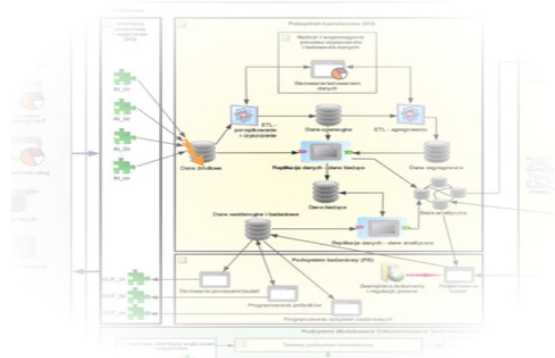


## Platforma informatyczna systemu badań i diagnozowania właściwości usług komunikacji elektronicznej PIBUK



### BIULETYN INFORMACYJNY NR 7

sierpień 2015

*Egzemplarz bezpłatny*

#### Kalendarium

1. Ogłoszenie postępowania o kompleksową organizację i przeprowadzenie szkoleń ogólnych i specjalistycznych dla pracowników Instytutu Łączności: 27.05.2015 r.
2. Składanie ofert na szkolenia do: 10.07.2015 r.
3. Rozstrzygnięcie postępowania o udzielenie zamówienia na szkolenia: 29.07.2015 r.
4. Podpisanie umowy na realizację szkoleń: 29.07.2015 r.

*Uwaga: Koncepcja szkoleń ogólnych została przedstawiona w Biuletynie nr 3.*

#### Przykładowy System Badań i Diagnozowania

Wykonawca umowy na dostawę platformy informatycznej PIBUK ma zaprojektować, zaimplementować i wdrożyć na tej platformie Przykładowy System Badań i Diagnozowania (PSBD). PSBD zostanie zrealizowany z użyciem oprogramowania Business Intelligence dostarczonego w ramach umowy.

Celem PSBD jest praktyczne zademonstrowanie metody projektowania i implementacji aplikacji z wykorzystaniem dostarczonego oprogramowania. Demonstracja PSBD ma także pokazać funkcjonalności i sprawność działania platformy PIBUK.

Szczegółowe założenia i wymagania dotyczące PSBD są zawarte w dokumentacji postępowania o zamówienie platformy dostępnej na stronie Instytutu Łączności.

Poniżej został przedstawiony PSBD na podstawie dokumentów przetargowych. Wyniki uzyskane w trakcie realizacji PSBD zostaną zaprezentowane w jednym z następnych biuletynów.

#### Dane wejściowe

Źródłem danych dla PSBD będą pliki tekstowe oraz baza danych MySQL. Źródłowe pliki tekstowe będą

w jednym z dwóch formatów: CSV lub KML. Zostaną wdrożone mechanizmy wczytywania źródłowych plików tekstowych dla ośmiu przykładowych plików: dla pliku CSV z zapisem badania QoS, dla pliku KML z zapisem badania QoS, 3 plików CSV z zapisem badania parametrów telefonii mobilnej i 3 plików KML z zapisem badania parametrów telefonii mobilnej. Wszystkie 4 rodzaje plików CSV mają zbliżoną do siebie strukturę, różnica występuje w kolumnach 5-10 i polega na zmianie nazwy zapisanego w tej kolumnie parametru liczbowego lub tekstowego. Źródłowe pliki tekstowe będą umieszczane w Repozytorium Danych Źródłowych Platformy w sposób nieregularny. PSBD będzie monitorować pojawienie się nowych plików. Harmonogram ich wczytywania będzie konfigurowany przez Operatora Eksploatacji Systemu.

#### Metadane

Wykonawca zapewni możliwość wykorzystania w PSBD następujących metadanych, które będą wersjonowane i będą definiowane przez uprawnionych Użytkowników:

1. Metadane pozwalające definiować źródła danych (pliki lub tabele i kolumny bazy danych) za pomocą masek nazw plików (np. wildcard lub wyrażenia regularne).
2. Metadane opisujące badania, których pomiary będą przedmiotem przetwarzania, w tym identyfikator źródła zawierającego pomiary oraz metadane identyfikujące jakiego operatora badanie dotyczy.
3. Słowniki wartości danych pobierane ze źródeł i słowniki wartości umieszczane w raportach.
4. Dane o poprawnych zakresach wartości danych pobieranych ze źródeł.
5. Dane o poprawnych dziedzinach danych pobieranych ze źródeł.

## Funkcje PSBD

PSBD zapewni zaawansowane funkcje wykrywania i obsługi anomalii w danych źródłowych i procesach ETL, w tym realizację poszczególnych funkcji, procedur, ścieżek przetwarzania i procesów ETL w zakresie czyszczenia danych źródłowych według zadanych reguł (usuwanie rekordów nieniosących informacji, usuwanie duplikatów, usuwanie kolejnych rekordów różniących się o wartość poniżej zadanego progu, uzupełnianie danych o wartości z rekordów poprzedzających je chronologicznie) oraz w zakresie wykrywania i obsługi wykrytych anomalii w danych źródłowych i procesach ETL. PSBD umożliwi realizację weryfikacji danych wejściowych z użyciem narzędzi ETL w zakresie obsługi wykrytych anomalii, poprzez obsługę braków konfiguracji w słownikach. W przypadku pojawienia się nowych niezdefiniowanych wartości w słownikach opisujących dane źródłowe (daną kolumnę) nowe wartości automatycznie zostaną dodane do słownika, a do Operatora Eksploatacji Systemu zostanie wysłany e-mail z listą dodanych wartości. W przypadku, gdy w słownikach brakuje danych niezbędnych do wykonania dalszego przetwarzania, zatrzymane zostaną tylko te części przetwarzania, które są od tego braku zależne. Następnie

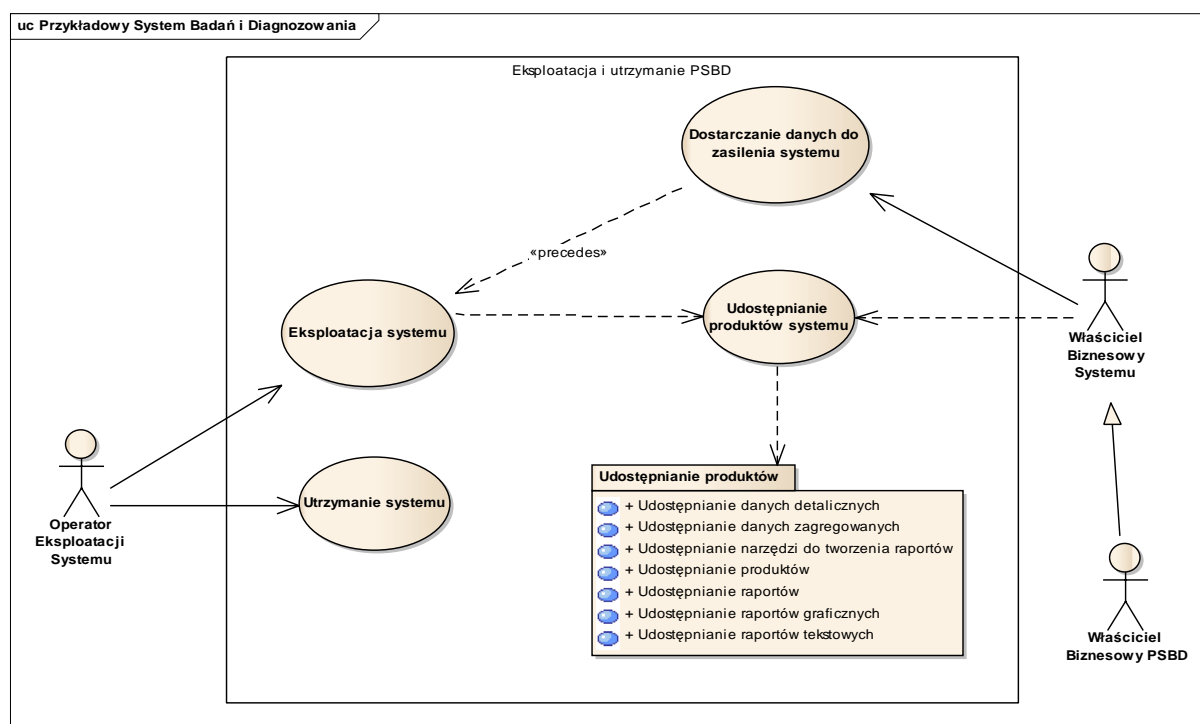
odbędzie się raportowanie wykrytej anomalii w ramach logów Systemu i wykonanie zdefiniowanej procedury obsługi napotkanie anomalii. PSBD oprócz danych detalicznych będzie zawierał agregaty danych.

## OLAP - Wymiary i miary

W PSBD będzie dostępna kostka OLAP zawierająca oczyszczone dane. W kostce zostaną zdefiniowane wymiary bazujące na danych źródłowych i metadanych. Wymiary dostępne w kostce zostaną wybrane na etapie Projektu. PSBD zapewni możliwość prowadzenia wielowymiarowych analiz danych z kostki OLAP (w tym możliwość agregacji, filtrowania, drążenia danych) oraz prezentacji wyników przeprowadzonych analiz w formie interaktywnych raportów dynamicznych oraz raportów statycznych prezentujących aktualny stan raportu dynamicznego.

## Raporty wyjściowe

PSBD umożliwi generowanie, na podstawie miar i wymiarów zawartych w kostce OLAP, raportów statycznych i dynamicznych zawierających zestawienia tabelaryczne, zestawienia graficzne w postaci wykresów, statyczne teksty, wizualizacje wartości danych detalicznych i zagregowanych na mapach.



Przypadki użycia systemu PSBD

Projekt realizowany przez