


RAPORT

w zakresie integracji systemów w celu automatyzacji aktualizacji wskaźnika

Warszawa, 01.2021

Cenatorium Sp. z o.o., ul. Piękna 68, 00-672 Warszawa, www.cenatorium.pl
KRS 0000392156, NIP 7393840515, REGON 281078837, kapitał zakładowy: 1.500.000,00 PLN



Spis treści

1. Wstęp	3
2. Udostępnienie danych.....	4
2.1. Bank Danych Lokalnych	4
2.1.1. Ocena.....	5
2.2. Centralny Rejestr Form Ochrony Przyrody.....	5
2.2.1. Ocena.....	6
2.3. GIOŚ – Bank Danych Pomiarowych	7
2.3.1. Ocena.....	8
2.4. Dane publiczne IMGW-PIB	8
2.4.1. Ocena.....	8
2.5. CEPiK.....	9
2.5.1. Ocena.....	9
2.6. Rejestry lokalne	9
2.6.1. Ocena.....	9
3. Podsumowanie	10
3.1. Rekomendacja	10

1. Wstęp

Celem niniejszego opracowania jest przedstawienie rekomendacji w zakresie integracji systemów w celu budowy bazy danych niezbędnej do wyznaczenia/wyliczenia wskaźników audytu miejskiego i analizy warunków życia opracowywanego w ramach tego projektu uniwersalnego dla polskich miast modelu danych.

W trakcie audytu danych miejskich przeanalizowano zbiory danych gromadzone w systemach na poziomie

- lokalnym m.in.:
 - wydziały/referaty Urzędu Miasta Kielce:
 - Wydział Budżetu i Księgowości,
 - Wydział Edukacji, Kultury i Sportu,
 - Ewidencja Gruntów i Budynków,
 - Wydział Gospodarki Komunalnej i Środowisk,
 - jednostki organizacyjne i spółki Miasta Kielce:
 - Miejski Zarząd Dróg w Kielcach,
 - Zarząd Transportu Miejskiego w Kielcach,
 - Komenda Straży Miejskiej w Kielcach,
 - Miejskie Przedsiębiorstwo Energetyki Ciepłej Sp. z o.o. w Kielcach,
 - Przedsiębiorstwo Gospodarki Odpadami Sp. z o.o. w Kielcach,
 - Wodociągi Kieleckie Sp. z o.o. ,
 - Miejski Urząd Pracy w Kielcach,
 - Miejski Ośrodek Pomocy Rodzinie w Kielcach,
 - jednostki zewnętrzne:
 - Publiczne uczelnie wyższe w Kielcach,
 - Świętokrzyska Izba Pielęgniarek i Położnych,
 - Polska Grupa Energetyczna Dystrybucja S.A. Oddział Skarżysko-Kamienna,
 - Świętokrzyskiego Centrum Ratownictwa Medycznego i Transportu Sanitarnego w Kielcach,
 - Delegatura UKE w Kielcach,
- centralnym:
 - Komenda Główna Państwowej Straży Pożarnej,

- Komenda Główna Policji,
- Narodowy Fundusz Zdrowia,
- Główny Urząd Statystyczny,
- Generalna Dyrekcja Ochrony Środowiska,
- Główny Inspektorat Ochrony Środowiska,
- Ministerstwo Cyfryzacji/CEPiK,
- Instytut Meteorologii i Gospodarki Wodnej - IMGW PIB,
- Państwowy Instytut Geologiczny - Państwowy Instytut Badawczy.

2. Udostępnienie danych

W ramach tego rozdziału opisane i podsumowane zostaną wybrane rejestry lub grupy rejestrów ze względu na możliwości ich integracji z systemami zewnętrznymi lub możliwości porównywania miast w zakresie wskaźników wyznaczonych na podstawie danych pochodzących z tych rejestrów.

2.1. Bank Danych Lokalnych

Bank danych lokalnych stanowi bazę ponad 40 tys. cech statystycznych o gospodarce, społeczeństwie i środowisku. W ramach systemu publikowanych jest 1662 statystyk dla Kielc, z których :

- 9263 dostępnych jest w odniesieniu do powiaty,
- 6075 dostępnych jest w odniesieniu do gminy,
- 1324 dostępnych jest w odniesieniu do miejscowości statystycznej.

System prowadzony jest przez Główny Urząd Statystyczny.

Aplikacja API pozwala przeglądać i pobierać pełen zakres danych znajdujących się w Banku Danych Lokalnych.

Na API nałożone są limity dotyczące liczby odpytań w zadanych okresach. Poniższa tabela prezentuje wartości limitów dla użytkowników anonimowych oraz zarejestrowanych. Dostęp do zwiększonych wartości limitów wymaga rejestracji – użytkownik zarejestrowany uzyskuje osobisty klucz API.

Tabela 1. Limity odpytań API BDL w zależności od okresu i rodzaju użytkownika

Okres	Użytkownik anonimowy	Użytkownik zarejestrowany
1 s	5	10
15 m	100	500

12 h	1000	5000
7 d	10000	50000

REST-we API udostępnia dane w formatach:

- XML
- JSON

Dostęp do danych poprzez API odbywa się na dwa sposoby:

- dla jednej zmiennej i wielu jednostek terytorialnych
- dla jednej jednostki terytorialnej oraz wielu zmiennych

2.1.1. Ocena

Bank Danych Lokalnych stanowi główne źródło danych dla parametrów lub składowych parametrów w przypadku ponad 100 wskaźników. Statystyki gromadzone w ramach tego zasobu stanowią również alternatywne źródło danych dla większości wskaźników ustalonych w ramach Etapu 1 w odniesieniu do poziomu miasta/gminy/powiatu – wynika to z faktu obowiązku przekazywania do Głównego Urzędu Statystycznego sprawozdań/zestawień dotyczących funkcjonowania poszczególnych analizowanych oddzielnie jednostek.

Dane gromadzone w systemie nie pozwalają na analizy rozkładu przestrzenne badanych aspektów funkcjonowania miasta w ramach granic miasta.

Ze względu na to, że zbiór jest prowadzony na poziomie centralnym, a dane w nim gromadzone choć bardzo często źródłowo pochodzą ze sprawozdań przekazywanych przez jednostki poziomu lokalnego to posiadają wspólną definicję i jednolitą metodologię wyznaczania dla wszystkich jednostek terytorialnych na obszarze całego kraju. Dzięki temu zbiór pozwala na szybkie i wiarygodne porównanie różnych jednostek terytorialnych.

2.2. Centralny Rejestr Form Ochrony Przyrody

Rejestr prowadzony jest przez Generalnego Dyrektora Ochrony Środowiska (Generalna Dyrekcja Ochrony Środowiska). Zawiera dane reprezentujące formy ochrony przyrody, m.in.:

- parki narodowe,
- parki krajobrazowe,

- rezerваты,
- obszary chronionego krajobrazu,
- specjalne obszary ochrony ("siedliskowa" Natura 2000),
- obszary specjalnej ochrony ("ptasia" Natura 2000),
- zespoły przyrodniczo-krajobrazowe,
- stanowiska dokumentacyjne,
- użytki ekologiczne,
- pomniki przyrody,
- obszary RAMSAR,
- korytarze ekologiczne,
- mezoregiony fizycznogeograficzne,
- regiony biogeograficzne.

Dane gromadzone w rejestrze poza tym, że publikowane są w ramach interaktywnych Geoserwisów, udostępniane są również w postaci:

- danych wektorowych w formacie plików ESRI Shapefile,
- usługa przeglądania WMS (dane rastrowe),
- usługa pobierania WFS (dane wektorowe).

Na szczególną uwagę zasługuje usługa WFS, która pozwala na pobranie jednocześnie danych geometrycznych i opisowych (tabelarycznych) poszczególnych form ochrony przyrody. Usługa ta pozwala na automatyzację procesu integracji pomiędzy systemami.

2.2.1. Ocena

Rejestr Form Ochrony Przyrody jest źródłem danych potrzebnych do wyznaczenia zaledwie kilku wskaźników. Jest on jednak źródłem, które integruje formy ochrony ustanawiane lub utworzone przez kilka organów:

- rady gmin w zakresie pomników przyrody, stanowisk dokumentacyjnych, zespołów przyrodniczo-krajobrazowych i użytków ekologicznych,
- sejmiki wojewódzkie w zakresie parków krajobrazowych i obszarów ochronionego krajobrazu,
- regionalnych dyrektorów ochrony środowiska w zakresie rezerwatów przyrody,
- Generalnego Dyrektora Ochrony Środowiska w zakresie obszarów Natura 2000,
- dyrektorów parków narodowych w zakresie parków narodowych.

Dodatkowo jest to źródło danych, które można łatwo zintegrować z systemami zewnętrznym:

- usługa WMS w zakresie przeglądania danych (dane rastrowe),
- usługa WFS w zakresie pobierania danych (dane wektorowe i opisowe).

2.3. GIOŚ – Bank Danych Pomiarowych

Bank danych pomiarowych stanowi rejestr archiwalnych danych pomiarów (lub ich agregacji) dotyczących zanieczyszczeń powietrza.

Dane gromadzone w zasobie pochodzą ze stacji lub stanowisk pomiarowych rozlokowanych na terenie całego kraju.

Zbiór pozwala na analizę wyników pomiarów następujących rodzajów zanieczyszczeń powietrza:

- pył zawieszony PM10,
- pył zawieszony PM2.5,
- ozon,
- dwutlenek siarki,
- dwutlenek azotu,
- benzen,
- tlenki azotu,
- ołów w PM10,
- benzo(a)piren w PM10,
- arsen w PM10,
- nikiel w PM10,
- kadm w PM10,
- tlenek węgla,
- propen.

Wyniki pomiarów mogą zostać udostępnione w postaci następujących agregatów:

- wyniki pomiarów (uśrednienie 1-godzinne),
- wyniki pomiarów (uśrednienie 24-godzinne),
- agregaty danych (Średnia dobową),
- agregaty danych (Maksimum ze średnich 8-godzinnych),
- parametry statystyczne.

Rejestr prowadzony jest przez Główny Inspektorat Ochrony Środowiska.

Zasób umożliwia pobieranie danych archiwalnych w postaci np.: plików .csv.

Dodatkowo GIOŚ umożliwia pobieranie bieżących danych pomiarowych (dla C6H6, PM10,SO2,NO2,CO,PM2.5,O3), informacji o stacjach i stanowiskach pomiarowych oraz polskim indeksie jakości powietrza przy pomocy usługi API.

2.3.1. Ocena

Zbiór danych umożliwia wyliczenie kilkunastu wskaźników dotyczących jakości powietrza.

Ze względu na fakt, iż pomiary, jak i ich opracowanie prowadzone są w oparciu o tę samą metodologię na obszarze całego kraju, wskaźniki wyznaczone na podstawie tych danych są porównywalne dla różnych jednostek.

Niestety rozkład stacji i stanowisk pomiarowych nie zapewnia pełnego pokrycia dla Polski. Punkty pomiarowe rozlokowane są w okolicy lub na terenie większych miast oraz na obszarach szczególnie narażonych na występowanie podwyższonych, niekorzystnych wartości indeksu jakości powietrza.

2.4. Dane publiczne IMGW-PIB

Rejestr prowadzony przez Instytut Meteorologii i Gospodarki Wodnej Państwowego Instytut Badawczy, który zawiera dane:

- dane synoptyczne,
- dane hydrologiczne,
- dane meteorologiczne.

Dane pochodzą ze stacji pomiarowych rozlokowanych na obszarze całego kraju (w przypadku Kielc stacja znajduje się poza granicami miasta).

Dostęp do danych odbywa się poprzez API. Usługa umożliwia pobieranie danych w następujących formatach:

- xml,
- csv,
- html.

W ramach zasobu możliwe jest pozyskanie danych aktualnych lub wyników pomiarów archiwalnych.

2.4.1. Ocena

Dane gromadzone w zasobie pozwalają na wyznaczenie wartości kilkunastu wskaźników.

Dane dla obszaru całej Polski gromadzone są w jednym, centralnym miejscu, a pomiary dokonywane są w oparciu o tę samą metodologię – pozwala to na łatwe porównanie wyników z różnych stacji.

2.5. CEPiK

Centralna Ewidencja Pojazdów i Kierowców (CEPiK) gromadzi dane o pojazdach zarejestrowanych w Polsce oraz o osobach uprawnionych do kierowania pojazdami. Dane udostępniane poprzez API nie zapewniają możliwości analizy rozkładu przestrzennego zjawisk związanych z pojazdami zarejestrowanymi na terenie miasta ale umożliwiają wyznaczenie wartości wskaźników, które opisują strukturę pojazdów ze względu na ich rodzaj, wiek, pojemność silnika oraz rodzaj paliwa na poziomie miasta.

Dostęp do danych odbywa się poprzez API. Usługa umożliwia pobieranie danych w formacie JSON.

2.5.1. Ocena

Dane gromadzone w zasobie pozwalają na wyznaczenie wartości kilkunastu wskaźników.

Dane dla obszaru całej Polski gromadzone są w jednym, centralnym miejscu, a sposób ich udostępniania pozwala na automatyzację procesu importu danych oraz wyliczania wskaźników.

2.6. Rejestry lokalne

Kolejną grupą są rejestry prowadzone w ramach systemów lokalnych (np. w ramach wydziałów Urzędu Miasta lub jednostek organizacyjnych spółek miasta).

Dane gromadzone w tych rejestrach bardzo często stanowią dane/informacje źródłowe, na podstawie których tworzone są opracowania ogólnokrajowe. Poszczególne jednostki w ramach swojej działalności zobligowane są do przekazywania zestawień do rejestrów centralnych.

W przypadku analizowanych rejestrów lokalnych w większości przypadków możliwości udostępniania danych ograniczały się do ich przekazania w formie generowanych/eksportowanych ręcznie plików excel, csv, .shp. Choć rozwiązania te integrują się z różnymi systemami zewnętrznymi, są to połączenia zdefiniowane w trakcie tworzenia i rozwoju aplikacji. Brakowało narzędzi/usług wymiany danych (np. API) z nowymi systemami.

2.6.1. Ocena

Dane oferowane w ramach rejestrów lokalnych umożliwiają analizę rozkładu przestrzennego badanych zjawisk na obszarze miasta.

Ze względu na manualny lub półautomatyczny sposób eksportu danych lub dostęp do bazy danych źródłowej/widoku bazy danych możliwa jest większa elastyczności w zakresie definiowane przekazywanych parametrów na potrzeby wyznaczania wskaźników.

Zdarza się, że odpowiadające sobie zasoby w ramach różnych jednostek terytorialnych (np. różnych miast/powiatów) prowadzone są w innych systemach – w przypadku Rejestru Cen i Wartości

Nieruchomości systemów do prowadzenie tego zasobu było 12¹. O ile model danych zasobu nie jest regulowanych jednolitymi przepisami, może to oznaczać, że dane (ich zakres i format) oraz sposób udostępniania mogą się różnić. Ze względu na to wykorzystanie tych danych na potrzeby wyznaczenia wskaźnika będzie uniemożliwiało wykorzystanie uniwersalnego dla polskich miast modelu transformacji danych, będzie on wymagał dostosowania w ramach różnych jednostek (miast/powiatów).

3. Podsumowanie

W ramach tego etapu przeanalizowane zostały zarówno rejestry prowadzone na szczeblu lokalnym, jak i centralnym.

W przypadku zasobów prowadzonych na poziomie lokalnym istnieje możliwość dostępu do danych źródłowych. Ze względu sposób eksportu danych lub dostęp do baz danych źródłowych możliwa jest większa elastyczność w zakresie dostosowywania udostępnianych danych do struktury parametrów wskazanych w definicji wskaźników.

Dostęp do danych źródłowych umożliwia również analizę rozkładu przestrzennego badanego aspektu życia miejskiego.

Sposób i zakres prowadzenia rejestrów lokalnych może się różnić w poszczególnych miastach, co może uniemożliwiać wykorzystanie modelu transformacji danych opracowanego w ramach uniwersalnego modelu danych dla polskich miast. W konsekwencji może to utrudnić porównywanie miast w ramach wybranych wskaźników.

W przypadku rejestrów centralnych zakres i forma ich prowadzenia jest jednolita pod względem struktury, definicji danych oraz metodologii ich pozyskiwania/wyznaczania. Dzięki temu łatwiej jest zastosować jednolity w skali całego kraju model danych niezbędnych do automatyzacji procesu wyznaczania wskaźników miejskich.

Wybrane serwisy publiczne oferują również usługi udostępniania danych, które pozwalają na integrację z zewnętrznymi systemami.

W przypadku rejestrów centralnych zakres i struktura udostępnianych danych są określone i nie można ich dowolnie definiować.

3.1. Rekomendacja

W zakresie integracji z rejestrami rekomenduje się w pierwszej kolejności integracje ze zbiorami prowadzonymi na szczeblu centralnym:

- Bank Danych Lokalnych – API BDL,
- Centralny Rejestr Form Ochrony Przyrody – usługa WFS,
- Bank Danych Pomiarowych – GIOŚ,
- Dane publiczne IMGW-PIB – API,
- CEPiK – API.

¹ Halik Ł., 2018, Information and Communication Systems Used for Keeping the Register of Real Estate Prices and Values (RREPV) in Poland

Dla wskaźników, dla których powyższe zbiory nie są źródłem pozwalającym na wyznaczenie ich wartości zgodnie z definicją lub zakresem przestrzennym, rekomenduje się wykorzystanie danych i zestawień cyklicznie generowanych przez poszczególne jednostki szczebla lokalnego.

W przypadku większości wskaźników proponowany okres ich aktualizacji wynosi nie mniej niż 12 miesięcy, dlatego ze względu na ograniczoną uciążliwość prac związanych z przygotowaniem/eksportem poszczególnych zestawień lub danych nie rekomenduje się rozpoczęcia prac wdrożeniowych dedykowanych modułów w ramach pierwszej fazy integracji.