Základní parametry projektu

|  |  |
| --- | --- |
| Název projektu  *Oficiální název, pod kterým bude projekt vedený, možno stanovit později* | DF Fabric - prověření infrastruktury pro datovou analytiku DF Partner |
| Cíle projektu  *Proč se má project implementovat a k čemu by měl vest – ideálně měřitelné cíle* | Navrhnout a nastavit architekturu lake house jako základ datové integrace pro datovou analytiku.  Nastavit načítání dat z provozního systému do data lake.  Navrhnout a vytvořit prezentační vrstvu pro 1. fázi zaměřenou na obchodní data.  Vybudovat první sadu data science / AI modelů pro podporu obchodního procesu.  Vytvořit sémantickou vrstvu pro Power BI.  Integrovat výstupy s aplikací DF Tablet. |
| Sponzor projektu a způsob financování  *Kdo je držitelem rozpočtu na projekt* | Pavel Drha – výkonný ředitel |
| Uživatelé výstupů  *Kdo bude koncovým příjemcem výstupů* | Útvar obchodu (p. Daněk) |
| Projektový tým  *Role a zodpovědnosti na projektu* | Jiří Krov – vedoucí analytického týmu, manažer projektu za DF partner  Martin Daněk – zástupce byznysu  Tomáš Břečka – systémový analytik, technologická stránka |
| Začátek a konec  *Požadovaný termín realizace projektu* | 3.4.2025 |
| Fáze projektu  *Pokud bude project realizován v několika fázích, jaké jsou* | Návrh architektury a ověření její funkčnosti a vlivu na provozní systém.  Vytvoření data lake a plnění daty.  Návrh a odsouhlasení nezávislé struktury prezentační vrstvy.  Vytvoření prezentační vrstvy a plnění daty.  Vytvoření sémantické vrstvy pro Power BI.  Návrh data science a AI modelů.  Testování a ladění funkčnosti data science a AI modelů.  Zpřístupnění API AI modelů. |
| Rozsah projektu  *Co se bude v rámci projektu řešit* | Integrovaná data:   * Bude definováno na následujícím workshopu.   Generování podnětů:   * Podnět na základě pravidlového systému. * Podnět na základě Data Science analýzy. * Podnět na základě generativní AI. * Logování a monitoring. |
| Součástí projektu není  *Co je z rozsahu projektu explicitně vyloučeno* | Přímá integrace do DF Tablet. |
| Kritéria úspěšnosti  *Co musí být splněno, aby byl projekt vyhodnocen jako úspěšný* | Splnění rozsahu projektu.  Zvolení správných Data Science modelů.  Spolehlivost Data Science modelů.  Spolehlivost komunikace s Data Science / AI modely.  Dostatečný rozsah a srozumitelná dokumentace řešení.  Předání do správy DF Partner tak, aby byl možný samostatný rozvoj a správa. |
| Rizika projektu  *Co může ohrozit úspěšnou realizaci projektu* | Aktuální funkčnost generativní AI – možné halucinování.  Kvalita dat ve zdrojových systémech.  Dopady na výkonnost zdrojového systému. |
| Vazby projektu  *Další probíhající projekty / aktivity, které projekt mohou ovlivnit* | Aktuálně nejsou známé. |
| Požadovaná součinnost  *Jaká aktivita je vyžadována na straně klienta* | Zajištění Open VPN přístupů do prostředí DF Partner.  Vysvětlení datových struktur a výběr dat pro 1. fázi.  Vytvoření subskripce MS Azure.  Nastavení prostředí v MS Azure podle instrukcí dolphin.  Provádění administrátorských změn na úrovni zdrojového SQL serveru.  Zajištění administrátorských přístupů k databázi zdrojového SQL serveru.  Návrh podnětů pro zpracování v rámci pilotního projektu.  Ladění a testování podnětů. |

Technické parametry projektu

|  |  |
| --- | --- |
| Zdrojové systémy   * *Znalost dat – ER model, dokumentace…* * *Způsob přístupu k datům* * *Počet tabulek* * *Objem zpracovávaných dat* |  |
| BI technologie   * *Databáze* * *ETL* * *Vizualizace* * *Modelovací nástroj* |  |
| Reporty a dashboardy   * *Počet / složitost* * *Distribuce* * *Zabezpečení* |  |
| Datová kvalita   * *Známé problémy* * *Požadavky* * *Nástroje* |  |
| Technická omezení   * *Co může projekt ovlivnit po technické stránce* |  |
| Zvolená architektura   * *Data lake* * *Lakehouse* * *Datový sklad* * *…..* |  |
| Cílový datový model   * *Dimenzionální* * *Relační* * *Data vault* |  |
| Frekvence načítání   * *Denně* * *Near-real time* * *Real-time* |  |
| Historizace dat   * *Jaká data budou historizována a jakým způsobem* |  |
| Úvodní načtení   * *Jaká historie dat bude načtená* |  |
| Časové okno pro ETL   * *Kdy je možné spouštět načítací procedury* |  |
| Výkonnostní parametry   * *Specifické požadavky na výkonnost systému* |  |
| Prostředí   * *Dev / test / prod* |  |
| Provoz   * *Zodpovědnosti* * *Podpora* * *SLA* |  |

|  |
| --- |
| Další důležité informace |