

NOTATKA WSTĘPNE KONSULTACJE RYNKOWE
„System Hurtownia Danych wraz z narzędziem raportowym jako zadanie w ramach Projektu
Transformacji Cyfrowej i Procesowej w obszarze DSM”, znak sprawy: CUI-ZZ.322.1.2021

Data:	26.03.2021r.
Miejsce/forma:	Konsultacje prowadzone za zgodą Stron na łączach video konferencji
Uwagi:	<p>Wstępne konsultacje rynkowe pozwolą na uzyskanie informacji w zakresie najlepszych i najkorzystniejszych: technicznie, technologicznie, organizacyjnie oraz ekonomicznie rozwiązań w zakresie przedmiotu konsultacji.</p> <p>Informacje o minimalnych środkach, które zamawiający podejmie w celu zapobieżenia zakłócenia konkurencji na etapie postępowania:</p> <ul style="list-style-type: none"> - upublicznienie informacji uzyskanych w trakcie konsultacji - upublicznienie informacji o podmiotach biorących udział w konsultacjach

PORUSZANE ZAGADNIENIA:

L.p.	Poruszane kwestie	Ustalenia
1.	Prezentacja firmy, wdrożenia referencyjne	<p>O firmie:</p> <ul style="list-style-type: none"> • Firma od 10 lat na rynku (od 2011r.); 12897 zrealizowanych projektów • Jest członkiem Grupy Orange • 3 miejsce na rynku Integratorów • 4 miejsce wśród Integratorów w Oświęcimie • Partner Oracle • Główne kierunki działalności: Duże kompleksowe projekty integracyjne dla miast Warszawa, Łódź - wytwarzanie oprogramowania wraz z tworzeniem infrastruktury <ul style="list-style-type: none"> - dedykowane Hurtownie Danych - e-usługi dla jednostek publicznych - rozwiązania BI • Wymagane zasoby osobowe po Stronie Zamawiającego: Zespół po stronie Zamawiającego powinien zapewnić po swojej stronie do współpracy odpowiednie zasoby ludzkie (Osoby dedykowane do realizacji poszczególnych strumieni tematycznych) – zgodnie z poniższymi etapami przygotowania do wdrażania Hurtowni Danych: <ol style="list-style-type: none"> 1. Analiza źródeł danych (duży wpływ na narzędzia i koszty) 2. Analiza potrzeb klienta (dane, raporty, wizualizacje) 3. Zdefiniowanie procesów biznesowych i danych z nimi związanych 4. Określenie jak dane mają być wykorzystywane (m.in. przedstawiciele właściciela danych) 5. Osoby odpowiedzialne za rozwiązania techniczne

w zakresie infrastruktury
Podstawą są dobrze komunikujące się ze sobą
Zespoły Architektoniczne powołane po stronie
Zamawiającego i Wykonawcy.

Wdrożenia referencyjne:

- **Wielkopolska Cyfrowa e-szkola**
- **Warszawa – Platforma elektronicznych Usług Miejskich**
Technologia oparta na rodzinie produktów ORACLE.
Składniki:
 - Magazyny Danych
 - Mikrousługi (Moduł Podatnika, Płatnika, e-dokumenty, Moduł karty miejskiej)
 - Systemy dziedziczone (CRM, JSP, 19115)
 - Moduł Administracji
 - Moduł bezpieczeństwa
 - API
 - Portal Warszawiaka
 - Karta Warszawiaka
- **IPS LOG – Usługi Elektroniczne dla Mieszkańca Łodzi**
IPS – Inteligentna Polityka Społeczna (świadczenia socjalne, opieka społeczna, rynek pracy, mieszkalnictwo, stypendia i zasiłki dla uczniów). Składa się z kilkudziesięciu modułów IPS (m.in. PZ, ePUAP)
Wdrożenie nowych usług i udoskonalenie istniejących usług na platformie elektronicznej
Wdrażanie i udoskonalanie e-usług z zakresu polityki społecznej.
Główne elementy Architektury:
 - Portal Mieszkańca
 - Integracja i wymiana danych (integrowanie różnych bazach danych, zabezpieczanie przed nadużyciami poprzez integrację i wymianę danych pomiędzy 5 głównymi jednostkami, robotyzacja procesów biznesowych)
 - Portal Pracownika
 - Analiza DanychUsługi zintegrowane z platformami e-PUAP, Węzeł Krajowy, platformą Ministerstwa Pracy i Polityki Społecznej. Usługi dostępu do informacji oraz analizy informacji. Rozwój usług elektronicznych oferowanych przez Centrum Obsługi Mieszkańca
Integracja z systemem Elektronicznego Zarządzania Dokumentami. Integracja z GIS.
Metody realizacji:
 - Budowa i wykorzystanie bazy wiedzy o usługach
 - Budowa miejskiego magazynu danych (SQL, NoSQL, dane przestrzenne)
 - Budowa mechanizmów ETL
 - Budowa narzędzi do analizy danych
 - Powtórne wykorzystanie danych
 - Przetwarzanie danych – oparte na Platformie Wymiany Danych (ETL) – głównie dane w postaci XML z jednostek:
 - Centrum Świadczeń Socjalnych (55+)
 - MOPS
 - PUP
 - Zarząd Lokali Mieszkalnych
 - Wydziały UMŁ

		<p>Strefa raportowania: kilkadziesiąt dashboardów raportowych dla Jednostek Miasta Łodzi. IPS w liczbach (wolumen danych z ostatnich 5 lat) Ponad 300 tys. klientów Ponad 100 tys. gospodarstw domowych 2 mln świadczeń Około 12 mln wypłat (m.in. świadczenia 500+)</p> <p>Weryfikacja tożsamości Mieszkańców: - ePUAP - integracja z Węzłem Krajowym</p> <p>Integracja skrzynek ESP (ePUAP) poszczególnych Jednostek z EZD (kierowanie dokumentów do odpowiednich jednostek miasta). Moduł generowania wniosków o zaświadczenia i generowania samych zaświadczeń. Generowanie dokumentów zgodne z Instrukcją Kancelaryjną i KPA.</p>
2.	Prezentacja systemu	<ul style="list-style-type: none"> •Prezentacja przykładów modeli Hurtowni Danych: <ul style="list-style-type: none"> Kontenerowa Hurtownia Danych Obszary zastosowania hurtowni w sferze publicznej: <ul style="list-style-type: none"> - Edukacja, - Polityka społeczna, - Infrastruktura - Zdrowie, - Majątek i finanse miejskie, - Rewitalizacja i inwestycje Mechanizm zarządzania danymi i pobierania wiedzy z danych - główne założenia techniczne: <ul style="list-style-type: none"> - Przetwarzanie różnych typów danych (w tym przestrzennych) - Wysoka niezależność architektury logicznej HD od fizycznej implementacji - Oszczędność zasobów Podstawowy koncept: <ul style="list-style-type: none"> - Uniwersalna baza danych MS SQL Server (oszczędność środków – licencje, duża ilość narzędzi do zarządzania, integracji (ETL), raportowania BI, narzędzia Data Science) - Technologia Polybase (obsługa zapytań do różnych źródeł za pomocą jednego silnika RDBMS) - Polybase jako kontenery danych - Big Data Cluster – obsługa wielkich zbiorów danych ustrukturyzowanych (SQL) i nieustrukturyzowanych (NoSQL) •Procesy ETL <ul style="list-style-type: none"> - W procesach ETL stosowane są zasady unifikacji danych umożliwiające późniejsze wykorzystanie danych za pomocą jednego typu zapytań języka SQL. Zastosowane narzędzia ETL: <ul style="list-style-type: none"> 1. Altova Map Force (dla plików XML, JSON itp.) Serwery Map Force <ul style="list-style-type: none"> - Flow Force – przepływy i przetwarzanie danych - Map Force Server - Style Version Server (dla plików XML) 2. MSSQL Integration Services

- Budowanie wzorców i schematów przetwarzania danych załadowanych do MMSQL Server, lub danych z Big Data Cluster podłączonych do serwera

- sekwencje kroków
- osadzanie opracowanych procesów na serwerze w celu zaprogramowania przetwarzania.

Bardzo dobry do zastosowania dla danych relacyjnych.

3. ETL Microsoft w chmurze – Azure Data Factory

Upload zadań opracowanych w SS Integration Services – do Azure Data Faktory umożliwia przeniesienie procesów ETL do środowiska chmurowego.

Typowy proces przetwarzania danych:

- Geokodowanie – nadawanie i odczytywanie atrybutów geolokalizacyjnych dla wszystkich informacji i danych, dla których jest to możliwe
- Anonimizacja danych – zastosowanie reguł zgodnych z wymogami RODO – umożliwienie publikacji danych jako „open-data”
- Modelowanie danych – dla celów przetwarzania i analizy danych – Budowanie specjalistycznych modeli: przestrzennych, czasowych, zależności, hierarchicznych, sieciowych (społecznych)
- Agregacja danych – zmierzająca do logicznej unifikacji i jednorodności danych na potrzeby raportowe – dane z wielu źródeł są ze sobą łączone w celu realizacji analiz wielowymiarowych i przekrojowych.
- Udostępnianie – wykorzystanie danych zgromadzonych w Hurtowni Danych: BI (Qlik, Power BI), data science,

- **Narzędzia raportowo – analityczne**

1. Qlik Sense

Często wykorzystywany przez analityków do realizacji analiz Ad-hoc. Nadaje się bardzo dobrze do samodzielnej pracy z danymi i raportami. Użytkownik ma większą swobodę operowania danymi (dodawanie własnych zmiennych, tworzenie swoistych modeli danych).

Działa w oparciu o pamięć in-memory.

Bardzo dobre narzędzie do manipulowania zmiennymi modelu analitycznego, w tym do podłączania wielu źródeł danych.

Większa elastyczność i wygoda dla analityków biznesowych.

Rozbudowany interfejs graficzny budowy raportowania (self-service)

Tworzenie modelu analitycznego na podstawie samodzielnie określanych przez użytkownika:

- wymiarów

- faktów

Możliwość korzystania z warstwy skryptowej.

- Qlik Sense Master Objects – możliwość utworzenia abstrakcyjnej warstwy obiektów biznesowych, bazujący na fizycznym schemacie danych.

Funkcjonalność budowania „narracji” (prezentacji) opartej na opracowanych wynikach wizualnych.

		<p>2. Power BI Praca na gotowych strukturach baz danych, mocno związana z fizycznym modelem hurtowni danych. Wydajniejszy w zakresie obsługi dużych zbiorów danych. Pozwala na szerokie spektrum transformacji danych, tworzenie własnych wymiarów Tworzenie wizualizacji, które są udostępniane dla innych użytkowników końcowych (nie potrzebują wykonywać dodatkowych działań na raportach). Doskonały w zakresie manipulowania wizualizacjami, bazującymi na gotowym modelu analitycznym. Power BI – przede wszystkim narzędzie chmurowe, ale również dostępne w wersji on-premise. Pozwala na tworzenie publicznie dostępnych dashboardów, publikowanych np. na stronach urzędu miasta. Edytor techniczny – wygodna praca nad wizualizacjami raportów Można osadzać skrypty innych języków (np. Python, R)</p> <p>3. Report builder / Report services – rozszerzenie Power BI lub jako wersja standalone. Budowa zaawansowanych raportów i obszarów raportowych.</p> <p>4. Altova StyleVision Przygotowanie wzorców dokumentów elektronicznych – na podstawie formularza usługi elektronicznej. Umożliwia zaprojektowanie wszystkich transformacji umożliwiających przetworzenie danych na konkretny dokument elektroniczny.</p> <p>5. Możliwe rozwiązania Software / hardware</p> <ul style="list-style-type: none"> - Kubernetes MSSQL Server Standard Edition – Docker Image Qlik – Kubernetes edition - Hypervisor (Windows Server na Hyper-V lub VM-Ware) MSSQL Enterprise Edition Power BI (zawarty w MS SQL Server) lub Qlik - Cloud Azure SQL Azure Data Factory Power BI lub Qlik Cloud - Scenariusze hybrydowe Np. hurtownia Danych on-premise + analizy danych w Power BI w O 365.
3.	Integracja z systemami	<ul style="list-style-type: none"> • Sposoby integracji danych: <ol style="list-style-type: none"> 1. Altova Map Force (dla plików XML, JSON itp.) 2. MSSQL Integration Services 3. Integracja w chmurze – Azure Data Factory
4.	Wymagania techniczne	<ul style="list-style-type: none"> • Czynniki i wymagania wpływające na wybór architektury rozwiązania • Kiedy korzystać z usług w chmurze • Elementy procesu tworzenia Hurtowni Danych • Nowoczesna Hurtownia Danych – architektura

		<ul style="list-style-type: none"> - Unikalny Model Danych (Analytical Data Repository) - Staging Area - Data Quality - ODS (On-line Data Stage) – warstwa dostępu on-line do danych źródłowych - ADR – Analytical Data Repository Ujednolicony model danych – zastosowanie m.in. słowników biznesowych ujednolicających dane do jednorodnego modelu (np. dane adresowe, dane klientów). Dane osobowe powinny być oddzielone od pozostałych danych (np. finansowych) - MDM – Master Data Management - Bazy tematyczne – Data Marty - Warstwa semantyczna – baza tematyczna in-memory - BI / Analityka danych (narzędzia raportowo – analityczne) • Kwestie RODO w Hurtowniach Danych • Nowoczesna Hurtownia Danych – Dodatkowe narzędzia
5.	Sposób licencjonowania i szacowany czas i koszt wdrożenia	<ul style="list-style-type: none"> • LICENCJONOWANIE: • Licencja Power BI on-premise wymaga wykupienia licencji na MSSQL Server Enterprise. • Wersja Licencji na MSSQL Server Enterprise z dodatkową opłatą pozwalającą na bezpłatne korzystanie z pełnej funkcjonalności Power BI. • Wkorzystanie Power BI on-premise wymaga licencji MS SQL Server Enterprise z wykupionym Software Assurance inną opcją jest korzystanie z Power BI w ramach licencji Office 365.

SPORZĄDZILI:
Jacek Tyburczy/Jerzy Sawczyszyn
.....
[imię i nazwisko]