**OPISU ZAŁOŻEŃ PROJEKTU INFORMATYCZNEGO**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tytuł projektu** | WROTA STATYSTYKI | | |
| **Wnioskodawca** | Kancelaria Prezesa Rady Ministrów | | |
| **Beneficjent** | Główny Urząd Statystyczny | | |
| **Partnerzy** | - | | |
| **Źródło finansowania** | Program Operacyjny Polska Cyfrowa II oś priorytetowa E-administracja i otwarty rząd 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). | | |
| **Całkowity koszt projektu** | 41 243 705,00 zł | | |
| **Planowany okres realizacji projektu** | 01.04.2019 do 31.03.2022 | | |
| **Osoba kontaktowa** | Janusz Dygaszewicz | j.dygaszewicz@stat.gov.pl | 22 608-33-41 |

# POWODY PODJĘCIA PROJEKTU

## Identyfikacja problemu i potrzeb

Misją statystyki publicznej jest dostarczanie wiarygodnych, rzetelnych, niezależnych i wysokiej jakości informacji statystycznych. W celu jej realizacji niezbędna jest znajomość potrzeb odbiorców danych statystycznych, oczekujących szybszych, wiarygodniejszych informacji. Ważna dla nich jest synteza informacji, interpretacja danych, a także ich międzynarodowa porównywalność.

Konieczne jest skoncentrowanie i zintensyfikowanie działań nad podniesieniem jakości i wydajności w obszarze wytwarzania, zarządzania i promocji produktów statystycznych oraz zarządzania wsparciem użytkowników. Obecny proces produkcji statystycznej wymaga udoskonalenia. Nie ma systemu oceniającego proces na poszczególnych etapach produkcji, który wspierany byłby wydajnym systemem metainformacji.

Informacja na temat potrzeb zgłaszanych przez użytkowników trafia do statystyki różnymi kanałami -  **rozproszona i trudno zarządzalna. Brakuje jednego miejsca do przechowywania wszystkich pierwotnych potrzeb informacyjnych.** 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. Pojawia się coraz większe zapotrzebowanie na udostępnianie informacji wynikowych, w wersji mobilnej, z uwzględnieniem profilu użytkownika. Udostępnianie użytkownikom, także w wersji mobilnej, spójnych, zintegrowanych, scentralizowanych zbiorów, pozwalających na zaawansowaną analizę danego zjawiska musi opierać się na dobrze funkcjonującym całym systemie produkcji, który obecnie wymaga znaczącej modernizacji w zakresie mocno rozproszonej infrastruktury.

|  |  |  |
| --- | --- | --- |
| **Interesariusz** | **Zidentyfikowany problem** | **Szacowana wielkość grupy** |
| Obywatele | Oczekiwanie szybciej udostępnianych, w tym w wersji mobilnej, uwzględniającej profil użytkownika, wiarygodniejszych danych; z uwzględnieniem udostępniania w formacie otwartych danych publicznych | ok. 150 000 osób (1% z ok. 15 mln osób) |
| Przedsiębiorcy | ok. 50 000 podmiotów gospodarczych (2.5% z ok. 2 mln) |
| Instytucje publiczne | ok. 10 tys. klientów instytucjonalnych (1.5% z ok. 640 000) |
| Pracownicy statystyki publicznej | Brak jednolitego, zestandaryzowanego modelu realizacji produkcji statystycznej | ok. 6 tys. osób |

## Opis stanu obecnego

Pierwszym etapem procesu produkcji statystycznej jest rozpoznanie potrzeb użytkowników, które obecnie wymaga zestandaryzowania działań.

Badania statystyczne są realizowane w oparciu o odrębnie pozyskane i przetworzone dane o określonym zakresie informacyjnym. Sprzyja to redundancji zbieranych informacji.

Pozyskane dane w formie elektronicznej, z wykorzystaniem dedykowanych systemów i  nośników danych są przetwarzane w odrębnych dla każdego opracowania systemach informatycznych, tworzonych przez różne zespoły informatyków, na podstawie autorskich założeń. Współużywalność rozwiązań informatycznych jest niewielka. Przetworzone dane są gromadzone w wielu, odseparowanych bazach danych. Wynikowe informacje są gromadzone w niezależnych systemach/bazach, zasilanych z różnych źródeł.

# EFEKTY PROJEKTU

## Cele i korzyści wynikające z projektu

|  |  |
| --- | --- |
| **Cel - 1** | Poprawa dostępności wynikowych informacji statystycznych dzięki stworzeniu nowych kanałów udostępniania danych i modernizacji obecnych. |
| **Cel strategiczny** | * sprawne zarządzanie informacją publiczną, w tym umożliwienie ponownego wykorzystywania publicznych zasobów informacyjnych, oraz dostępność dla jak najszerszych kręgów odbiorców, o których mowa w Programie Zintegrowanej Informatyzacji Państwa – kierunek interwencji Otwarty Rząd i wspieranie rozwoju społeczeństwa obywatelskiego; * Budowa i rozwój e-administracji – orientacja administracji państwa na usługi cyfrowe, działanie: Zwiększenie dostępu do informacji sektora publicznego, w tym poprzez interfejsy dla programistów, zgodnie ze Strategią na rzecz Odpowiedzialnego Rozwoju – obszar E‑państwo. |
| **Korzyść:** | W wyniku projektu użytkownikom zapewniony zostanie szeroki i łatwy dostęp do scentralizowanych, zintegrowanych danych publicznych, tj. oficjalnych danych statystycznych o sytuacji ekonomicznej, demograficznej, społecznej oraz środowisku naturalnym, organom władzy państwowej, administracji publicznej rządowej i samorządowej, sektorowi gospodarczemu i społeczeństwu. Dane te są istotne z punktu widzenia całej gospodarki, a także kluczowe dla procesów monitorowania rozwoju społecznego i gospodarczego. W wyniku projektu możliwe będzie zapewnienie odbiorcom dostępu do wynikowych informacji statystycznych z wykorzystaniem nowoczesnych kanałów dostępu i komunikacji, w oparciu o jednolitą, uspójnioną platformę udostępniania wykorzystującą narzędzia open-source (z upublicznionym kodem źródłowym) i API, umożliwiającą prowadzenie samodzielnych analiz i zestawień. |
| **KPI:** | 1. Liczba pobrań/odtworzeń dokumentów zawierających informacje sektora publicznego 2. Liczba podmiotów, które udostępniły on-line informacje sektora publicznego 3. Liczba udostępnionych on-line dokumentów zawierających informacje sektora publicznego |
| **Wartość aktualna i docelowa KPI:** | 1. Aktualnie:0 ; docelowo: 1 000 000  2. Aktualnie:0 szt., docelowo: 1 szt.  3. Aktualnie: 0 szt., docelowo: 80 000 szt. |
| **Metoda pomiaru KPI** | 1/ Pomiar KPI będzie następował dzięki narzędziom na stronach internetowych  2/ Raport z testów końcowych |
| **Cel - 2** | Podniesienie jakości udostępnionych danych poprzez umożliwienie kontroli procesu ich produkcji oraz publikację metadanych. |
| **Cel strategiczny** | * Budowa i rozwój e-administracji – orientacja administracji państwa na usługi cyfrowe, zgodnie ze Strategią na rzecz Odpowiedzialnego Rozwoju – obszar E‑państwo; * Poprawa jakości usług świadczonych przez instytucje publiczne (m.in. uproszczenie procedowania spraw drogą elektroniczną, eliminacja barier prawnych, maksymalizacja automatyzacji procesów), upowszechnianie najlepszych praktyk w zakresie obsługi biznesu, oraz zastosowanie instrumentów analitycznych służących identyfikacji potrzeb obywateli w ramach e-administracji oraz wspieraniu działań podnoszących aktualność i dostępność danych znajdujących się w rejestrach państwowych, zgodnie ze Strategią na rzecz Odpowiedzialnego Rozwoju – obszar: Instytucje prorozwojowe i strategiczne zarządzanie rozwojem. |
| **Korzyść:** | W wyniku projektu możliwe będzie zapewnienie odbiorcom szerszego dostępu do metadanych, pozwalających na lepsze wykorzystanie wynikowych informacji statystycznych. Zwiększenie poziomu świadomości pracowników poprzez uczestnictwo w szkoleniach, wpłynie na poprawę jakości udostępnianych informacji wynikowych. Odbiorcy będą mieli dostęp do Systemu Metainformacji, w którym będą znajdowały się publiczne metadane dotyczące całego procesu produkcji, w tym statystycznych metod badawczych, sposobu interpretacji i wykorzystania wyników badań statystycznych. Pozwoli to użytkownikom na lepsze, pełniejsze wykorzystanie informacji wynikowych, zapoznanie się z metodologią badań i sposobem produkcji danych statycznych. |
| **KPI:** | 1. Liczba utworzonych API do Składnicy Metadanych Statystycznych  2. Liczba baz danych udostępnionych on-line poprzez API  3. Procent danych zawartych w Składnicy Danych Publikacyjnych opisanych metadanymi  4. Liczba pracowników służb statystycznych niebędących pracownikami IT objętych wsparciem szkoleniowym. |
| **Wartość aktualna i docelowa KPI:** | 1. Aktualnie: 0 szt., docelowo: 2 szt.  2. Aktualnie: 0 szt., docelowo: 1 szt.  3. Aktualnie: 0 %, docelowo: 80%  4. Aktualnie 0 osób; docelowo: 100 osób |
| **Metoda pomiaru KPI** | 1/ Pomiar KPI będzie następował dzięki narzędziom na stronach internetowych;  2/ Raport z testów końcowych  3/ Dokumentacja szkoleniowa |
| **Cel - 3** | Zwiększenie otwartości danych poprzez dostosowanie do formatów umożliwiających odczyt maszynowy w celu ułatwienia ich ponownego wykorzystania. |
| **Cel strategiczny** | * Stymulowanie innowacyjności poprzez wzrost efektywności wiedzy i pracy, działanie 2.6.4. Otwarcie dostępu do informacji publicznej dla przedsiębiorstw i obywateli, o którym mowa w Strategii Innowacyjności i Efektywności Gospodarki „Dynamiczna Polska 2020” – cel 2 . * Udoskonalenie kompetencji analitycznych, w tym zdolności do dokonywania analizy kosztów i korzyści interwencji publicznych, ekonometrycznej i statystycznej analizy danych, wykorzystania nowoczesnych technik przetwarzania informacji (big data), zgodnie ze Strategią na rzecz Odpowiedzialnego Rozwoju – obszar: Instytucje prorozwojowe i strategiczne zarządzanie rozwojem; * Program otwierania danych publicznych: cel szczegółowy 3: „zwiększenie przejrzystości i skuteczności funkcjonowania organów administracji publicznej, usprawnienie działalności urzędów, umożliwienie obywatelom udziału w sprawowaniu władzy i procesie ponownego wykorzystywania danych publicznych”. |
| **Korzyść:** | W wyniku realizacji projektu powstaną API do Składnicy Metadanych Statystycznych oraz do Składnicy Danych Publikacyjnych. Publikowanie danych w sposób otwarty, umożliwiający odczyt maszynowy, zapewni użytkownikom zautomatyzowane pobieranie danych statystycznych i metadanych co w znacznym stopniu usprawni zarzadzanie i podejmowanie decyzji w oparciu o informacje o aktualnej sytuacji ekonomicznej, demograficznej, społecznej oraz środowisku naturalnym. |
| **KPI:** | 1. Liczba utworzonych API  2. Liczba baz danych udostępnionych on-line poprzez API |
| **Wartość aktualna i docelowa KPI:** | 1. Aktualnie: 0 szt., docelowo: 2 szt.  2. Aktualnie: 0 szt., docelowo: 1 szt. |
| **Metoda pomiaru KPI** | 1/ Pomiar KPI będzie następował dzięki narzędziom na stronach internetowych  2/ Raport z testów końcowych |

## 

## Udostępnione e-usługi

| **Lp.** | **Nazwa  e-usługi** | **Typ** | **Zakres oddziaływania** | **Poziom dojrzałości  e-usługi** |
| --- | --- | --- | --- | --- |
|  | **Nie dotyczy** |  |  |  |

## Udostępnione informacje sektora publicznego i zdigitalizowane zasoby

Czy wszystkie zdigitalizowane zasoby objęte projektem będą udostępniane bezpłatnie? TAK/~~NIE~~ [[1]](#footnote-2)

|  |  |  |
| --- | --- | --- |
| **Rodzaj**  **informacji / zasobów** | **Planowana data udostępnienia** | **Szacowana liczba obiektów objętych digitalizacją (udostępnianiem informacji)** |
| Składnica danych publikacyjnych – w których zgromadzone są dane przygotowane do opublikowania poprzez Platformę udostępniania wynikowych informacji statystycznych i innych zasobów informacyjnych. | 2021-09-30 | 80 000 zmiennych zagregowanych w różnych układach. |
| Składnica Metadanych Statystycznych zawierająca metadane definicyjne, badań, procesowe i strukturalne powiązane z danymi znajdującymi się w Składnicy Danych Publikacyjnych | 2021-09-30 | 80 000 grup metadanych powiązanych z udostępnianymi danymi. |

## Produkty końcowe projektu

| **Nazwa produktu** | **Planowana data wdrożenia** |
| --- | --- |
| System zarządzający i monitorujący sieć | 2019-12-31 |
| System backupowy | 2019-12-31 |
| System zarządzający i monitorujący sieć WAN | 2020-06-30 |
| System Serwis Desk | 2020-06-30 |
| System zarządzający dostępem do sieci | 2020-12-31 |
| Portale Informacyjne | 2021-07-31 |
| Składnica Jednostek Statystycznych | 2021-09-30 |
| Składnica Danych Surowych | 2021-09-30 |
| Składnica Danych Operacyjnych | 2021-09-30 |
| Składnica Danych Analitycznych | 2021-09-30 |
| Składnica Danych Geoprzestrzennych | 2021-09-30 |
| Składnica Potrzeb Użytkowników | 2021-09-30 |
| System Metainformacji | 2021-09-30 |
| Platforma Gromadzenia Danych Statystycznych | 2021-09-30 |
| Analityczny System Przetwarzania Danych | 2021-09-30 |
| Platforma Udostępniania Wynikowych Informacji Statystycznych i Innych Zasobów Informacyjnych | 2021-09-30 |

# KAMIENIE MILOWE

|  |  |
| --- | --- |
| **Kamienie milowe** | **Planowany termin osiągnięcia** |
| Wyłonienie wykonawcy usługi wsparcia informatycznego | 2019-09-30 |
| Opracowanie projektu technicznego Systemu Metainformacji | 2019-12-31 |
| Opracowanie koncepcji modernizacji Platformy Udostępniania Wynikowych Informacji Statystycznych i Innych Zasobów Informacyjnych | 2020-02-28 |
| Opracowanie projektów technicznych:  - Platformy Gromadzenia Danych Statystycznych  - Analitycznego Systemu Przetwarzania Danych | 2020-03-31 |
| Opracowanie prototypu Systemu Metainformacji | 2020-04-30 |
| Opracowanie prototypu Platformy Gromadzenia Danych Statystycznych | 2020-12-31 |
| Odbiór Portali Informacyjnych | 2021-07-31 |
| Opracowanie prototypu Analitycznego Systemu Przetwarzania Danych | 2021-03-31 |
| Przeprowadzenie testów akceptacyjnych i podpisanie protokołów odbioru poszczególnych komponentów systemu:  - System Metainformacji  - Platforma Gromadzenia Danych Statystycznych  - Analityczny System Przetwarzania Danych  - Platforma Udostępniania Wynikowych Informacji Statystycznych i Innych Zasobów Informacyjnych | 2021-09-30 |
| Testy akceptacyjne i odbiór systemu | 2022-01-31 |

# KOSZTY

## Koszty ogólne projektu wraz ze sposobem finansowania

|  |  |  |
| --- | --- | --- |
| **Całkowity koszt projektu (netto oraz brutto), w tym:** | 41 243 705,00 zł (brutto)  33 531 467,48 zł (netto)  Brak wydatków niekwalifikowanych | |
| **Procent dofinansowania ze środków UE (brutto)** | 84,63% | |
| **Procent środków z budżetu państwa (brutto)** | 15,37% | |
| **Podział całkowitego kosztu projektu na poszczególna lata (netto oraz brutto)** | rok 2019 | 8 002 967,48 zł netto  9 843 650,00 zł brutto |
| rok 2020 | 16 525 418,70 zł netto  20 326 265,00 zł brutto |
| rok 2021 | 7 785 195,12 zł netto  9 575 790,00 zł brutto |
| rok 2022 | 1 217 886,18 zł netto  1 498 000,00 zł brutto |

## Wykaz poszczególnych pozycji kosztowych

| **Nazwa pozycji kosztowej:** | | Przewidywany koszt brutto: | Uzasadnienie pozycji kosztowej (przeznaczenie): |
| --- | --- | --- | --- |
| Oprogramowanie | Koszt licencji | 14 998 500,00 | W związku z koniecznością utworzenia środowiska do pozyskiwania, przetwarzania, analizy i udostępniania danych niezbędny jest zakup oprogramowania w związku z modernizacją systemu Serwis Desk i budową Portali Informacyjnych, licencji oprogramowania operacyjnego i biurowego do dostarczanych stacji roboczych oraz licencji oprogramowania zarządzającego dostarczaną infrastrukturą sieciową, a także licencji rozszerzających zakres funkcjonalny systemu Call Center (4 783 000,00 zł).  Środki zostaną również przeznaczone na wynagrodzenia pracowników merytorycznych CIS (2 160 000,00 zł), wynagrodzenia pracowników merytorycznych US’ów (2 400 000,00 zł) oraz na koszty usług zewnętrznych (5 655 500,00 zł) w tym: usługi asysty eksperckiej dla opracowania szczegółowych projektów modernizacji lub budowy komponentów systemu oraz wsparcia wdrożeniowego. Wsparcie dla dostarczanej infrastruktury sprzętowo-systemowej oraz wdrażanych komponentów. |
| Infrastruktura (Środki trwałe i wartości niematerialne i prawne – Sprzęt informatyczny) | Koszt sprzętu | 16 778 500,00 | Dostawa infrastruktury sieciowej (urządzenia do akceleracji ruchu w sieci WAN i modernizacja punktów dostępowych) w CPD i LAN, dostawa bibliotek taśmowych do składowania backupów, modernizacja systemu Serwis Desk oraz zakup stacji roboczych (9 683 000,00 zł).  Środki zostaną również przeznaczone na wynagrodzenia pracowników merytorycznych CIS (1 440 000,00 zł) oraz na koszty usług zewnętrznych (5 655 500,00 zł) w tym: usługi asysty eksperckiej dla opracowania szczegółowych projektów modernizacji lub budowy komponentów systemu oraz wsparcia wdrożeniowego. Wsparcie dla dostarczanej infrastruktury sprzętowo-systemowej oraz wdrażanych komponentów. |
| Szkolenia | Koszt szkoleń w ramach budowania infrastruktury | 275 000,00 | Szkolenia dotyczące dostarczanych i wdrażanych elementów infrastruktury (głównie oprogramowanie). |
| Koszt szkoleń ogólnych | 1 000 805,00 | Podwyższanie kwalifikacji pracowników w zakresie nowych funkcjonalności oraz technologii IT wdrażanych w Projekcie, poznanie praktycznych aspektów zarządzania architekturą korporacyjną.  Szkolenie 1-dniowe dla Urzędów Statystycznych oraz departamentów z zakresu nowych funkcjonalności. |
| Działania informacyjno-promocyjne | Materiały i działania informacyjno-promocyjne | 500 000,00 | Działania promocyjne mają na celu przede wszystkim upowszechnienie informacji o realizacji Projektu, jego produktach i źródłach jego finansowania. Wpłynie to niewątpliwie na wzrost zainteresowania użytkowników końcowych możliwością korzystania z udostępnionych zasobów statystyki oraz nowych funkcjonalności. |
| Koszty UX i grafiki | Brak | 0,00 |  |
| Bezpieczeństwo | Brak | 0,00 |
| Wydajność rozwiązań | Brak | 0,00 |
| Koszty zarządzania i wsparcia (w tym wynagrodzenia personelu wspomagającego) | Wynagrodzenia zespołu obsługi, koszty delegacji  Usługi asysty eksperckiej przy realizacji Projektu | 7 690 900,00 | Środki przeznaczone będą na:   1. asystę ekspercką (1314900,00 zł) w zakresie:  * zarządzania i wdrażania Projektu * ustalenia wymagań użytkowników.  1. Wynagrodzenia zespołu obsługi oraz koszty delegacji (4648000,00 zł); 2. Wynagrodzenie bezpośrednie pracowników merytorycznych (1728000,00 zł) |

## Koszty ogólne utrzymania wraz ze sposobem finansowania (okres 5 lat)

|  |  |  |  |
| --- | --- | --- | --- |
| **Całkowity koszt utrzymania trwałości projektu (brutto)** | **1 870 000,81 zł netto**  **2 300 101,00 zł brutto** | | **Źródło finansowania** |
| **Podział całkowitego kosztu utrzymania trwałości projektu na poszczególna lata (netto oraz brutto)** | rok 2022 | **117 073,17 zł netto**  **144 000,00 zł brutto** | – budżet państwa |
| rok 2023 | **298 481,30 zł netto**  **367 132,00 zł brutto** | – budżet państwa |
| rok 2024 | **447 521,95 zł netto**  **550 452,00 zł brutto** | – budżet państwa |
| rok 2025 | **447 521,95 zł netto**  **550 452,00 zł brutto** | – budżet państwa |
| rok 2026 | **447 521,95 zł netto**  **550 452,00 zł brutto** | – budżet państwa |
| rok 2027 | **111 880,49 zł netto**  **137 613,00 zł brutto** | – budżet państwa |

# 

# 4.4 Planowane koszty ogólne realizacji (w przypadku projektu współfinansowanego – wkład krajowy z budżetu państwa) oraz koszty utrzymania projektu:

# zostaną pokryte w ramach budżetów odpowiednich dysponentów części budżetowych bez konieczności występowania o dodatkowe środki z budżetu państwa,

# ~~będą powodować konieczność przyznania dodatkowych kwot~~.[[2]](#footnote-3)

# GŁÓWNE RYZYKA

## Ryzyka wpływające na realizację projektu

|  |  |  |  |
| --- | --- | --- | --- |
| **Nazwa ryzyka** | Siła oddziaływania | Prawdopodobieństwo wystąpienia ryzyka | Sposób zarzadzania ryzykiem |
| Ryzyko braku zabezpieczenia wystarczających środków na realizację Projektu – z perspektywy finansowej 2014-2020 (POPC 2.3.1) lub ze środków budżetowych | Duża | Średnie | Określenie zakresu i harmonogramu projektu, aby możliwe było pozyskanie alternatywnych źródeł finansowania. Fazowanie Projektu. |
| Ryzyko zmian rynkowych związanych ze zmianami cen usług podczas realizacji Projektu – m.in. wzrost kosztów usług obcych, wzrost kosztów materiałów i energii, przekroczenie kosztów inwestycyjnych, wzrost wynagrodzeń, zmiana kursu walut. | Średnia | Małe | Dogłębna analiza związana z różnymi aspektami budowy i wdrożenia systemu. Ciągłe monitorowanie trendów rynkowych. Zabezpieczenie odpowiednich rezerw finansowych. |
| Ryzyko awarii spowodowane błędami dostawców (np. programistów), niewykrytymi w  trakcie testowania (luki bezpieczeństwa, ukryte błędy, niska wydajność aplikacji), wpływające na dostępność i bezpieczeństwo danych przetwarzanych przez system. | Duża | Średnie | Wdrożenie odpowiednich procedur odbiorów, testów akceptacyjnych. Testowanie zmian na środowisku testowym przed wprowadzeniem do produkcyjnego środowiska.  Wprowadzenie procedur odtworzenia systemu oraz systematycznego tworzenia kopii zapasowych.  Regularne kontrole jakości systemu.  Zaplanowanie testów bezpieczeństwa prowadzonych przez zewnętrzny podmiot – Audytora bezpieczeństwa. |
| Niezgodność i brak możliwości integracji rozwiązań obecnie użytkowanych z planowanymi do wdrożenia. | Duża | Małe | Precyzyjny opis wymagań na etapie SIWZ – szczegółowy opis obecnie funkcjonujących rozwiązań podlegających integracji.  Zapisy SIWZ uwzględniać będą aspekt integracji i zgodności rozwiązań używanych z wdrażanymi w Projekcie.  Wsparcie zewnętrzne do analizy i przygotowania przedmiotu zamówienia. Na etapie testowania duży nacisk położony na testy integracyjne. |
| Problemy z wykonaniem systemów przez Wykonawców wybranych do ich realizacji (brak potencjału i  doświadczenia, opóźnienia, brak możliwości technicznych lub organizacyjnych do realizacji zamówienia) | Duża | Średnie | W SIWZ zawarcie kryteriów formalnych na wybór podmiotu odpowiedzialnego za realizację zadania.  Wybór dostawców o odpowiednim do skali Projektu doświadczeniu i potencjale technicznym; stosowanie kryteriów jakościowych w ocenie ofert. Wsparcie zewnętrzne do analizy i przygotowania przedmiotu zamówienia i kryteriów oceny ofert. Monitorowanie na bieżąco postępów prac po stronie dostawców, zarówno usług, jak i sprzętu. Budowa silnego zespołu wsparcia po stronie Zamawiającego. |
| Złożoność Projektu. Projekt wymaga dostarczenia wielu komponentów. Poważne ryzyko związane jest z synchronizacją dostawy poszczególnych składników systemu (infrastruktura, oprogramowanie, wdrożenia itp.). Przedłużające się terminy przekazania systemu użytkownikom wewnętrznym i zewnętrznym. | Duża | Wysokie | Zapewnienie wysokiej jakości nadzoru na realizacją Projektu. Podjęcie standardowych działań zarządczych zgodnie z powszechnie znanymi metodykami. |
| Niski poziom dostarczanych produktów Projektu oraz niespełnienie kluczowych oczekiwań Zamawiającego będące wynikiem błędnie sformułowanych wymagań, nieprecyzyjnej komunikacji pomiędzy zespołami oraz zaangażowanymi użytkownikami. | Średnia | Małe | Precyzyjny opis wymagań na etapie SIWZ. Wsparcie zewnętrzne do analizy i przygotowania przedmiotu zamówienia. Włączenie reprezentatywnych przedstawicieli użytkowników końcowych w procesy ustalania wymagań na systemy, testowania systemów i odbioru systemów; włączenie wysokich rangą przedstawicieli wszystkich służb do Komitetu Sterującego Projektu lub do zespołu koordynującego tworzenie systemu. Precyzyjna komunikacja w formie elektronicznej i/lub pisemnej, z  zachowaniem ustaleń zapisanych w umowie, zatwierdzonych notatkach oraz zaproponowanych w dokumentacji prowadzenia Projektu (DIP). Precyzyjne formułowanie celów spotkań, warsztatów i innych wspólnych prac oraz respektowanie tych ustaleń. |
| Ryzyko wydłużenia czasu realizacji Projektu związanego z  prowadzeniem postępowań przetargowych (protesty i odwołania składane przez Wykonawców), jak również opóźnienia w procesie oceny ofert i wyboru Wykonawcy powodujący przekroczenie zakładanych terminów na przygotowanie i realizacje przetargów. | Duża | Średnie | Planowanie odpowiednich rezerw czasowych dla postępowań przetargowych.  Monitorowanie ścieżki krytycznej.  Opracowywanie dokumentacji przetargowych, minimalizujących liczbę protestów i odwołań KIO. |
| Ryzyko braku odpowiednich zasobów ludzkich po stronie GUS do nadzorowania i koordynowania przebiegu Projektu. | Średnia | Małe | Przydzielenie odpowiednich zasobów ludzkich do działań projektowych i zapewnienie wsparcia zewnętrznego, doradczego (Inżynier Projektu).  Odpowiednie pełnomocnictwa/uprawnienia dla Kierownika Projektu, pozwalające egzekwować od pracowników realizację nałożonych na nich obowiązków.  System motywacyjny dla członków zespołu projektowego premiujący za zakończenie całego Projektu z sukcesem. Wprowadzenie regularnego raportowania dla kierownictwa Projektu, ze wskazywaniem na opóźnienia w realizacji Projektu i ich przyczyny. |

## Ryzyka wpływające na utrzymanie efektów

|  |  |  |  |
| --- | --- | --- | --- |
| **Nazwa ryzyka** | Siła oddziaływania | Prawdopodobieństwo wystąpienia ryzyka | Sposób zarzadzania ryzykiem |
| Ryzyko niedoszacowania zasobów infrastruktury do zapewnienia prawidłowego przetwarzania danych w okresie eksploatacji. | Duża | Wysokie | Zapewnienie skalowalności rozwiązania w każdej warstwie.  Zapewnienie skalowalności infrastruktury teleinformatycznej (wirtualizacja lub zakupy uzupełniające). |
| Ryzyko związane z problemami z zasilaniem składnic danych i metadanych | Duża | Średnie | Bieżące informowanie jednostek autorskich o zmianie podejścia do procesu produkcji statystycznej. |
| Ryzyko związane z utratą trwałości projektu w okresie 5 lat. | Duża | Małe | Przygotowanie i wdrożenie skutecznej formuły organizacyjnej i prawnej, zapewniającej utrzymanie  i zarządzanie systemem oraz zapewnienie stabilnych źródeł utrzymania systemu po upływie okresu trwałości. Zapewnienie środków budżetowych na serwisowanie i rozwój oprogramowania, eksploatację. |
| Ryzyko braku zabezpieczenia środków finansowych na utrzymanie systemu po jego wdrożeniu | Duża | Małe | Analiza kosztów utrzymania systemu i zabezpieczenie odpowiednich środków w  budżecie Beneficjenta. |

# OTOCZENIE PRAWNE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Lp** | Tytuł aktu prawnego | Czy wymaga zmian? | Opis zmian (jeśli dotyczy) | Etap prac legislacyjnych (jeśli dotyczy) |
| 1 | Ustawa z dnia 29 czerwca 1995 r. o  statystyce publicznej (Dz. U. z 2018, poz. 997, z późn. zm.) | ~~TAK~~/NIE[[3]](#footnote-4) | Nie dotyczy | Nie dotyczy |

1. **ARCHITEKTURA**
   1. **Widok kooperacji aplikacji**

Celem Projektu jest zwiększenie jakości i ilości informacji udostępnianych przez statystykę publiczną. Nowe funkcjonalności zostaną udostępnione poprzez budowę, modernizację i wykorzystanie istniejących systemów teleinformatycznych zgodnie z przyjętym „Modelem Procesu Produkcji Statystycznej”   
(Rysunek 1).

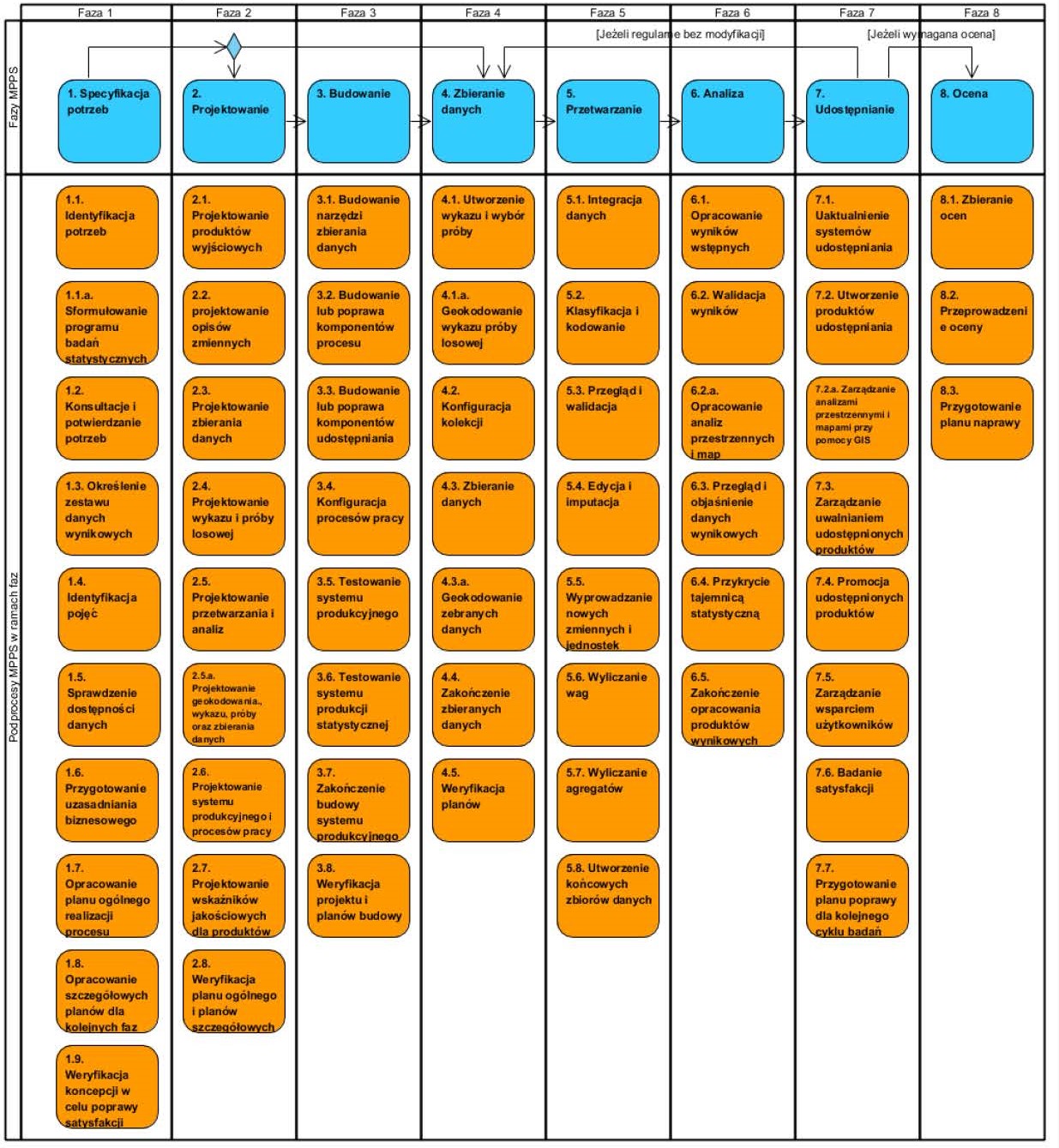
Na podstawie modelu stworzono ramy architektoniczne będące podstawą budowy oraz unormowanego rozwoju rozwiązań informatycznych wykorzystywanych do realizacji zadań statystyki publicznej. Podejście zgodne z architekturą korporacyjną ma na celu zapewnienie trwałości, spójności oraz obniżenie kosztów rozwiązania.

Na Rysunek 1 zgrupowano procesy produkcji statystycznej w fazy. Holistyczne podejście do procesu produkcji pozwoliło na wyodrębnienie systemów i aplikacji niezbędnych do realizacji procesów, od fazy specyfikacji do fazy udostępniania. Projektowane, modyfikowane i istniejące systemy aplikacyjne zostały wyszczególnione w Tabela 2 oraz oznaczone na schemacie kooperacji aplikacji (Rysunek 2).

Dane wynikowe i rejestrowe przeznaczone do udostępnienia, niezależnie od systemu i formy prezentacji, zostaną udostępnione z jednej struktury bazodanowej, tj. „Składnicy Danych Publikacyjnych”. Udostępnianie danych z jednej składnicy pozwoli na uporządkowanie procesu, co w konsekwencji doprowadzi do ograniczenia liczby kanałów prezentacji danych, wzrostu zasobu informacyjnego przeznaczonego do udostępniania i wydajniejszej kontroli jakości udostępnianych danych. Niezależnie od procesów fazy udostępniania przewiduje się w Projekcie zapewnienie dostępu do danych i informacji z innych faz procesu produkcji poprzez API, formularze i interfejsy systemowe (Rysunek 2).

Dla każdej z faz zostanie utworzona składnica danych, będąca logiczną lub rzeczywistą strukturą bazodanową, w celu odseparowania proponowanych rozwiązań architektonicznych od wpływu konkretnych produktów (głównie baz danych i hurtowni) oferowanych przez różnych producentów komercyjnych. Pozwoli to na uogólnienia opisu funkcji obiektów bazodanowych w procesie produkcji statystycznej przechowujących i udostępniających kolejne „stany” danych statystycznych. Wrota Statystyki, poprzez wdrażane usługi i  systemy operujące na składnicach, będą wspierać realizację podprocesów MPPS określonych w Tabela 1.

Kluczowymi produktami Projektu, poza systemami realizującymi usługi udostępniania danych w różnych fazach procesów statystycznych, będą System Metainformacji i Składnica Metadanych Statystycznych. Celem wdrożenia Systemu Metainformacji jest efektywne zarządzanie procesami produkcji poprzez gromadzenie metadanych i wykorzystanie ich do sterowania procesami. Wdrożenie Składnicy Metadanych Statystycznych pozwoli na zapewnienie jednego, spójnego źródła metadanych definicyjnych, badań, strukturalnych oraz procesowych. System Metainformacji zapewni i udostępni dla wszystkich faz przetwarzania zestaw metadanych strukturalnych stanowiących opis struktur produktów statystycznych na podstawie informacji zgromadzonych Składnicy Metadanych Statystycznych.

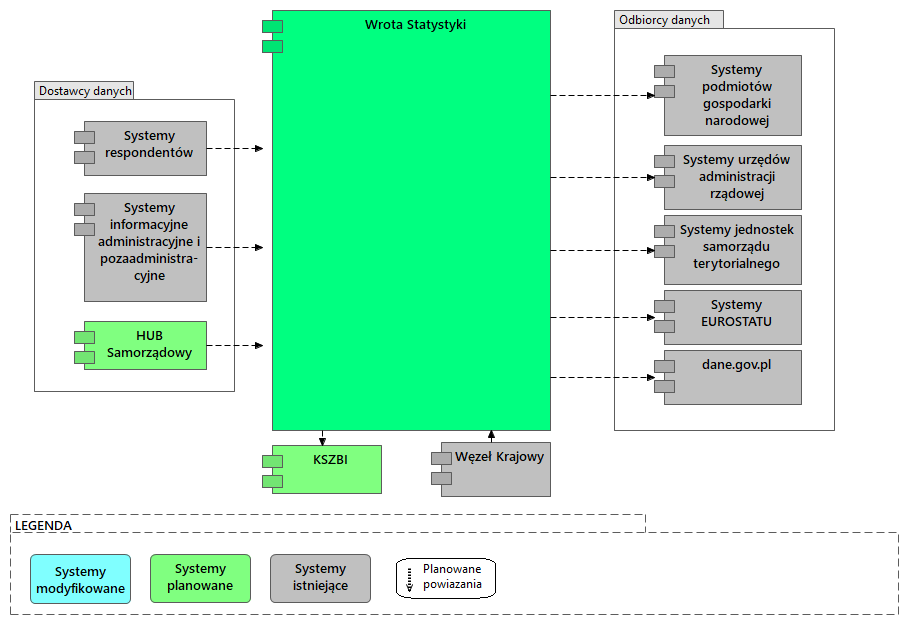


Rysunek 1 – Model Procesu Produkcji Statystycznej „MPPS”

**Tabela 1 - Zestawiezienie powiązań Wrót Statystyki z modelem MPPS**

| Fazy MPPS | Podprocesy | Składnica Metadanych Statystycznych | Składnica Potrzeb Użytkownika | Składnica Jednostek Statystycznych | Składnica Danych Geoprzestrzennych | Składnica Danych Surowych | Składnica Danych Operacyjnych | Składnica danych Analitycznych | Składnica Danych Publikacyjnych |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Specyfikacja potrzeb** | 1.1 | v | v |  |  |  |  |  |  |
| 1.1a | v | v |  |  |  |  |  |  |
| 1.3 | v |  |  |  |  |  |  |  |
| 1.4 | v |  |  |  |  |  |  |  |
| 1.5 | v |  |  |  |  |  |  |  |
| 1.7 | v |  |  |  |  |  |  |  |
| 1.8 | v |  |  |  |  |  |  |  |
| **Projektowanie** | 2.1 | v |  |  |  |  |  |  |  |
| 2.2 | v |  |  |  |  |  |  |  |
| 2.3 | v |  |  |  |  |  |  |  |
| 2.4 | v |  | v |  |  |  |  |  |
| 2.5 | v |  | v |  |  |  |  |  |
| 2.5a | v |  | v | v |  |  |  |  |
| 2.6 | v |  |  |  |  |  |  |  |
| 2.7 | v |  |  |  |  |  |  |  |
| **Budowanie** | 3.1 | v |  |  |  | v |  |  |  |
| 3.2 | v |  |  |  |  | v | v |  |
| 3.3 | v |  |  |  |  |  | v |  |
| 3.4 | v |  |  |  |  |  |  |  |
| 3.5 | v |  |  |  | v | v | v |  |
| 3.6 | v |  |  |  |  | v | v |  |
| 3.7 | v |  |  |  |  |  |  |  |
| 3.8 | v |  |  |  |  |  |  |  |
| **Zbieranie danych** | 4.1 | v |  | v |  |  |  |  |  |
| 4.1a | v |  |  | v |  |  |  |  |
| 4.2 | v |  |  |  |  |  |  |  |
| 4.3 | v |  | v |  | v |  |  |  |
| 4.3a | v |  |  | v | v |  |  |  |
| 4.4 | v |  |  |  | v |  |  |  |
| 4.5 | v |  |  |  |  |  |  |  |
| **Przetwarzanie** | 5.1 | v |  |  |  | v |  |  |  |
| 5.2 | v |  |  |  |  | v |  |  |
| 5.3 | v |  |  |  |  | v |  |  |
| 5.4 | v |  |  |  |  | v |  |  |
| 5.5 | v |  |  |  |  | v |  |  |
| 5.6 | v |  |  |  |  | v |  |  |
| 5.7 | v |  |  |  |  | v |  |  |
| 5.8 | v |  |  |  |  | v |  |  |
| **Analiza** | 6,1 | v |  |  |  |  | v |  |  |
| 6.2 | v |  |  |  |  |  | v |  |
| 6.2a | v |  |  | v |  |  | v |  |
| 6.3 | v |  |  |  |  |  | v |  |
| 6.4 | v |  |  |  |  |  | v |  |
| 6.5 | v |  |  |  |  |  | v |  |
| **Udostępnianie** | 7.1 | v |  |  |  |  |  |  | V |
| 7.2 | v |  |  |  |  |  | v | V |
| 7.2a | v |  |  | v |  |  |  | V |
| 7.3 | v |  |  |  |  |  |  |  |
| 7.4 | v |  |  |  |  |  |  |  |
| 7.5 | v |  |  | v |  |  |  | V |
| 7.6 | v |  |  |  |  |  |  |  |
| 7.7 | v |  |  |  |  |  |  |  |
| **Ocena** | 8.1 | v |  |  |  |  |  |  |  |
| 8.2 | v |  |  |  |  |  |  |  |
| 8.3 | v |  |  |  |  |  |  |  |

**Rysunek 2 – Schemat kooperacji aplikacji**



| **Status** | **Opis** |
| --- | --- |
| Planowany | System projektowany, w trakcie budowy, w trakcie wdrożenia. |
| Modyfikowany | System modyfikowany, rozszerzany na potrzeby projektu. |
| Istniejący | System działający produkcyjnie, gotowy do wykorzystania. |

**Tabela 2 – Lista systemów wykorzystywanych w projekcie**

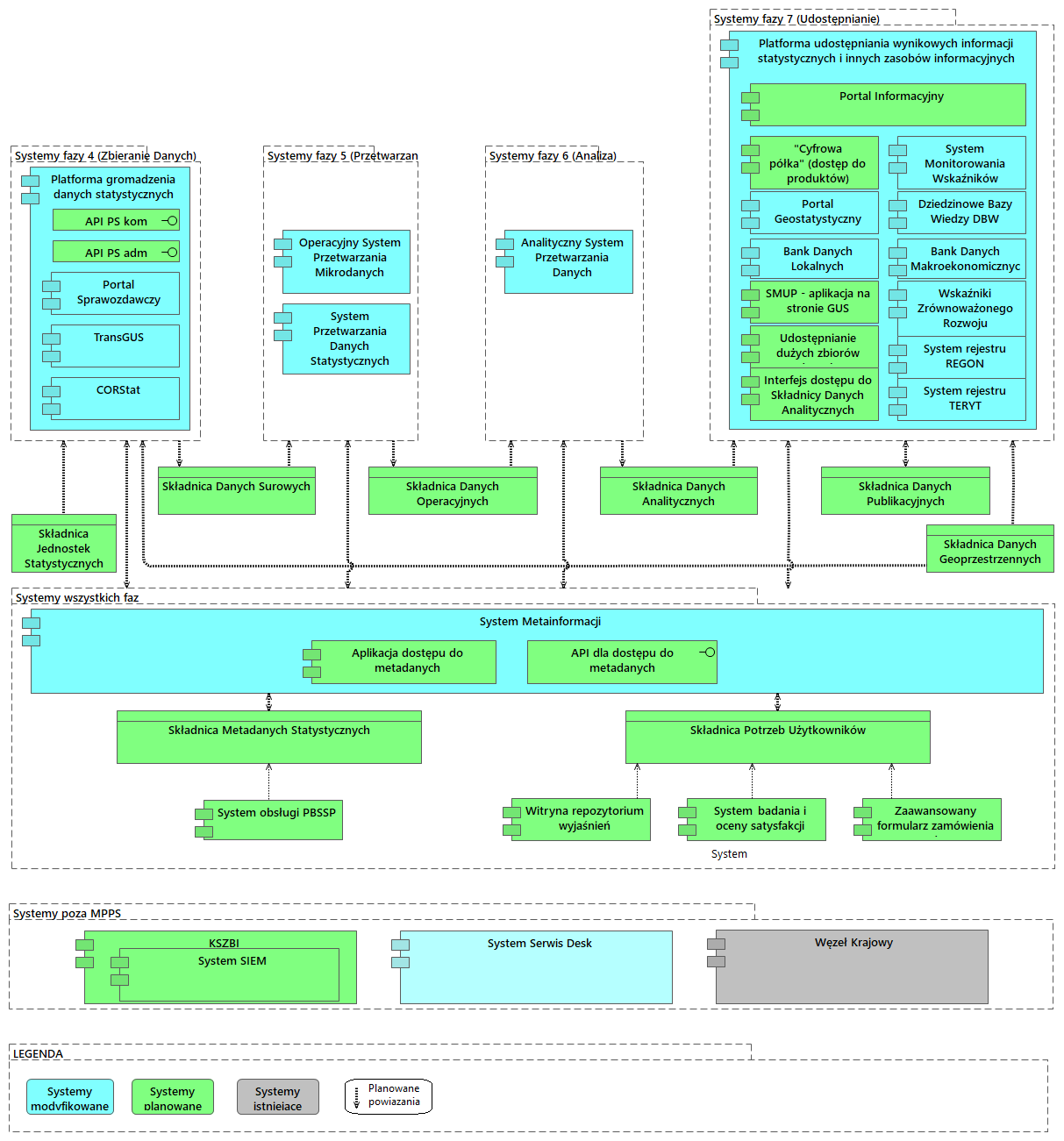
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Lp.** | **Nazwa systemu** | **Gestor systemu** | **Opis systemu** | **Status** | **Krótki opis ewentualnej zmiany** |
| 1 | Wrota Statystyki | GUS | System zapewni zmodernizowanie procesu produkcji statystycznej w celu:  - podniesienia jakości udostępnianych informacji oraz - udostępniania informacji poprzez planowane komponenty. | System planowany | Wprowadzenie nowych funkcjonalności:  **1. API PS kom**  Zautomatyzowane pozyskiwanie danych poprzez API;  **2. API PS adm**  Zautomatyzowane pozyskiwanie danych poprzez API od administracji;  **3. Informacja o obowiązkach sprawozdawczych**  Dostęp do repozytorium obowiązków sprawozdawczych, wybór preferowanego kanału odbierania powiadomień i kanału realizacji obowiązków sprawozdawczych, spersonalizowana informacja zwrotna dla podmiotów realizujących obowiązek sprawozdawczy;  **4. Udostępnienie narzędzi analitycznych**  Zapewnienie możliwości prowadzenia analiz dla użytkowników przez udostępnienie narzędzi analitycznych  **5. Profilowane Portale Informacyjne**  Zapewnią integrację źródeł udostępniania danych oraz wyszukiwarkę podpowiedzi kategorii informacji/działu/bazy danych/klasyfikacji;  **6. Cyfrowa półka (Dostęp do produktów)**  Utworzenie spersonalizowanego pulpitu klienta z dostępem do: złożonych zamówień, repozytorium zapewniającego ponowne wykorzystanie udostępnionych danych, funkcji zapisu wyszukanych danych z różnych podsystemów udostępniania;  **7. Udostępnianie zbiorów danych wynikowych**  Udostępnianie zbiorów danych wynikowych odbiorcom informacji, w tym w wersji mobilnej, uwzględniającej profil użytkownika  **8. Dostęp do zasobów Składnicy Metadanych Statystycznych**  Dostępu do metadanych poprzez aplikację webową i interfejs API;  **9. Zaawansowany formularz zamówienia na dane**  Dostęp do informacji statystycznej poprzez aplikację formularza i wirtualnego asystenta;  **10. Repozytorium wyjaśnień**  Dostęp do repozytorium odpowiedzi na najczęściej zadawane pytania dotyczące informacji statystycznych oraz wyjaśnień dot. metodologii;  **11. Badanie potrzeb i ocena satysfakcji**  Narzędzia pozwalające na automatyzację zbierania informacji o potrzebach użytkowników oraz umożliwiające uzyskanie oceny satysfakcji.  **12. PBSSP**  Obsługa przygotowania i uzgodnienia projektu PBSSP, w tym zarządzanie uwagami i stanowiskami do uwag. |
| 2 | Systemy respondentów | - | Systemy zewnętrzne stanowiące źródło danych dla procesu produkcji statystycznej. | Systemy istniejące | Możliwość integracji/komunikacji systemów zewnętrznych poprzez udostępnione API. |
| 3 | Systemy informacyjne administracyjne i pozaadminist-racyjne | - | Systemy zewnętrzne stanowiące źródło danych dla procesu produkcji statystycznej. | Systemy istniejące | Możliwość integracji/komunikacji systemów zewnętrznych poprzez udostępnione API. |
| 4 | Systemy podmiotów gospodarki narodowej | - | Systemy zewnętrzne odbierające informacje statystyczne oraz dane z rejestrów urzędowych prowadzonych przez Prezesa GUS. | Systemy istniejące | Możliwość integracji/komunikacji systemów zewnętrznych poprzez udostępnione API. |
| 5 | Systemy urzędów administracji rządowej | - | Systemy zewnętrzne odbierające informacje statystyczne oraz dane z rejestrów urzędowych prowadzonych przez Prezesa GUS. | Systemy istniejące | Możliwość integracji/komunikacji systemów zewnętrznych poprzez udostępnione API. |
| 68 | Systemy jednostek samorządu terytorialnego | - | Systemy zewnętrzne odbierające informacje statystyczne oraz dane z rejestrów urzędowych prowadzonych przez Prezesa GUS. | Systemy istniejące | Możliwość integracji/komunikacji systemów zewnętrznych poprzez udostępnione API. |
| 7 | Systemy Eurostatu | Eurostat | Systemy zewnętrzne odbierające informacje statystyczne. | System istniejący | Możliwość integracji/komunikacji systemów zewnętrznych poprzez wytworzone API. |
| 8 | Dane.gov.pl | MC | System zewnętrzny odbierający informacje statystyczne oraz dane z rejestrów urzędowych prowadzonych przez Prezesa GUS, wdrożony w ramach projektu Otwarte Dane. | Systemy istniejące | Integracja z systemami Wrót Statystyki. |
| 9 | HUB Samorządowy | GUS | System powstaje w ramach projektu SMUP. System będzie stanowił repozytorium dla danych zebranych w ramach projektu SMUP, które będą mogły być źródłem danych poprzez replikowanie tych danych do odpowiednich Składnic Danych w zależności od stopnia ich przetworzenia w HUBie. | System planowany | Integracja z systemami Wrót Statystyki. |
| 10 | Kompleksowy System Zarządzania Bezpieczeństwem Informacji | GUS | System, którego centralnym komponentem, istotnym z perspektywy projektu jest system typu SIEM, służący do analizy informacji o zdarzeniach, w czasie rzeczywistym. Jednymi ze źródeł informacji tego systemu będą podsystemy wchodzące w skład WRÓT STATYSTYKI. System powstaje w ramach innego projektu. | System planowany | - |
| 11 | Krajowy Węzeł Identyfikacji Elektronicznej | MC | System będzie pośredniczyć w uwierzytelnianiu w krajowych usługach online za pomocą środków identyfikacji elektronicznej wydanych przez różne podmioty. Zostanie wykorzystany do uwierzytelniania użytkowników w fazie Zbierania danych i fazie udostępniania. System wdrażany przez inny podmiot. | System istniejący | - |

**Tabela 3 – Lista przepływów**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Lp.** | **System źródłowy** | **System docelowy** | **Zakres wymienianych danych** | **Sposób wymiany danych** | **Typ modyfikacji** | **Typ Interfejsu** |
| 1 | Systemy respondentów | Wrota Statystyki | Dane pozyskiwane do badań | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP, API |
| 2 | Systemy informacyjne administracyj-ne i pozaadminis-tracyjne | Wrota Statystyki | Dane pozyskiwane do badań | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP, API |
| 3 | HUB Samorządowy | Wrota Statystyki | Dane pozyskiwane do badań | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP, API |
| 4 | Wrota Statystyki | Systemy podmiotów gospodarki narodowej | Informacje statystyczne | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP, API |
| 5 | Wrota Statystyki | Systemy urzędów administracji rządowej | Informacje statystyczne | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP, API |
| 6 | Wrota Statystyki | Systemy jednostek samorządu terytorialnego | Informacje statystyczne | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP, API |
| 7 | Wrota Statystyki | Systemy Eurostatu | Informacje statystyczne | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP |
| 8 | Wrota Statystyki | Dane.gov.pl | Informacje statystyczne | odwołania bezpośrednie / kopiowanie | Krytyczny | SOAP |
| 9 | Węzeł Krajowy | Wrota Statystyki | Dane uwierzytelniające | odwołania bezpośrednie | Niekrytyczny | SOAP |
| 10 | Wrota Statystyki | KSZBI – System SIEM | Logi | odwołania bezpośrednie | Niekrytyczny | API |

* 1. **Kluczowe komponenty architektury rozwiązania**

**Rysunek 3 – Kluczowe komponenty Wrót Statystyki**

****

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 Przetwarzania Produkcji Statystycznej poprzez utworzenie grup systemów odpowiadających fazom produkcji statystycznej i przypisanie im nowych funkcjonalności związanych z udostępnianiem informacji:

1. Systemy fazy zbierania danych:
   1. Modyfikowany Portal Sprawozdawczy, którego zadaniem będzie umożliwienie składnia sprawozdań w formie elektronicznej poprzez udostępnienie usługi przekazywania zestawów danych wyposażonej w mechanizm walidacji danych. Przewiduje się nowe funkcjonalności:
      * W zakresie automatyzacji procesów udostępnione zostaną interfejsy API umożliwiające transfer danych bezpośrednio z systemów respondentów.
      * Powiadomienie i wybór kanału sprawozdawczości pozwalające na wybór preferowanego kanału odbierania powiadomień o obowiązkach sprawozdawczych i preferowanego kanału realizacji obowiązku.
      * Informacja zwrotna, tj. udostępnianie informacji zwrotnej podmiotom, które zrealizowały obowiązek sprawozdawczy. Celem usługi jest podkreślenie znaczenia respondenta oraz przekazanie mu przydatnych informacji dotyczących branży związanej z jego działalnością.
      * Informacja o obowiązkach sprawozdawczych, tj. zapewnienie dostępu do zasobów wiedzy wyjaśniających obowiązki sprawozdawcze oraz zasady ich realizacji.
   2. Modyfikowany TransGUS, którego zadaniem jest umożliwienie przesyłania danych z jednostek administracji publicznej oraz innych instytucji. Przewiduje się nowe funkcjonalności:
      * W zakresie automatyzacji procesów udostępnione zostaną interfejsy API umożliwiające transfer danych bezpośrednio z systemów administracji publicznej i innych jednostek.
   3. Modyfikowany CORstat, którego zadaniem będzie obsługa zbierania danych kanałami CAPI/CATI/CAII oraz nowych komponentów.
2. Systemy fazy przetwarzania:
   1. Modyfikowany Operacyjny System Przetwarzania Mikrodanych, którego zadaniem jest wstępne przetwarzanie danych pobranych ze Składnicy Danych Surowych, tj. standaryzacja, transformacja i integracja danych zgodnie z regułami pobranymi ze Składnicy Metadanych Statystycznych. Jeżeli dane mają być poddane bardziej zaawansowanym metodom przetwarzania, wysyłane są do Systemu Przetwarzania Danych Statystycznych, jeśli nie – zapisywane są w postaci zbiorów danych w Składnicy Danych Operacyjnych.
      * System zostanie zoptymalizowany i zintegrowany z komponentami Wrót Statystyki zgodnie z przepływami wskazanymi na Rysunek 2.
   2. Modyfikowany System Przetwarzania Danych Statystycznych, którego zadaniem jest zaawansowane przetwarzanie danych pobranych ze Składnicy Danych Surowych lub z Operacyjnego Systemu Przetwarzania Mikrodanych, tj. wyliczenie nowych zmiennych, uzupełnienie braków danych (imputacja), korekta danych, walidacja, anonimizacja lub pseudonimizacja identyfikowalnych danych jednostkowych, wyliczanie wag, agregatów i wskaźników. Reguły przetwarzania pobierane są ze Składnicy Metadanych Statystycznych. Rezultatem procesu jest utworzenie zbiorów danych gotowych do procesów analitycznych. Przetworzone zbiory zasilają Składnicę Danych Operacyjnych.
      * System zostanie zoptymalizowany i zintegrowany z komponentami Wrót Statystyki zgodnie z przepływami wskazanymi na Rysunek 3.
3. Systemy fazy analizy:
   1. Modyfikowany Analityczny System Przetwarzania Danych, którego zadaniem jest wytworzenie i sprawdzenie analitycznych produktów statystycznych, tj. przeznaczonych do opracowania lub udostępniania wyników badań statystycznych. Przewiduje się nowe funkcjonalności:
      * Udostępnienie narzędzi analitycznych - funkcjonalność umożliwi spersonalizowany dostęp do udostępnionych przez statystykę publiczną narzędzi analizy statystycznej do wybranych danych analitycznych przykrytych tajemnicą statystyczną. Dostęp do funkcjonalności będzie wymagał uwierzytelnienia w domenie statystyki.
4. Systemy fazy udostępniania (Platforma udostępniania wynikowych informacji statystycznych i innych zasobów informacyjnych):
   1. Planowana Platforma udostępniania wynikowych informacji statystycznych i innych zasobów informacyjnych pobiera dane ze Składnicy Danych Publikacyjnych na potrzeby systemów udostępniania, pobiera metadane z Systemu Metainformacji i zapisuje w nim informacje pochodzące od użytkowników, obejmujące ich oceny i preferencje. Platforma będzie oferowała możliwość komunikacji poprzez API, w tym także dla systemów zrealizowanych w technologii open source (z upublicznionym kodem źródłowym), . Przewiduje się nowe funkcjonalności:
      * Zbudowanie, profilowanych Portali Informacyjnych, z możliwością wykorzystania systemów do zarządzania treścią w technologii open source (z upublicznionym kodem źródłowym),umożliwiających integrację źródeł udostępniania danych.
      * Wprowadzenie zaawansowanej wyszukiwarki dla podpowiedzi kategorii informacji/działu/bazy danych/klasyfikacji, w której występuje szukana fraza oraz prezentację wyników wyszukiwania według ich ilości i miejsc występowania frazy.
      * Personalizacja udostępniania danych poprzez utworzenie pulpitu klienta z dostępem do złożonych zamówień oraz funkcją zapisu wyszukanych danych z podsystemów udostępniania danych („Cyfrowa półka”).
      * Udostępnianie zbiorów danych wynikowych- funkcjonalność pozwalająca na udostępnianie zbiorów danych wynikowych odbiorcom informacji, w sposób wydajny i bezpieczny.
      * Systemy klienckie zewnętrzne uzyskują dostęp do zasobów platformy poprzez implementację udostępnianych API.
   2. Systemy udostępniania informacji składające się na Platformę udostępniania wynikowych informacji statystycznych i innych zasobów informacyjnych zostaną zmodernizowane w celu zapewnienia personalizacji udostępniania danych.

W ramach Wrót Statystyki przewiduje się wdrożenie systemów dla wszystkich faz procesu produkcji oraz struktur bazodanowych w postaci składnic:

1. Modyfikowany System Metainformacji zapewni i udostępni dla wszystkich faz procesu produkcji statystycznej metadane definicyjne, badań i metodologiczne, procesowe i strukturalne. Pozwoli to sterować całościowo procesem produkcji statystycznej. System zapewni nowe funkcjonalności związane z udostępnianiem informacji:
   1. Dostęp do zasobów Składnicy Metadanych Statystycznych poprzez dedykowaną aplikację webową.
   2. Dostęp do zasobów Składnicy Metadanych Statystycznych poprzez interfejs API.
   3. Zaawansowany formularz zamówienia na dostęp do informacji statystycznej poprzez aplikację formularza.
   4. Zaawansowany formularz zamówienia na dostęp do informacji statystycznej poprzez wirtualnego asystenta online, np. chat z pracownikiem informatorium.
   5. Wprowadzenie Repozytorium wyjaśnień, tj. funkcji dostępu do repozytorium odpowiedzi na najczęściej zadawane pytania dotyczące informacji statystycznych oraz wyjaśnień dot. metodologii.
   6. System do badania i oceny satysfakcji klientów poprzez dedykowane narzędzia pozwalające na dostęp do wyników użytkownikom wewnętrznym i zewnętrznym. Narzędzie będzie dostępne dla odbiorców informacji statystycznych, korzystających z Portali Informacyjnych.
   7. Zapewnienie obsługi przygotowania i uzgodnienia projektu PBSSP, w tym zarządzanie uwagami i stanowiskami do uwag.
2. Planowane składnice:
   1. Składnica Jednostek Statystycznych, której zadaniem będzie obsługa fazy 4 „Zbieranie danych”. Jest to 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.
   2. Składnica Danych Surowych, której zadaniem będzie obsługa fazy 4 . Struktura bazodanowa pozwalająca na przechowywanie i udostępnianie do dalszej obróbki nieprzetworzonych statystycznie danych, zebranych zarówno pierwotnie (od respondentów), jak i wtórnie pozyskanych z systemów administracyjnych oraz innych źródeł.
   3. Składnica Danych Operacyjnych, której zadaniem będzie obsługa fazy 5 „Przetwarzanie”. Struktura bazodanowa przechowująca przetworzone statystycznie dane jednostkowe. Wystandaryzowane dane w SDO podlegają niezbędnej integracji, pseudonimizacji lub anonimizacji oraz agregacji w celu ich przygotowania do analiz statystycznych.
   4. Składnica Danych Analitycznych, której zadaniem będzie obsługa fazy 6 „Analiza”. 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).
   5. Składnica Danych Publikacyjnych, której zadaniem będzie obsługa fazy 7 „Udostępnianie”. 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.
   6. 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.
   7. Składnica Metadanych Statystycznych. Baza wiedzy o wszystkich zasobach i procesach związanych z produkcją danych statystycznych. Jest niezbędna do monitorowania i zarządzania procesem produkcji statystycznej.
   8. Składnica Potrzeb Użytkowników - Struktura bazodanowa, która przechowuje i udostępnia zebrane różnymi kanałami komunikacyjnymi dane o istniejących i oczekiwanych potrzebach informacyjnych użytkowników statystyki publicznej.
3. Modernizowany System Serwis Desk, z możliwością wykorzystania technologii open source (z upublicznionym kodem źródłowym), który zostanie rozszerzony na wszystkie jednostki organizacyjne statystyki publicznej.

Ponadto przewiduje się wykorzystanie systemów zewnętrznych wobec projektu:

1. System klasy SIEM, planowany do wdrożenia w ramach projektu „KSZBI”, będzie zbierał i analizował dane napływające z systemów monitorujących oraz z kluczowych dla statystyki publicznej aplikacji, w tym wytworzonych w ramach Wrót Statystyki.
2. Krajowy Węzeł Identyfikacji Elektronicznej.
3. Inne systemy zewnętrzne przedstawione na Rysunek 2.

Produkty projektu i informacje statystyczne będą udostępniane użytkownikom zgodnie z wymaganiami standardu WCAG 2.0.Dane będą zgodne ze standardem 5 Star Open Data na poziomie co najmniej 3. Zostaną opatrzone metadanymi udostępnianymi niezależnie przez API. Systemy udostępniania będą zawierały udokumentowane interfejsy dla programistów (API).

* 1. **Przyjęte założenia technologiczne**

|  |  |  |
| --- | --- | --- |
| **Lp.** | **Obszar** | **Założenie technologiczne** |
|  | Infrastruktura | - |
|  | Sieć i bezpieczeństwo | - |
|  | Standardy wymiany danych | - |
|  | Systemy operacyjne serwerowe | System zostanie zbudowany ma dostosowanych do istniejących rozwiązań w Centrum Przetwarzania Danych Głównego Urzędu Statystycznego systemach operacyjnych Windows Serwer lub opartych na Linux. Wirtualizacja zostanie zapewniona przez aktualnie wdrożony Vmware. |
|  | Bazy danych | - |
|  | Serwery aplikacji | Dla aplikacji wykorzystywane będą wdrożone serwery z oprogramowaniem wymienionym w pkt 4 z użyciem technologii Internet Information Services oraz Apache. |
|  | Portale | - |
|  | Inne | - |

Założenia technologiczne w pkt 4 i 6 wynikają z konieczności dostosowania projektowanych rozwiązań do istniejącej infrastruktury wykorzystywanej w Centrum Przetwarzania Danych Głównego Urzędu Statystycznego.

* 1. **Opis zasobów danych przetwarzanych w planowanym rozwiązaniu**

Czy nowy system będzie tworzył zasoby danych o charakterze rejestru publicznego?

~~TAK~~/**NIE** [[4]](#footnote-5)

|  |  |  |
| --- | --- | --- |
| **Lp.** | **Tworzony rejestr publiczny** | **Opis** |
|  |  |  |

Czy nowy system będzie przetwarzał (używał, zmieniał) zawartość innych rejestrów publicznych?

**TAK**/~~NIE~~ [[5]](#footnote-6)

|  |  |  |  |
| --- | --- | --- | --- |
| **Lp.** | **Rejestr publiczny** | **Opis** | **Zakres przetwarzania** |
| 1 | REGON | Rejestr REGON będzie podstawą SJS w zakresie przedsiębiorstw | Użycie – w pełnym zakresie dostępnym dla podmiotów państwowych |
| 2 | TERYT | Rejestr TERYT będzie podstawą SJS w zakresie kodowania adresów | Użycie – w pełnym zakresie dostępnym dla podmiotów państwowych |

* 1. **Bezpieczeństwo**

Planowany poziom zapewnienia bezpieczeństwa (w rozumieniu przepisów §20 rozporządzenia Rady Ministrów z dnia 12 kwietnia 2012 r. w sprawie Krajowych Ram Interoperacyjności […] (Dz. U. 2012, poz. 526 z późn. zm.) w zakresie dot. systemu zarządzania bezpieczeństwem informacji:

* ~~system nie podlega rygorom KRI – należy wyjaśnić czy istnieją inne normy bezpieczeństwa, które będą spełnione przez system zgodnie z wymogami KRI,~~
* ~~dodatkowe zabezpieczenia powyżej wymogów KRI: należy wskazać uzasadnienie>>~~~~[[6]](#footnote-7)~~

Projekt bezpośrednio realizuje wymagania KRI w zakresie zarządzania bezpieczeństwem informacji poprzez wdrożenie nowych technologii zwiększających bezpieczeństwo informacji w sieciach i systemach teleinformatycznych zgodnie z ustawą z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne (Dz. U. z 2014 r. poz. 1114) dzięki zapewnieniu właściwej ochrony zarówno przed zagrożeniami wewnętrznymi, mającymi źródło wewnątrz organizacji, jak i zewnętrznymi mającymi swoje źródło na zewnątrz organizacji min. poprzez sieć komputerową z wykorzystaniem elementów technicznych w postaci rozwiązań sprzętowych oraz oprogramowania.

W ramach Projektu zrealizowane zostaną działania w obszarze: bezpiecznego dostępu do sieci, bezpieczeństwa sieci, bezpieczeństwa danych i aplikacji.

Projekt uwzględniać będzie normy i standardy:

* PN-ISO/IEC 20000-1 i PN-ISO/IEC 20000-2 – w zakresie zarządzania usługami;
* grupy norm z serii PN-ISO/IEC 2700x – w zakresie zarządzania bezpieczeństwem informacji;
* PN-EN ISO 22301:2014-11 – w zakresie wymagań dla systemów zarządzania ciągłością działania.

Zakresem Projektu objęte zostaną obszary:

* bezpieczeństwo i zarządzanie siecią,
* dostęp do sieci i usług sieciowych,
* zarządzanie i monitorowanie wydajności sieci i systemów.

Z punktu widzenia architektury systemu zabezpieczeń w GUS, wdrożenie niezależnie systemu zabezpieczeń wewnętrznych segmentów sieci, zapewni szeroką wielopoziomową ochronę sieci GUS oraz znajdujących się w niej urządzeń, aplikacji i danych, jednocześnie dostarczając nowy poziom ochrony ruchu sieciowego i zapewniając proaktywną ochronę wymaganą dla dzisiejszych i przyszłych centrów danych. Nowy system Firewall/IPS będzie komplementarny z równoległym projektem dotyczącym zwiększenia bezpieczeństwa „KSZBI – Kompleksowy System Bezpieczeństwa Informacji.

1. Niepotrzebne skreślić. [↑](#footnote-ref-2)
2. Niepotrzebne skreślić [↑](#footnote-ref-3)
3. Niepotrzebne skreślić. [↑](#footnote-ref-4)
4. Niepotrzebne skreślić. [↑](#footnote-ref-5)
5. Niepotrzebne skreślić [↑](#footnote-ref-6)
6. Niepotrzebne skreślić [↑](#footnote-ref-7)