

1 Informacje o Systemie Hurtownia Danych i Biznes.gov.pl

Hurtownia danych jest źródłem wiarygodnych danych i raportów, z których może korzystać każdy zainteresowany, po wcześniejszej rejestracji. Informacje o przedsiębiorcach udostępniane są bezpłatnie, wg idei otwartych danych (open data), czyli formy dzielenia się danymi publicznymi w sposób przejrzysty, przydatny i zgodny ze standardami oraz przepisami prawa.

W skład Hurtowni Danych wchodzi trzy aplikacje webowe:

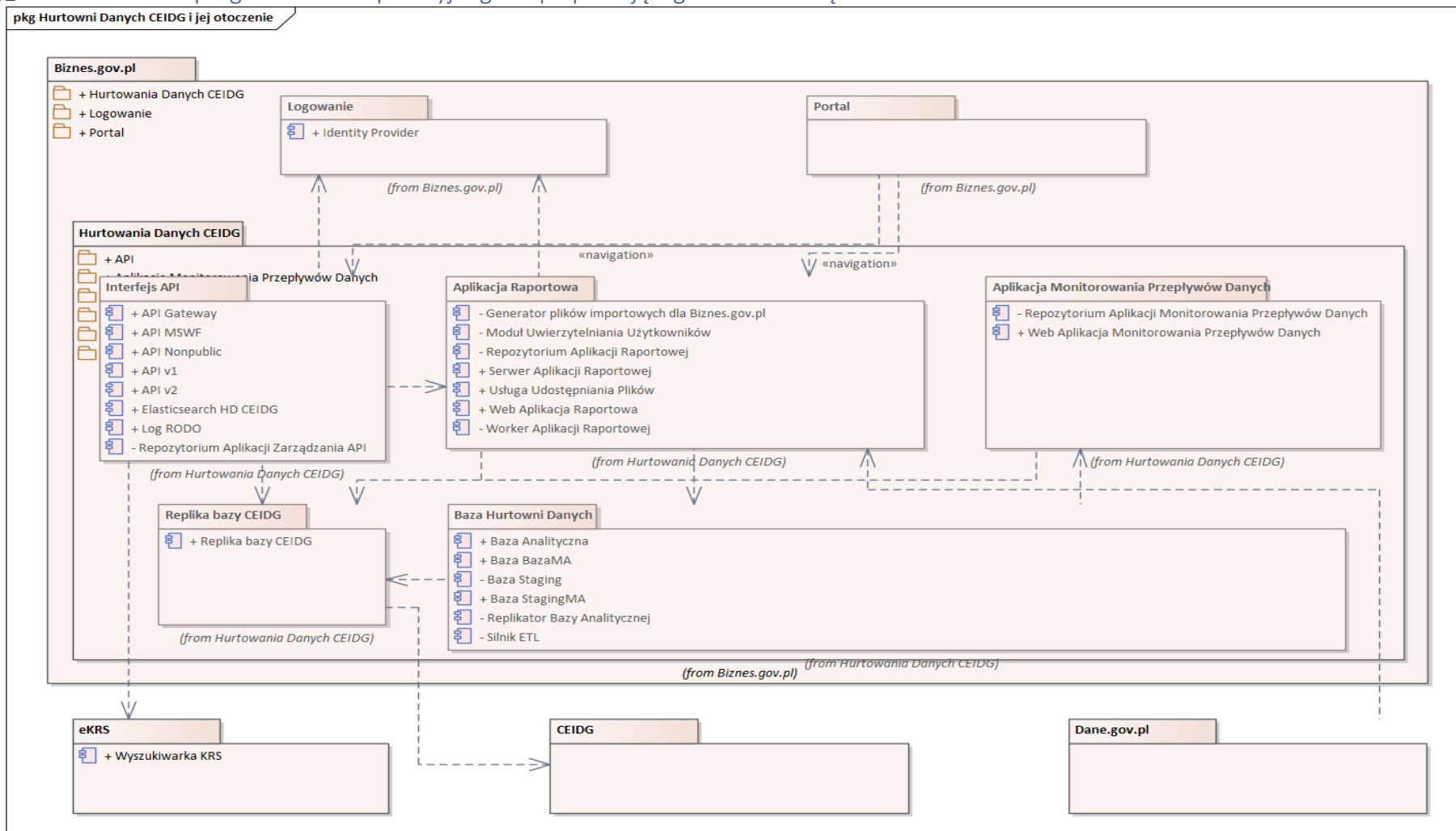
- **Interfejsu API (Gravitee) służącej do administrowania interfejsem,**
- **Aplikacji Raportowej (Wyn Enterprise) służącej do administrowania i generowania raportów,**
- **Aplikacji Monitorowania Przepływów Danych (Havran) służącej do oceny parametrów jakościowych i ilościowych transferów danych Hurtowni.**

Możliwość skorzystania z poszczególnych aplikacji oraz funkcji uzależniona jest od ról które są dostępne dla konkretnego, uwierzytelnionego Użytkownika.

Hurtowni Danych korzysta z usług infrastruktury:

- Ośrodek Podstawowy Hurtowni (Infrastruktura Chmury). Ośrodek podstawowy jest usługą typu PaaS. W ośrodku jest eksploatowana Instancja Produkcyjna Hurtowni Danych.
- Ośrodek Zapasowy Hurtowni wchodzący w skład Infrastruktury MRiT. Ośrodek zapasowy należy traktować jako fizyczny sprzęt komputerowy.

1.1 Architektura oprogramowania aplikacyjnego współpracującego z Hurtownią



Rysunek 1-1 Hurtownia Danych CEIDG i jej otoczenie

Opis modułów z którymi współpracuje Hurtownia

Moduł	Przeznaczenie i integracje
Portal Biznes.gov.pl	Serwis, w ramach którego udostępniona jest strona dane.biznes.gov.pl, gdzie zamieszczone są informacje, artykuły i link do nadawania uprawnień do Hurtowni Danych.
Logowanie Biznes.gov.pl	Serwis dostarczający usługę uwierzytelniania i autoryzacji Użytkowników.
eKRS	Serwis eKRS to serwis prowadzony przez Ministerstwo Sprawiedliwości udostępniający dokumenty i informacje z Krajowego Rejestru Sądowego (KRS). Dane i informacje z eKRS udostępniane są poprzez API i stronę internetową https://ekrs.ms.gov.pl/ . Usługi eKRS umożliwiają wyszukiwanie podmiotów z rejestru KRS.
CEiDG	Rejestr CEiDG zawiera m.in. informacje o osobach fizycznych prowadzących działalność gospodarczą oraz o ich uczestnictwie w spółkach cywilnych zgodnie z przepisami Ustawy Liczba wpisów w rejestrze ponad 6 mln.
Dane.gov.pl	Serwis realizujący cel Centralnego Repozytorium Informacji Publicznej, wskazanego w Ustawie o dostępie do informacji publicznej (Dz. U. Nr 112, poz. 1198, z późn. zm.) jako jeden z trybów dostępu i ponownego wykorzystywania informacji publicznej. W serwisie tym cyklicznie publikowane są wybrane raporty.

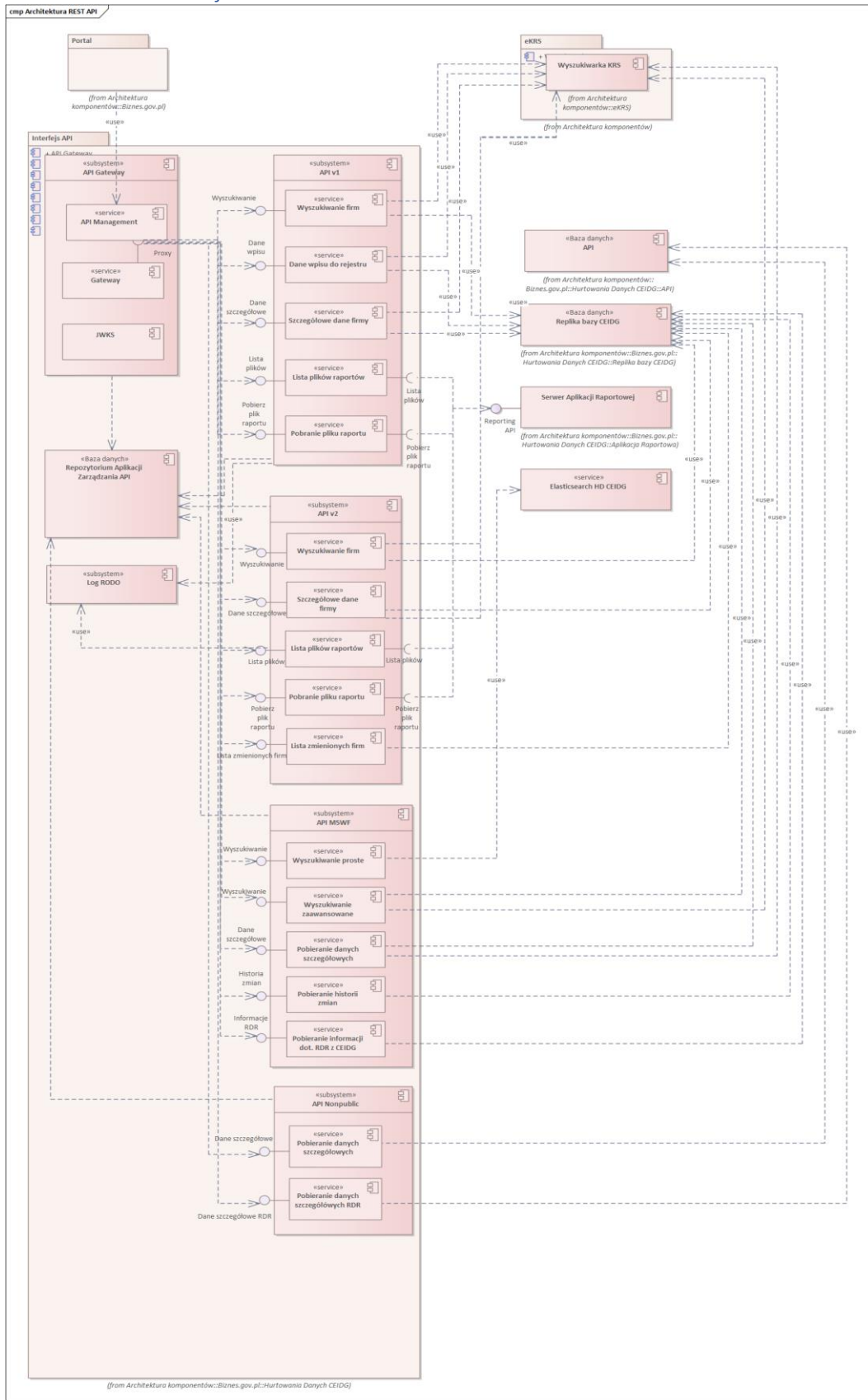
Tabela 1-1 Opis modułów współpracujących z hurtownią

Opis modułów funkcjonalnych Hurtowni

Moduł	Przeznaczenie i integracje
Interfejs API	Moduł realizujący usługi dostępu do danych z Repliki bazy CEiDG oraz Bazy Hurtowni Danych za pośrednictwem API (application programming interface). Integralną częścią Interfejsu API jest moduł administracyjny służący do zarządzania API.
Aplikacja Raportowania	Aplikacja umożliwiająca tworzenie oraz wyświetlanie raportów bazujących na danych Bazy Hurtowni Danych. Aplikacja służy również do zarządzania źródłami, zbiorami danych, szablonami raportów, regułami dystrybucji raportów, itp.
Aplikacja Monitorowania Przepływów Danych	Aplikacja przeznaczona dla Administratorów do monitorowania procesów ETL.
Replika Bazy CEiDG	Replika operacyjnej bazy danych Rejestru CEiDG, aktualizowana na bieżąco (tzw. replika transakcyjna); stanowi bezpośrednio źródło danych biznesowych dla Interfejsu API oraz Bazy Hurtowni Danych, jak również jest źródłem informacji o rolach urzędników dla Aplikacji Raportowej.
Baza Hurtowni Danych (Analityczna)	Główna baza danych zasilająca Aplikację Raportową, w której składowane są dane z Repliki Bazy CEiDG w formie zoptymalizowanej pod kątem raportowania.

Tabela 1-2 Opis Modułów Hurtowni

1.2 Architektura Interfejsów API



Rysunek 1-2 Obecna Architektura Interfejsu API Hurtowni Danych

Opis struktury modułów interfejsu API Hurtowni

Moduł	Przeznaczenie i integracje
API Gateway	Umożliwia nakładanie limitów na przesyłanie danych lub liczby zapytań w wybranym okresie czasu. Udziela dostępu do poszczególnych usług API lub odmawia dostępu w przypadku podania niezgodnego tokenu JWT. Pozwala na zarządzanie dostępem i użytkownikami API poprzez moduł API Management. Użytkownik może subskrybować wybrane API. Konieczna jest rejestracja użytkownika w API Gateway (poprzez udostępnioną usługę REST API w ramach API Gateway). Rejestracja jest dokonywana w Biznes.gov.pl w momencie generowania tokenu JWT dla użytkownika.
API	Występuje pięć usług zapewniających realizację wyszukiwania podmiotów w Replice bazy CEIDG oraz bezpośrednio w systemie zewnętrznym eKRS za pośrednictwem Wyszukiwarki KRS. API łączy wyniki znalezione w Replice bazy CEIDG oraz otrzymane przez WebService Wyszukiwarka KRS. Pozostałe usługi do pobierania listy raportów oraz konkretnego raportu wykorzystują do tego Aplikację Raportową działającą w ramach Hurtowni Danych CEIDG.
API MSWF	Występuje pięć usług zapewniających wyszukiwanie podmiotów zarejestrowanych w CEIDG oraz eKRS: <ul style="list-style-type: none"> - Wyszukiwanie proste umożliwia poprzez podanie w języku naturalnym zapytania i otrzymania wyników najbardziej zbliżonych do zapytania - Wyszukiwanie zaawansowane umożliwia przeszukiwanie zasobów CEIDG oraz eKRS przy użyciu zamkniętego katalogu kryteriów wyszukiwawczych. - Pobieranie informacji o wskazanych wpisach z CEIDG lub eKRS. - Pobieranie na podstawie unikalnego numeru KRS informacji zgromadzonych w CEIDG na temat koncesji, zezwoleń, licencji oraz wpisów do rejestrów działalności regulowanej jak również kwalifikacjach zawodowych w kontekście podmiotu zarejestrowanego w eKRS. - Pobieranie historii zmian zachodzących we wpisach pochodzących z CEIDG.
API NONPUBLIC	Występują dwie usługi zapewniające wyszukiwanie podmiotów zarejestrowanych w CEIDG zawierające dane niepubliczne: <ul style="list-style-type: none"> - Pobieranie danych szczegółowych - Pobieranie danych szczegółowych RDR
Repozytorium Aplikacji Zarządzania API	Zawiera informacje o kontaktach użytkowników, subskrypcje do API, konfigurację limitów ilości żądań dla użytkowników.
Log RODO	Logi generowane podczas wyświetlania danych osobowych w raporcie. W logu rejestrowana jest informacja o tym kiedy, komu i jakie raporty z danymi osobowymi zostały udostępnione (format JSON).
Replika bazy CEIDG	Replika operacyjnej bazy danych Rejestru CEIDG, aktualizowana na bieżąco (tzw. replika transakcyjna); stanowi bezpośrednio źródło danych biznesowych dla Interfejsu API oraz Bazy Hurtowni Danych, jak również jest źródłem informacji o rolach urzędników dla Aplikacji.
eKRS	Serwis eKRS to serwis prowadzony przez Ministerstwo Sprawiedliwości udostępniający dokumenty i informacje z Krajowego Rejestru Sądowego (KRS). Dane i informacje z eKRS udostępniane są poprzez API i stronę internetową https://ekrs.ms.gov.pl/ . Usługi eKRS umożliwiają wyszukiwanie podmiotów z rejestru KRS.

Tabela 1-3 Opis modułów wchodzących w skład Hurtowni Danych