

Publiczna prezentacja projektu

WROTA STATYSTYKI

Janusz Dygaszewicz

Dyrektor Departamentu Systemów Teleinformatycznych, Geostatystyki i Spisów

Agenda

1. Podstawowe informacje o projekcie
2. Interesariusze projektu
3. Zdiagnozowane problemy i potrzeby
4. Cel główny i cele szczegółowe projektu
5. Cele projektu wyrażone mierzalnymi wskaźnikami
6. Korzyści projektu
7. Usprawnienie procesów w oparciu o Model Procesu Produkcji Statystycznej (MPPS)
8. Schemat kooperacji Systemu Wrota Statystyki z aplikacjami zewnętrznymi
9. Schemat kooperacji systemów (aplikacji) w ramach Wrót Statystyki
10. Produkty i podprodukty projektu
11. Komplementarność z innymi działaniami
12. Kamienie milowe
13. Harmonogram zamówień publicznych i koszty

Podstawowe informacje o projekcie

Okres realizacji:

1 grudnia 2019 r. – 30 listopada 2022 r.

Koszt realizacji:

41 243 705,00 zł

Źródło finansowania:

Działanie 2.3 Cyfrowa dostępność i użyteczność informacji sektora publicznego Poddziałanie 2.3.1 „Cyfrowe udostępnienie informacji sektora publicznego ze źródeł administracyjnych i zasobów nauki” (Typ I projektu: Cyfrowe udostępnienie ISP ze źródeł administracyjnych) w ramach II Osi priorytetowej – „E-administracja i otwarty rząd” Programu Operacyjnego Polska Cyfrowa

Interesariusze projektu

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
Osoby indywidualne	Dostęp do pełnych, zintegrowanych zbiorów danych wynikowych i ich analiza, mimo że znajdują się w zasobach statystyki, jest utrudniony.	ok. 150 000 osób (1% z ok. 15 mln osób)
Podmioty gospodarcze (strefa biznesowa)		ok. 50 000 podmiotów gospodarczych (2.5% z ok. 2 mln)
Klienci instytucjonalni (strefa publiczna)		ok. 10 tys. klientów instytucjonalnych (1.5% z ok. 640 000)

Interesariusze projektu (cd.)

Interesariusz	Zidentyfikowany problem	Szacowana wielkość grupy
Statystyka publiczna	<p>Brak systemu do oceny procesu produkcji statystycznej na poszczególnych etapach, który wspierany byłby wydajnym systemem metainformacji.</p> <p>Informacja na temat potrzeb zgłaszanych przez użytkowników trafia do statystyki różnymi kanałami - jest rozproszona i trudno zarządzalna. Brakuje jednego miejsca do przechowywania wszystkich pierwotnych potrzeb informacyjnych.</p>	ok. 6 tys. osób

Zdiagnozowane problemy i potrzeby

Problem	Potrzeba interesariusza
Brak systemu do oceny procesu produkcji statystycznej na poszczególnych etapach, który wspierany byłby wydajnym systemem metainformacji.	Konieczne jest skoncentrowanie i zintensyfikowanie działań nad podniesieniem jakości i wydajności w obszarze wytwarzania, zarządzania i promocji produktów statystycznych oraz zarządzania wsparciem użytkowników.
Informacja na temat potrzeb zgłaszanych przez użytkowników trafia do statystyki różnymi kanałami - jest rozproszona i trudno zarządzalna. Brakuje jednego miejsca do przechowywania wszystkich pierwotnych potrzeb informacyjnych.	Pojawia się coraz większe zapotrzebowanie na udostępnianie informacji wynikowych, w wersji mobilnej, z uwzględnieniem profilu użytkownika.

Zdiagnozowane problemy i potrzeby (cd.)

Problem	Potrzeba interesariusza
<p>Dane są w niewielkim stopniu współdzielone. Przeważa wykorzystywanie danych pozyskiwanych bezpośrednio od respondentów. Dane z zewnętrznych systemów w dużej części są pozyskiwane na nośnikach elektronicznych. Dostęp do pełnych, zintegrowanych zbiorów danych wynikowych i ich analiza, mimo że znajdują się w zasobach statystyki, jest utrudniony.</p>	<p>Udostępnianie użytkownikom, także w wersji mobilnej, spójnych, zintegrowanych, scentralizowanych zbiorów, pozwalających na zaawansowaną analizę danego zjawiska opartego się na dobrze funkcjonującym całym systemie produkcji, który obecnie wymaga znaczącej modernizacji w zakresie mocno rozproszonej infrastruktury.</p>

Cel projektu

CEL GŁÓWNY:

Poszerzenie zakresu i usprawnienie udostępniania informacji statystycznych oraz podniesienie ich użyteczności. Cel ten przyczyni się do realizacji Celu szczegółowego 4. POPC Cyfrowa dostępność i użyteczność informacji sektora publicznego. Efektem projektu będzie zwiększenie dostępności oraz poprawa jakości informacji sektora publicznego (wynikowych informacji statystycznych).

Cel projektu

CELE SZCZEGÓŁOWE:

- 1) Poprawa dostępności wynikowych informacji statystycznych dzięki stworzeniu nowych kanałów udostępniania danych.
- 2) Podniesienie jakości udostępnionych danych poprzez umożliwienie kontroli procesu ich produkcji oraz publikację metadanych.
- 3) Zwiększenie otwartości danych poprzez dostosowanie do formatów umożliwiających odczyt maszynowy w celu ułatwienia ich ponownego wykorzystania.

Cele projektu wyrażone mierzalnymi wskaźnikami

Cel nr	Wskaźnik projektu	j.m.	Wartość docelowa
Wskaźniki rezultatu			
1	Liczba pobrań/odtworzeń dokumentów zawierających informacje sektora publicznego	szt.	1 000 000
Wskaźniki produktu			
1	Liczba podmiotów, które udostępniły on-line informacje sektora publicznego	szt.	1
1	Liczba udostępnionych on-line dokumentów zawierających informacje sektora publicznego	szt.	80 000
2; 3	Liczba utworzonych API	szt.	3
2; 3	Liczba baz danych udostępnionych on-line poprzez API	szt.	2

Korzyści projektu

- Najważniejszą korzyścią, jaką niesie za sobą realizacja niniejszego projektu, jest **usprawnienie udostępniania informacji statystycznych oraz podniesienie ich użyteczności**.
- Wprowadzone w ramach projektu rozwiązania przyczynią się do ułatwienia prowadzenia działalności gospodarczej, m.in. poprzez **likwidację barier na styku biznes-administracja**, a także poprzez **zwiększenie możliwości powtórnego wykorzystania zasobów informacji publicznej** przez przedsiębiorców, obywateli, a także administrację rządową i samorządową.

Korzyści projektu

- Konstruowany model organizacji badań statystycznych zapewni statystyce publicznej w długim czasie dużą elastyczność organizacji i działania, **co wzmocni system informacyjny państwa.**
- Produkty powstałe w wyniku realizacji projektu przyczynią się do **wzmocnienia procesów decyzyjnych w państwie**, co wymaga wprowadzenia systemów wsparcia realizacji wielu procesów statystycznych. Przełoży się to również na produktywność, efektywność organizacji oraz poprawę jej zarządzania.
- Wdrożenie Projektu wpłynie na zmianę funkcjonowania samej organizacji.

Usprawnienie procesów

- W projekcie zakłada się budowę lub modernizację systemów / aplikacji / funkcjonalności w celu udostępniania zdigitalizowanych informacji sektora publicznego.
- Dla podniesienia jakości i zakresu udostępnianych informacji systemy zostaną dostosowane do Modelu Procesu Produkcji Statystycznej poprzez utworzenie grup systemów odpowiadających fazom produkcji statystycznej i przypisanie im nowych funkcjonalności związanych z udostępnianiem informacji.

Usprawnienie procesów

- Wdrożenie podejścia procesowego poprzez usprawnienie procesów wg. MPPS (Model Procesu Produkcji Statystycznej) obejmującego 8 faz:
 - 1) specyfikacja potrzeb,
 - 2) projektowanie,
 - 3) budowanie,
 - 4) zbieranie danych,
 - 5) przetwarzanie,
 - 6) analiza,
 - 7) udostępnianie,
 - 8) ocena.
- Odejście od obecnego „modelu rurkowego” realizowanego przez ponad 200 niezależnych aplikacji działających w różnych środowiskach sprzętowo-programowych.

Model Procesu Produkcji Statystycznej - MPPS

1 Specyfikacja potrzeb	2 Projektowanie	3 Budowa	4 Zbieranie danych	5 Przetwarzanie	6 Analiza	7 Udostępnianie	8 Ocena
1.1 Identyfikacja potrzeb	2.1 Projektowanie produktów wyjściowych	3.1 Ponowne użycie lub budowa narzędzi zbierania danych	4.1 Utworzenie wykazy i wybór próby	5.1 Integracja danych	6.1 Opracowanie wyników wstępnych	7.1 Uaktualnienie systemów udostępniania	8.1 Zbieranie ocen
1.1a Sformułowanie programu badań statystycznych	2.2 Projektowanie opisów zmiennych	3.2 Ponowne użycie lub budowa komponentów przetwarzania i analizy	4.1a Geokodowanie wykazu i próby losowej	5.2 Klasyfikacja i kodowanie	6.2 Walidacja wyników	7.2 Utworzenie produktów udostępniania	8.2 Przeprowadzenie oceny
1.2 Konsultacje i potwierdzanie potrzeb	2.3 Projektowanie zbierania danych	3.3 Ponowne użycie lub budowa komponentów udostępniania	4.2 Konfigurowanie zbierania danych	5.3 Przegląd i walidacja	6.2a Opracowanie analiz przestrzennych i map	7.2a Zarządzanie analizami przestrzennymi i mapami przy pomocy GIS	8.3 Przygotowanie planu naprawy
1.3 Określanie zestawu danych wynikowych	2.4 Projektowanie wykazu i próby losowej	3.4 Konfiguracja procesów pracy	4.3 Zbieranie danych	5.4 Edycja i imputacja	6.3 Przegląd i objaśnienia danych wynikowych	7.3 Zarządzanie rozpowszechnianiem udostępnianych produktów	
1.4 Identyfikacja pojęć	2.5 Projektowanie przetwarzania i analiz	3.5 Testowanie systemu produkcyjnego	4.3a Geokodowanie zebranych danych	5.5 Wyprowadzanie nowych zmiennych i jednostek	6.4 Nałożenie tajemnicy statystycznej	7.4 Promocja udostępnianych produktów	
1.5 Sprawdzenie dostępności danych	2.5a Projektowanie geokodowania wykazu, próby oraz zbieranych danych	3.6 Testowanie procesu produkcji statystycznej	4.4 Zakończenie zbierania danych	5.6 Wyliczenie wag	6.5 Zakończenie opracowania produktów wynikowych	7.5 Zarządzanie wsparciem użytkowników	
1.6 Przygotowanie i zgłoszenie uzasadnienia biznesowego	2.6 Projektowanie systemu produkcyjnego i procesów pracy	3.7 Zakończenie budowy systemu produkcyjnego	4.5 Weryfikacja planów	5.7 Wyliczenie agregatów		7.6 Badanie satysfakcji	
1.7 Opracowanie planu ogólnego realizacji procesu produkcji	2.7 Projektowanie wskaźników jakościowych	3.8 Weryfikacja projektu i planów budowy		5.8 Utworzenie końcowych zbiorów danych		7.7 Przygotowanie planu naprawczego dla kolejnego cyklu badań	
1.8 Opracowanie szczegółowych planów dla kolejnych faz	2.8 Weryfikacja planu ogólnego i planów szczegółowych						
1.9 Weryfikacja koncepcji w celu poprawy satysfakcji							

Model Procesu Produkcji Statystycznej - MPPS

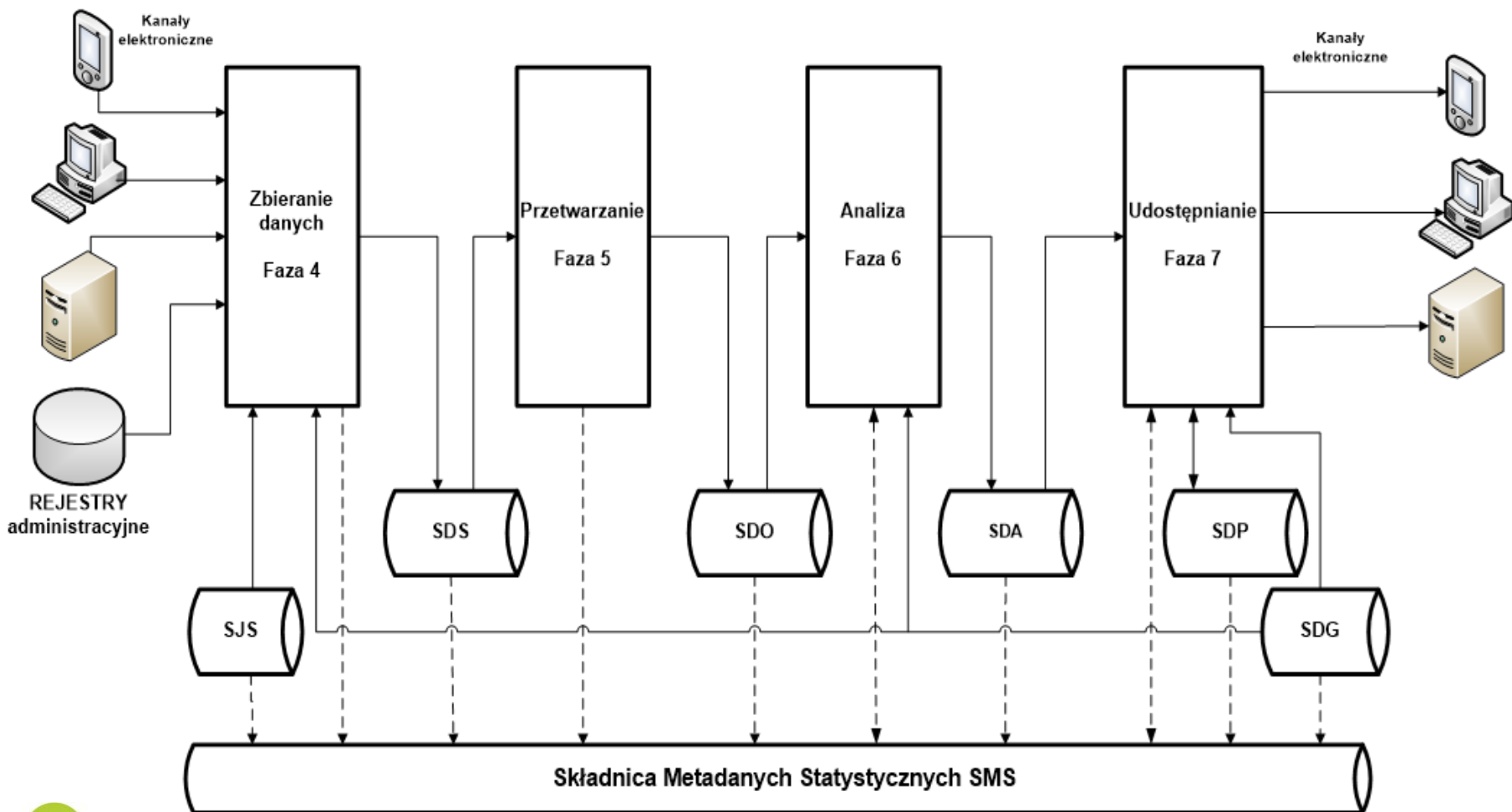
Na podstawie modelu MPPS stworzono ramy architektoniczne będące podstawą budowy oraz unormowanego rozwoju rozwiązań informatycznych wykorzystywanych do realizacji zadań statystyki publicznej. Podejście zgodne z architekturą korporacyjną ma na celu zapewnienie trwałości, spójności oraz obniżenie kosztów rozwiązania.

Fazy i składnice danych według modelu MPPS

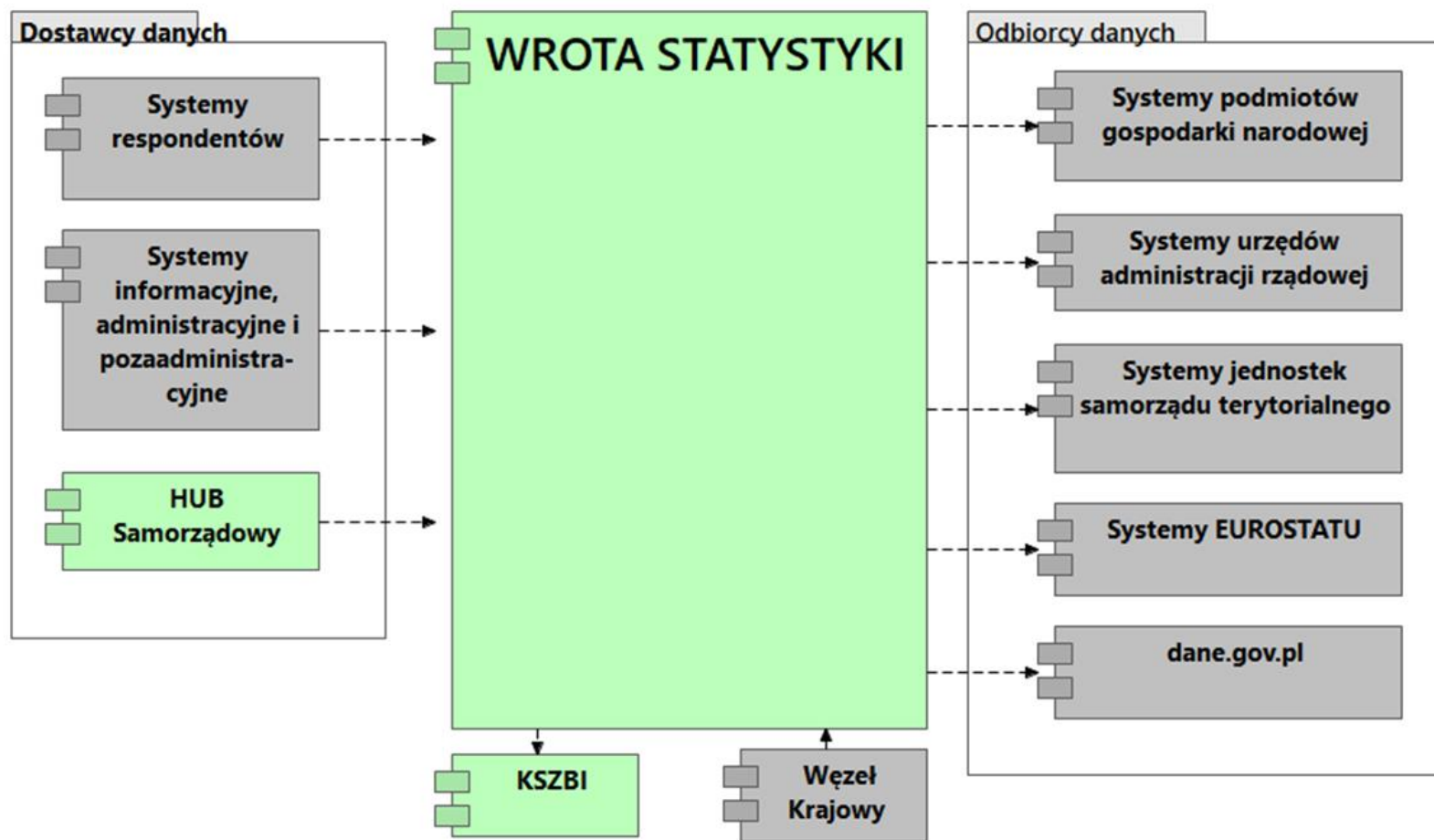
Strefa publiczna

Statystyka publiczna – platforma wsparcia informatycznego

Strefa publiczna



Schemat kooperacji Systemu Wrota Statystyki z aplikacjami zewnętrznymi



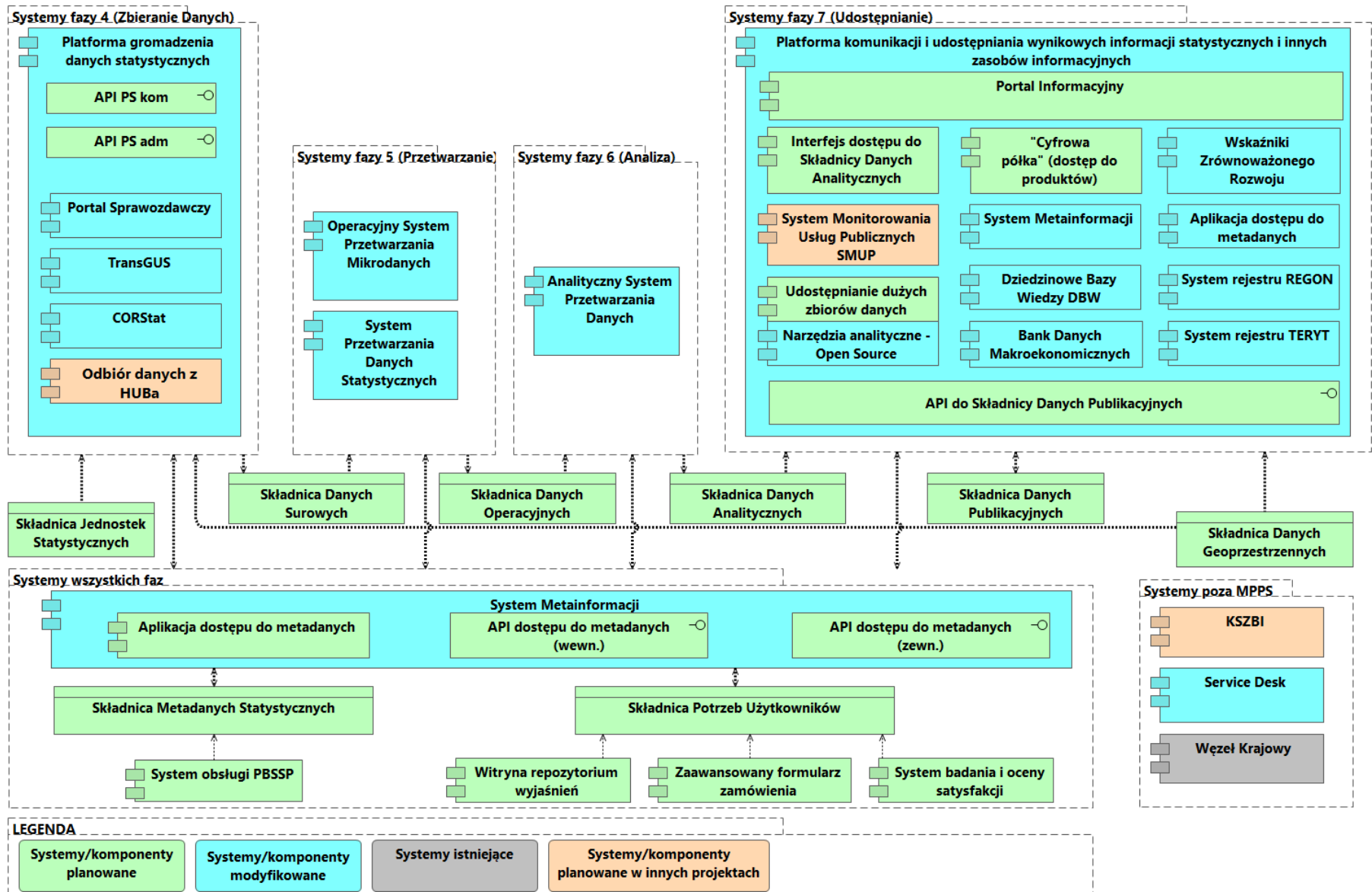
LEGENDA

Systemy modyfikowane

Systemy planowane

Systemy istniejące

Schemat kooperacji systemów (aplikacji) w ramach Wrót Statystyki



Produkty projektu

W ramach projektu przewidziano do wdrożenia następujące nowe lub modernizowane produkty projektu:

- Bazy danych
- Podsystemy aplikacyjne Wrót Statystyki
- Produkty infrastrukturalne

Bazy danych



Składnica Jednostek Statystycznych – baza niezbędnych danych identyfikacyjnych, która pełni rolę operatu do badań statystycznych, obejmująca zarówno całą populację obiektów poddawanych obserwacji statystycznej, jak i celowo lub stochastycznie przygotowane podzbiory populacji.



Składnica Danych Surowych – struktura bazodanowa pozwalająca na przechowywanie i udostępnianie do dalszej obróbki nieprzetworzonych statystycznie danych, zebranych zarówno pierwotnie (od respondentów), jak i wtórnie pozyskanych z rejestrów administracyjnych oraz innych źródeł.



Składnica Danych Operacyjnych – struktura bazodanowa przechowująca przetworzone statystycznie dane jednostkowe. Wystandardyzowane dane w SDO podlegają niezbędnej integracji, pseudonimizacji lub anonimizacji oraz agregacji w celu ich przygotowania do analiz statystycznych.



Składnica Danych Analitycznych – struktura bazodanowa przechowująca i udostępniająca odpersonalizowane (nieidentyfikowalne) dane jednostkowe, mikro- i makroagregaty na potrzeby wykonywania analiz statystycznych (z uwzględnieniem analiz przestrzennych) m.in. w postaci tabel i kostek (hypercubes).

Bazy danych



Składnica Danych Geoprzestrzennych – baza zawierająca dane geometryczne obiektów geoprzestrzennych wraz z ich lokalizacją przestrzenną, z dokładnością do współrzędnych x,y. Obiekty bazy pozwalają na geokodowanie (powiązanie z lokalizacją przestrzenną) zarówno punktów adresowych, jak i jednostek podziału administracyjnego lub statystycznego kraju.



Składnica Potrzeb Użytkowników – struktura bazodanowa, która przechowuje i udostępnia zebrane różnymi kanałami komunikacyjnymi dane o istniejących i oczekiwanych potrzebach informacyjnych użytkowników statystyki publicznej.



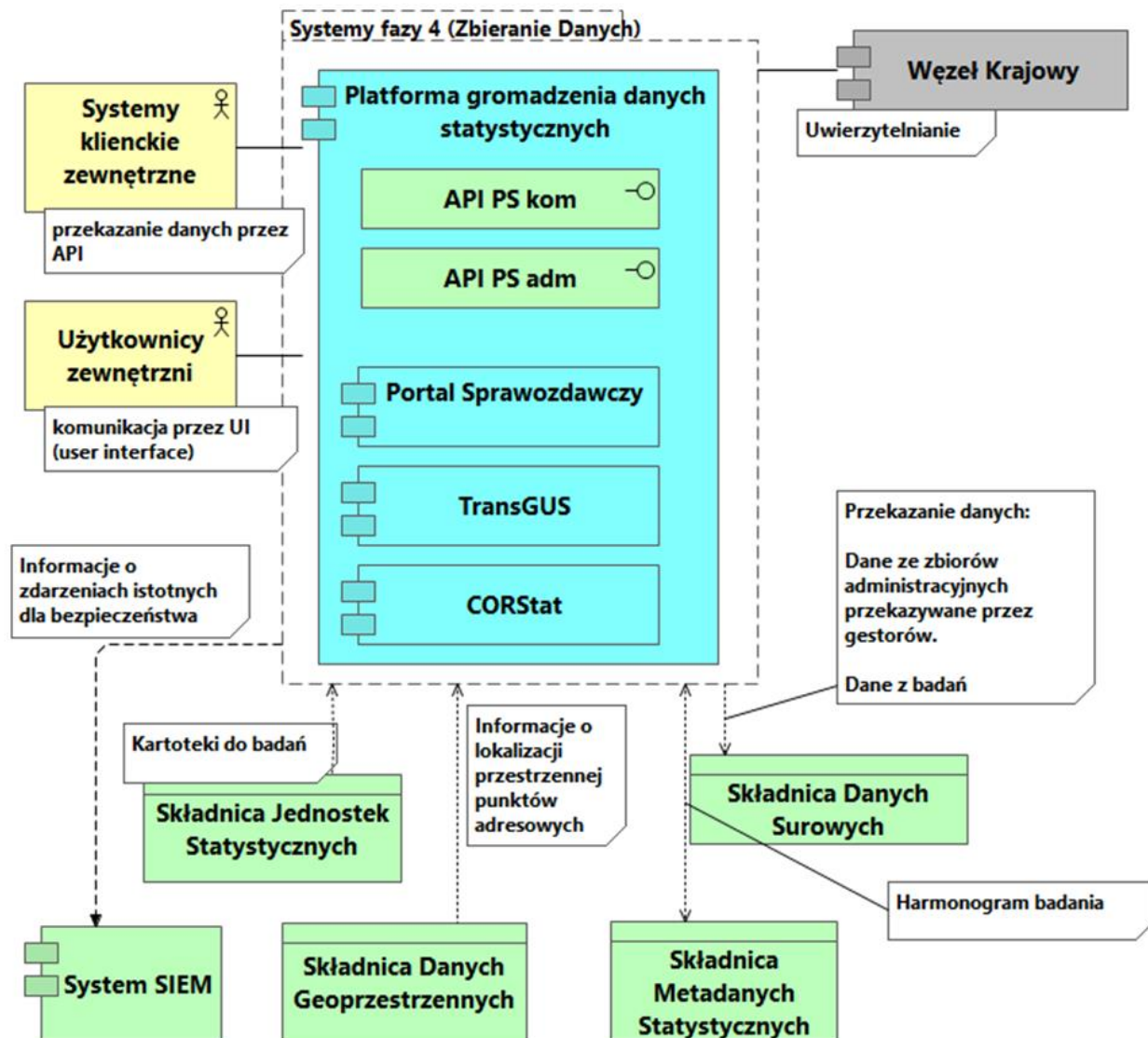
Składnica Danych Publikacyjnych – struktura bazodanowa umożliwiająca przechowywanie gotowych do publicznego udostępniania produktów końcowych procesu produkcji statystycznej w postaci tabel, kostek wielowymiarowych, map oraz rozwiązań bazodanowych dostosowanych do elektronicznych lub tradycyjnych kanałów udostępniania wynikowych informacji statystycznych.



Składnica Metadanych Statystycznych – baza wiedzy o wszystkich zasobach i procesach związanych z produkcją danych statystycznych. Jest niezbędna do monitorowania i zarządzania procesem produkcji statystycznej.

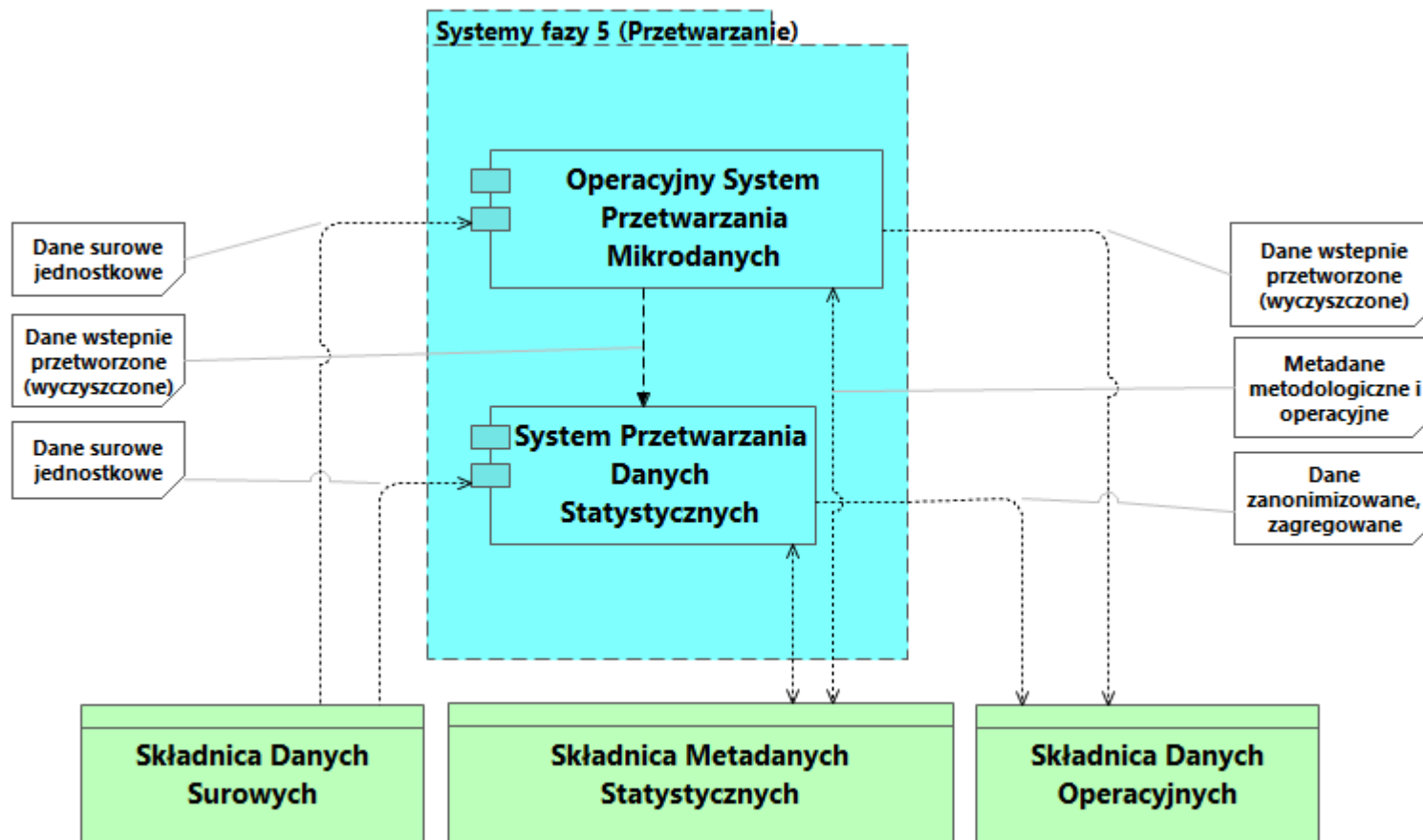
Podsystemy aplikacyjne Wrót Statystyki

- Platforma Gromadzenia Danych Statystycznych



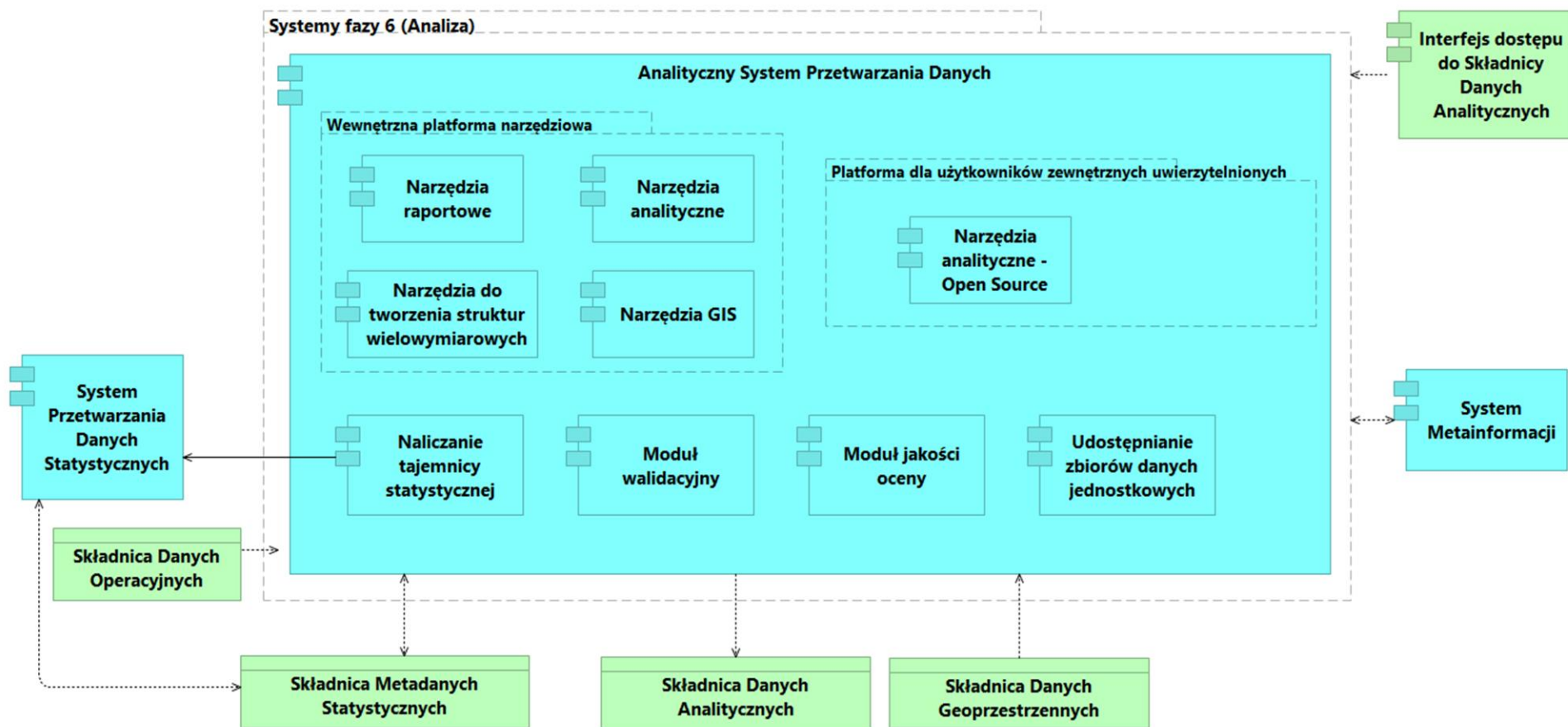
Podsystemy aplikacyjne Wrót Statystyki

- Operacyjny System Przetwarzania Mikrodanych



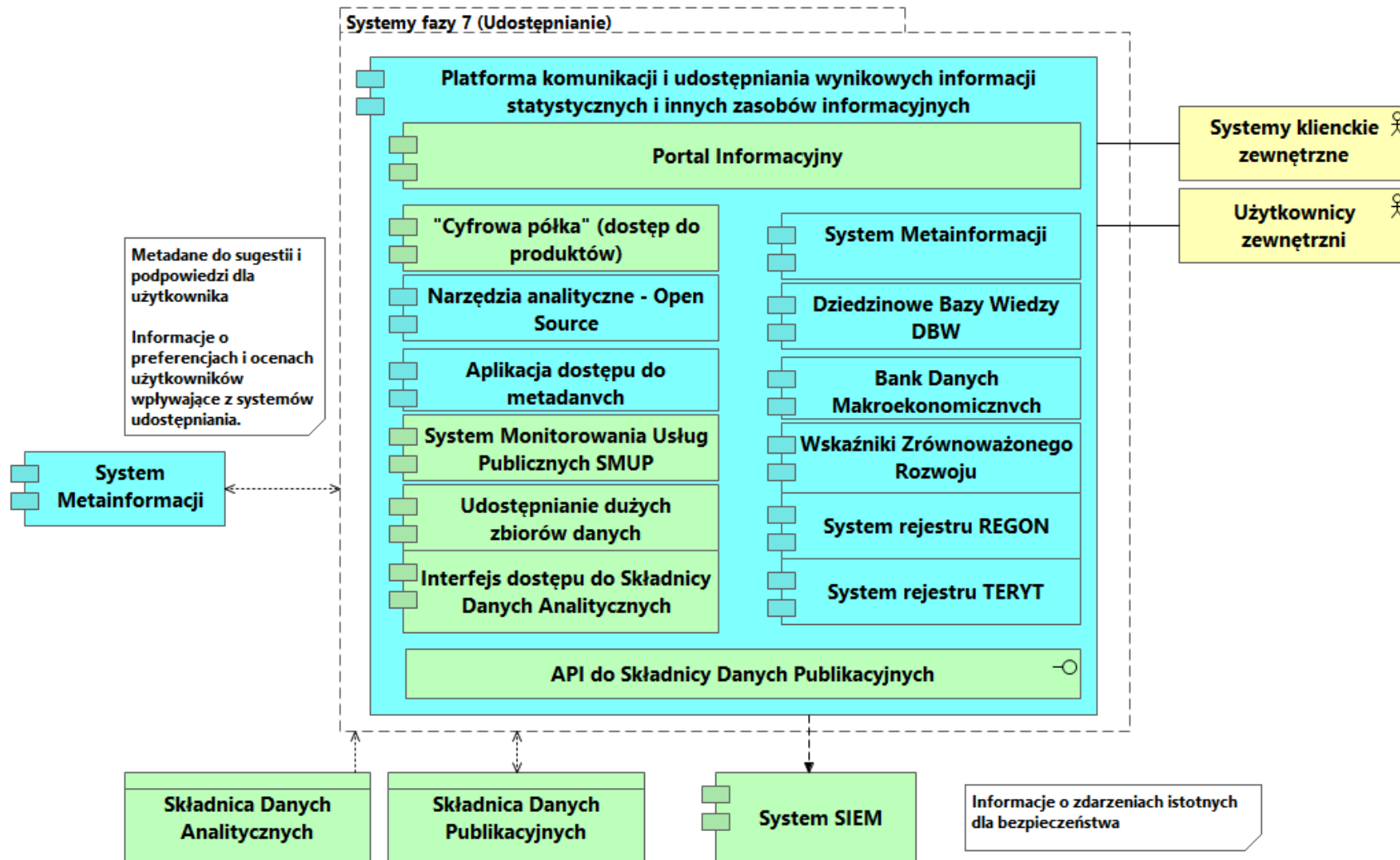
Podsystemy aplikacyjne Wrót Statystyki

- Analityczny System Przetwarzania Danych



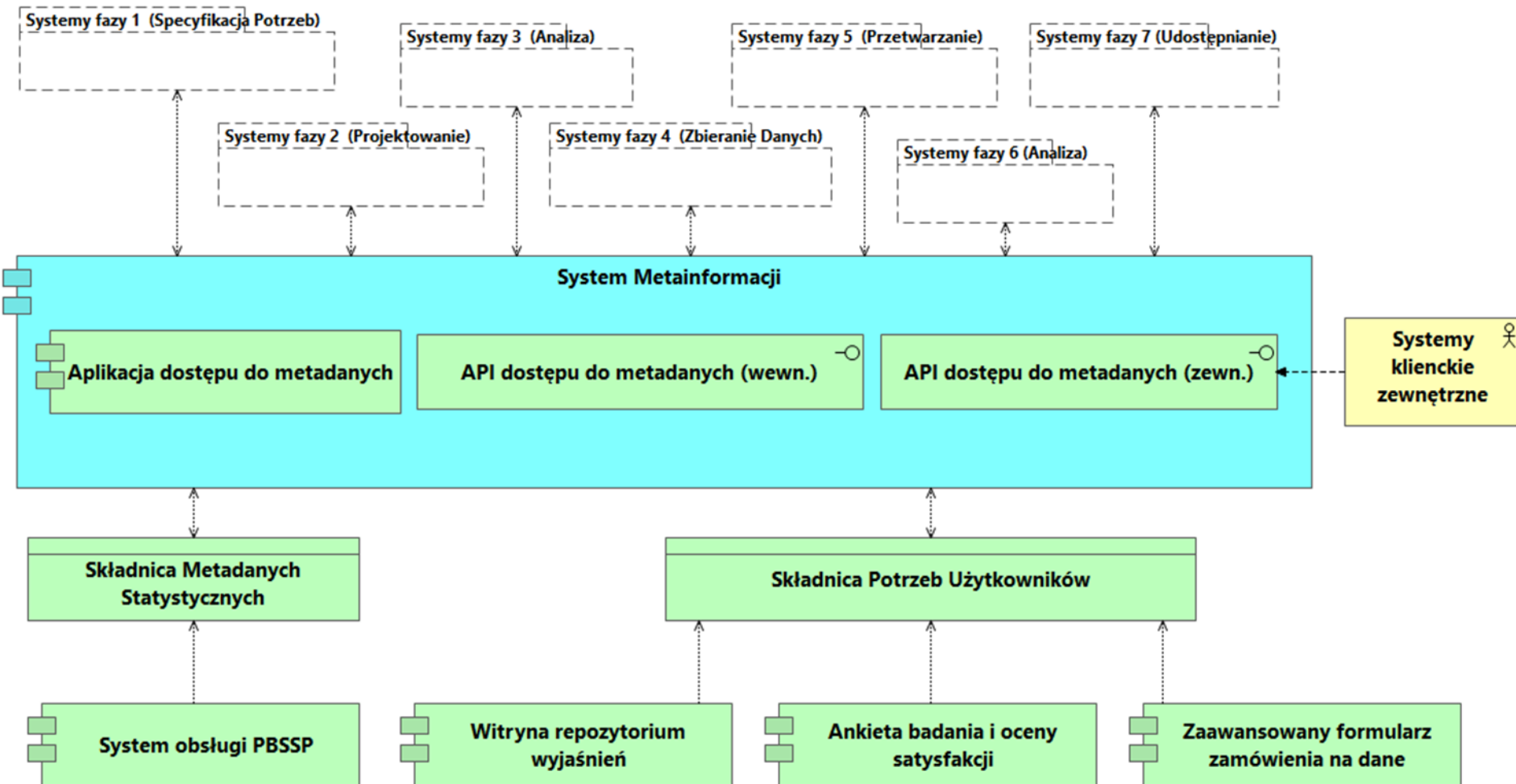
Podsystemy aplikacyjne Wrót Statystyki

- Platforma Komunikacji i Udostępniania Wynikowych Informacji Statystycznych i Innych Zasobów Informacyjnych



Podsystemy aplikacyjne Wrót Statystyki

- System Metainformacji



Produkty infrastrukturalne

System backupowy – modernizacja środowiska sprzętowo-systemowego CPD dla składnic danych i informacji. Dostawa infrastruktury zasobowej, oprogramowania backupowego oraz bibliotek taśmowych do składowania backupów.

System zarządzający i monitorujący sieć WAN – modernizacja systemu akceleracji ruchu w sieci WAN (urządzenia do akceleracji ruchu w sieci WAN i modernizacja punktów dostępowych) w CPD i LAN.

System Serwis Desk – rozbudowa systemu Serwis Desk polegająca na rozszerzeniu zakresu terytorialnego świadczonych usług na cały resort statystyki.

System zarządzający dostępem do sieci – rozbudowa sieci bezprzewodowej oraz modernizacja systemów administracyjnych zapewniających bezpieczny dostęp do sieci; centralny punkt kontroli dostępu do sieci zapewni zwiększenie poziomu bezpieczeństwa poprzez każdorazowe uwierzytelnienie klienta próbującego uzyskać dostęp do zasobów informatycznych.

Modernizacja stanowisk pracy – zakup sprzętu i oprogramowania planowany w związku z wdrożeniem zaplanowanych zadań, udostępnianiem ISP, a także działań uzupełniających w postaci promocji.

Komplementarność z innymi działaniami

- **Węzeł krajowy** – w zakresie dostępu do środowiska integracyjnego Węzła Krajowego na potrzeby realizacji integracji środków identyfikacji wykorzystywanych do uwierzytelniania użytkowników usług elektronicznych
- **Projekt SMUP** (*System Monitorowania Usług Publicznych*) – jedno ze źródeł zasilających Platformę Gromadzenia Danych Statystycznych oraz upublicznianych w zintegrowanej postaci na Platformie Udostępniania
- **Projekt KSZBI** (*Kompleksowy System Zarządzania Bezpieczeństwem Informacji*) – w zakresie objęcia projektu mechanizmami bezpieczeństwa na poziomie organizacyjnym i technologicznym

Komplementarność z innymi działaniami (cd.)

- **Projekt PDS** (*Przestrzenne Dane Statystyczne w Systemie Informacyjnym Państwa*) – wykorzystanie bazy danych geoprzestrzennych, utworzonej w ramach projektu PDS, stanowiącej podstawę Składnicy Danych Geoprzestrzennych (SDG)
- **Otwarte Dane i Otwarte Dane+** – BDL oraz DBW jako element planowanej do budowy platformy udostępniania wynikowych informacji
- **Projekt SISP-2** - wykorzystanie elementów infrastruktury techniczno-systemowej zakupionej w ramach innych projektów
- **Spisy powszechne** – przy realizacji projektu wykorzystana zostanie infrastruktura wirtualizacyjna dostosowana do Modelu Procesu Produkcji Statystycznej

Kamienie milowe

Kamienie milowe	Data osiągnięcia kamienia
Wyłonienie wykonawcy usługi wsparcia informatycznego	2020-04-30
Opracowanie projektu technicznego Systemu Metainformacji	2020-06-30
Opracowanie koncepcji modernizacji Platformy Udostępniania Wynikowych Informacji Statystycznych i Innych Zasobów Informacyjnych	2020-08-31
Opracowanie projektów technicznych: - Platformy Gromadzenia Danych Statystycznych - Analitycznego Systemu Przetwarzania Danych	2020-09-30
Opracowanie prototypu Platformy Gromadzenia Danych Statystycznych	2021-06-30
Opracowanie prototypu Analitycznego Systemu Przetwarzania Danych	2021-09-30

Kamienie milowe (cd.)

Kamienie milowe	Data osiągnięcia kamienia
Opracowanie prototypu Systemu Metainformacji	2021-10-30
Odbiór Portali Informacyjnych	2022-01-31
Przeprowadzenie testów akceptacyjnych i podpisanie protokołów odbioru poszczególnych komponentów systemu: <ul style="list-style-type: none">- System Metainformacji- Platforma Gromadzenia Danych Statystycznych- Analityczny System Przetwarzania Danych- Platforma Komunikacji i Udostępniania Wynikowych Informacji Statystycznych i Innych Zasobów Informacyjnych	2022-03-31
Testy akceptacyjne, ewentualne modyfikacje i odbiór systemu WROTA STATYSTYKI	2022-11-30

Harmonogram zamówień publicznych i ich koszty

	Zakres zamówienia	Termin ogłoszenia	Koszt (zł brutto)
1	Inżynier Projektu	Grudzień 2019	1 312 320,00
2	Wsparcie ds. IT	Grudzień 2019	9 000 000,00
3	Zakup sprzętu i oprogramowania dla: <ul style="list-style-type: none"> • Modernizacji środowiska sprzętowo-systemowego CPD dla składnic danych i informacji • Modernizacji systemu akceleracji ruchu w sieci WAN • Rozbudowy sieci bezprzewodowej oraz modernizacja systemów administracyjnych • Serwis Desk • Modernizacji stanowisk pracy w Urzędach Statystycznych 	Czerwiec 2020	16 052 000,00

Harmonogram zamówień publicznych i ich koszty (cd.)

	Zakres zamówienia	Termin ogłoszenia	Koszt (zł brutto)
4	Promocja projektu	Styczeń 2020 Czerwiec 2022	500 000,00
5	Budowa Platformy Komunikacji i Udostępniania Wynikowych Informacji Statystycznych i innych zasobów informacyjnych	Czerwiec 2020	1 000 000,00
6	Szkolenia	Marzec 2021	928 805,00

Dziękuję za uwagę