/rest	Opad, poziom wody, siła wiatru, kierunek wiatru. Rozdzielczość 1h, historia kilku miesięcy.	-	-	-	-