Analiza realizacji pierwszego etapu wdrożenia hurtowni danych Global Terrorism

Kraków, 10.12.2019

Autorzy opracowania:

Ania Gajda

Ania Gogól

Spis treści

[1. Wstęp 3](#_Toc32079455)

[1.1. Cele hurtowni „Global Terrorism” 3](#_Toc32079456)

[1.2. Terroryzm 3](#_Toc32079457)

[1.3. Zasady wdrażania hurtowni danych i raporty 4](#_Toc32079458)

[2. Harmonogram i organizacja projektu 4](#_Toc32079459)

[2.1. Lista zadań i harmonogram 4](#_Toc32079460)

[2.2. Struktura organizacyjna projektu 5](#_Toc32079461)

[2.2.1. Proponowana struktura organizacyjna projektu i schemat zespołu. 5](#_Toc32079462)

[2.2.2. Podział ról w zespole tworzącym i obsługującym hurtownię danych 5](#_Toc32079463)

[2.2.3. Nadzór projektu 6](#_Toc32079464)

[2.2.4. Kierownik projektu po stronie Bunny Hop 7](#_Toc32079465)

[2.2.5. Kierownik projektu po stronie YYY 7](#_Toc32079466)

[2.3. Obszary odpowiedzialności 7](#_Toc32079467)

[2.3.1. Obszar odpowiedzialności za dane źródłowe 7](#_Toc32079468)

[2.3.2. Obszar odpowiedzialności za integrację i uporządkowanie danych 7](#_Toc32079469)

[2.3.3. Obszar odpowiedzialności za nadawanie uprawnień 7](#_Toc32079470)

[2.4. Rodzaje dokumentów 7](#_Toc32079471)

[3. Wstęp do specyfikacji interfejsu hurtowni danych 8](#_Toc32079472)

[3.1.Model interfejsów 8](#_Toc32079473)

[3.2. Spis interfejsów 8](#_Toc32079474)

[4. Analiza wymagań biznesowych 8](#_Toc32079475)

[4.1. Wymagania funkcjonalne 8](#_Toc32079476)

[4.2. Wymagania niefunkcjonalne 9](#_Toc32079477)

[4.3. Przypadki użycia 9](#_Toc32079478)

[5. Aspekty techniczne 11](#_Toc32079479)

[5.1. Warstwa zasilania – interfejsy wymiany danych 11](#_Toc32079480)

[5.1.1. Szczegóły dotyczące warstwy zasilania interfejsów 11](#_Toc32079481)

[5.1.2. Szczegóły bazy danych z Wikipedii 11](#_Toc32079482)

[5.1.3. Szczegóły bazy danych Global Terrorism Database 11](#_Toc32079483)

[5.1.4. Porównanie obydwu źródeł danych 12](#_Toc32079484)

[5.1.5. Web scraping 12](#_Toc32079485)

[5.2. Metodyka zespolenia rozwoju i eksploatacji oraz zapewnienia jakości 13](#_Toc32079486)

[5.2.1. Dostępność infrastruktury w chmurze i jej wirtualizacja 13](#_Toc32079487)

[5.2.2. Narzędzia automatyzacji i zarządzania konfiguracją w centrum danych 13](#_Toc32079488)

[5.2.3. Planning 13](#_Toc32079489)

[5.3 Warstwa prezentacji i analizy 13](#_Toc32079490)

[5.3.1 Raporty dynamiczne (OLAP, SQL) 13](#_Toc32079491)

[6. Szczegóły dotyczące wdrożenia modułu „Hurtownia danych” w YYY 13](#_Toc32079492)

[6.1 OLAP: Zagrożone lokalizacje 13](#_Toc32079493)

[6.2 OLAP: Motyw a cel ataku 13](#_Toc32079494)

[6.3 OLAP: Rodzaj broni i straty 13](#_Toc32079495)

[7. Czas i koszty wdrożenia 13](#_Toc32079496)

[7.1. Czas wdrożenia i licencje 14](#_Toc32079497)

[7.2. Wynagrodzenie zespołu 15](#_Toc32079498)

# 1. Wstęp

## 

## 1.1. Cele hurtowni „Global Terrorism”

Celem stworzenia hurtowni danych o nazwie Global Terrorism jest łatwy i szybki dostęp do danych potrzebnych następnie do analizy dla potrzeb instytucji rządowych, którym przesyłane są wyniki w formie comiesięcznych raportów.

## 1.2. Terroryzm

Firma YYY pozyskuje dane o atakach terrorystycznych na całym świecie z wielu źródeł, aby następnie je oczyścić, uzupełnić i uporządkować..

Wiele sygnałów wskazuje, iż na polskim rynku informatycznym nadchodzi czas na hurtownie danych. Technologia ta obecna jest już w Polsce od kilku lat - przede wszystkim za sprawą pionierskiej działalności SAS Institute i jego partnerów wdrożeniowych.

## 1.3. Zasady wdrażania hurtowni danych i raporty

Jest kilka podstawowych zasad, na jakich oparty jest efektywny proces budowy hurtowni danych. Należą do nich:

* zapewnienie strategicznego kontekstu analiz dostarczanych przez hurtownię danych
* przyrostowa, iteracyjna budowa systemu
* oparcie budowy składników hurtowni danych na definicjach architektury
* ścisła współpraca z użytkownikami.

Celem raportów, w spojrzeniu biznesowym, jest…

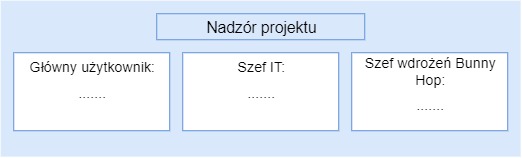
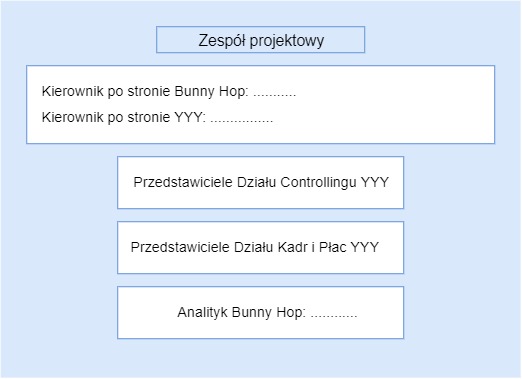
# 2. Harmonogram i organizacja projektu

## 2.1. Lista zadań i harmonogram

|  |  |  |  |
| --- | --- | --- | --- |
| **Nazwa zadania** | **Osoby odpowiedzialne** | **Data rozp. plan.** | **Data zak. plan.** |
| **Wdrożenie modułu Hurtowni Danych w YYY** | **Bunny Hop, YYY** | **17/01/2020** | **20/03/2020** |
| Prace analityczne | Bunny Hop | 19/01/2020 | 02/02/2020 |
| Specyfikacja interfejsów | Bunny Hop | 19/01/2020 | 24/01/2020 |
| Analiza przedwdrożeniowa | Bunny Hop | 28/01/2020 | 02/02/2020 |
| Prace wdrożeniowe |  |  |  |
| Wykonanie interfejsów |  |  |  |
| Interfejs X |  |  |  |
| Obszar Q |  |  |  |
| Arkusz Excel z mapowaniem czego? |  |  |  |
| Wykonanie kostki Z |  |  |  |
| Wykonanie kostki X |  |  |  |
| Wykonanie kostki W |  |  |  |
| Testy kostki Z |  |  |  |
| Obszar E |  |  |  |
| Arkusz Excel z definicjami pozycji |  |  |  |
| Integracja definicji pozycji kilku |  |  |  |
| Wykonanie kostki ZZZ |  |  |  |
| Testy kostki ZZZ |  |  |  |

## 2.2. Struktura organizacyjna projektu

### 2.2.1. Proponowana struktura organizacyjna projektu i schemat zespołu.



### 

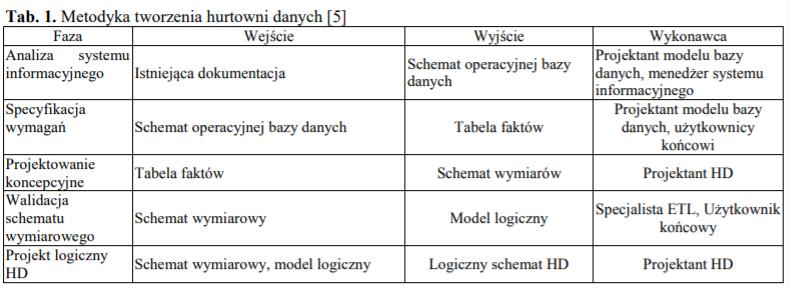
### 2.2.2. Podział ról w zespole tworzącym i obsługującym hurtownię danych

Hurtownia danych nie jest produktem gotowym, który można kupić „z półki”. Budowa hurtowni danych wymaga powołania zespołu tworzącego hurtownię oraz jej eksploatację. Osoby zaangażowane w proces tworzenia i eksploatacji hurtowni danych to między innymi:

* administrator hurtowni danych,
* kierownik zespołu tworzącego hurtownię danych,
* analityk biznesowy,
* architekt systemowy,
* specjalista ETL,
* projektant interfejsu użytkownika,
* specjalista OLAP,
* projektant modelu danych,
* trener użytkowników końcowych.

Administrator HD jest odpowiedzialny za integrację i koordynację metadanych i danych z różnych źródeł danych oraz za zarządzanie źródłami danych, fizyczny projekt BD, tworzenie kopii zapasowych i odzyskiwania bazy po awarii, bezpieczeństwa, wydajności i strojenia. Kierownik zespołu projektowego jest odpowiedzialny za zarządzanie zespołem, za przestrzeganie wymagań biznesowych i koordynację planów strategicznych, za rozwój projektu hurtowni danych, oraz za wybór członków zespołu projektującego HD o określonych kompetencjach i umiejętnościach do wykonywania zadań Analityk biznesowy jest odpowiedzialny za określenie jakie informacje z hurtowni danych są potrzebne dla celów zarządzania. Architekt systemowy jest odpowiedzialny za rozwój i wdrażanie infrastruktury technicznej HD, od konfiguracji sprzętowej i programowej serwera bazy danych po stanowisko użytkownika końcowego. Specjalista ETL jest odpowiedzialny za pobieranie danych ze źródeł (ekstrakcję danych), dokonywaniem przekształceń, unifikacji i czyszczenia danych (transformację danych) i umieszczenia danych w strukturach hurtowni (ładowanie danych). Projektant interfejsu użytkownika jest odpowiedzialny za interfejs użytkownika końcowego na komputerach stacjonarnych w modelu klient-serwer, witryny internetowej i projektowania interfejsu na urządzenia mobilne. Specjalista OLAP jest odpowiedzialny za tworzenie wielowymiarowych kostek i widoków danych, umożliwiający przeprowadzenie bieżącej analizy, Projektant modelu danych jest odpowiedzialny za opracowanie modelu istniejących danych w organizacji oraz za tworzenie schematu odpowiedniego do analizy OLAP. Trener użytkownika końcowego jest odpowiedzialny za przeprowadzenie szkoleń dla końcowych użytkowników, aby mogli w efektywny sposób przeprowadzić analizę OLAP.

Metodyka tworzenia hurtowni danych:



### 2.2.3. Nadzór projektu

Ścisła współpraca z użytkownikami jest podstawowym zadaniem w każdym projekcie informatycznym, jednak w przypadku budowy hurtowni danych wzrasta zarówno rola, jak i wyzwanie niesione przez ten postulat. "Ścisła współpraca" to w tym przypadku postulat dotyczący użytkownika, którego czas jest wyjątkowo cenny. Użytkownik ten jest dla nas źródłem dwóch rodzajów wymagań.

1. Przegląd dokumentów projektu: harmonogramu, notatek ze spotkań, raportów okresowych, analizy, rejestru zagrożeń, kluczowych procedur
2. Podejmowanie decyzji w sytuacji, gdy kierownik projektu sygnalizuje opóźnienie zadań, za które odpowiada

### 2.2.4. Kierownik projektu po stronie Bunny Hop

* Koordynowanie zadań wykonywanych po stronie Bunny Hop
* Uczestniczenie we wszystkich spotkaniach roboczych w YYY a w sytuacji gdy jest to niemożliwe wyznaczanie zastępców
* Tworzenie i aktualizacja dokumentacji projektu: harmonogramu, notatek ze spotkań, raportów okresowych, analizy
* Ścisła współpraca z kierownikiem po stronie YYY

### 2.2.5. Kierownik projektu po stronie YYY

* Koordynowanie zadań wykonywanych po stronie YYY
* Koordynowanie zadań wykonywanych po stronie dostawców systemów IT, z którymi łączyć się będzie hurtownia
* Organizowanie spotkań roboczych w YYY oraz uczestniczenie w nich, a w sytuacji gdy jest to niemożliwe wyznaczanie zastępców
* Ścisła współpraca z kierownikiem po stronie Bunny Hop

## 2.3. Obszary odpowiedzialności

Kluczowe dla projektu wdrożenia hurtowni danych jest wyodrębnienie następujących obszarów odpowiedzialności:

### 2.3.1. Obszar odpowiedzialności za dane źródłowe

Odpowiedzialność po stronie KLIENTA lub dostawców systemów źródłowych. Dotyczy jakości i rzetelności danych wprowadzanych do systemów, z których później przekazywane są do hurtowni. Hurtownia danych poprzez reguły walidacyjne zapewni odpowiednią jakość danych.

### 2.3.2. Obszar odpowiedzialności za integrację i uporządkowanie danych

Odpowiedzialność po stronie Bunny Hop. Dotyczy wszelkiej obróbki danych źródłowych do postaci, w której są one przetwarzane i prezentowane użytkownikom końcowym.

### 2.3.3. Obszar odpowiedzialności za nadawanie uprawnień

Odpowiedzialność po stronie KLIENTA. Dotyczy przyporządkowywania użytkowników końcowych do ról, bądź administrowania struktura organizacyjną z której wynikają uprawniania dynamiczne

## 2.4. Rodzaje dokumentów

Głównym dokumentem projektu jest analiza przedwdrożeniowa, która jest cały czas uaktualniana. Pozostałe dokumenty:

**Notatki ze spotka**

Każde spotkanie jest dokumentowane notatką wysyłaną mailem do uczestników oraz kierownictwa projektu

**Raporty miesięczne**

Do 5 dnia kolejnego miesiąca sporządzany jest raport miesięczny, który zawiera:

* Omówienie harmonogramu (co się udało zrobić, czego się nie udało),
* Aktualizację analizy przedwdrożeniowej (wraz z harmonogramem),
* Listę kluczowych zagrożeń dla projektu (rejestr zagrożeń).

**Rejestr zagrożeń**

Lista spraw, które mogą zagrozić terminowej realizacji projektu. Każdej ze spraw przypisywana jest waga i stopień prawdopodobieństwa wystąpienia. W przypadku zagrożeń przekraczających określony próg, informowany jest nadzór projektu.

# 3. Wstęp do specyfikacji interfejsu hurtowni danych

W przypadku wszystkich systemów działających w ramach XXX S.A. wymiana danych będzie miała na razie charakter jednostronny – dane będą przepływać z systemów zewnętrznych do hurtowni danych. Dodatkowo interfejs analityczny powinien być możliwie wygodny w użyciu i łatwy w obsłudze także dla osób, które mogą nie być bezpośrednio związane z informatyką, ale za to potrafiące bardzo dobrze analizować informacje przechowywane w tym systemie.

## 3.1.Model interfejsów

Dane będą wymieniane w ramach interfejsów. Pod pojęciem tym kryje się sposób, rozumiany jako ściśle określony zestaw reguł i ich opisów, w jaki programy komputerowe komunikują się między sobą. Zadaniem interfejsu programowania aplikacji jest dostarczenie odpowiednich specyfikacji podprogramów, struktur danych, klas obiektów i wymaganych protokołów komunikacyjnych. W ramach danego interfejsu dane będą:

* Udostępniane w postaci tabelarycznej jako odpowiedź zapytań SQL
* Udostępniane w postaci raportów
* Segregowane według kryteriów wybranych przez konkretnego użytkownika
* Importowane na serwer hurtowni danych (za pomocą usług SSIS)
* Przetwarzane i wstawiane do bazy hurtowni danych

## 3.2. Spis interfejsów

W ramach integracji z systemami w XXX S.A.:

# 4. Analiza wymagań biznesowych

## 4.1. Wymagania funkcjonalne

1. tworzenie raportów o procentowym udziale ataków na danym terytorium (kontynencie, stanie) kraju
2. tworzenie raportów o czasie dokonywanych ataków
3. tworzenie raportów o skali wybuchu, ilości poniesionych strat (materialnych i osobowych) do jednego zamachu
4. możliwość przeglądania danych ze względu na parametry
5. możliwość sortowania danych według kryteriów
6. możliwość archiwizacji danych
7. wspomaganie decyzji
8. agregacja wielu zamachów przy braku jednolitej struktury
9. wsparcie dla zewnętrznych systemów np po przez dobieranie strategii do walki z terroryzmem

## 4.2. Wymagania niefunkcjonalne

1. Tworzenie raportu nie może trwać dłużej niż 20 sekund
2. Raporty powinny tworzyć się płynnie, bezproblemowo
3. Raporty tworzone są przez program Excel

## 4.3. Przypadki użycia

1. **Przypadek użycia:**

* Utworzenie raportu o procentowym udziale ataków na danym terytorium:

**Aktorzy:**

* Użytkownik, System

**Scenariusz Główny:**

* Użytkownik zaznaczając wybiera z listy wyboru interesujące go dane:
* Country
* Region, State, City
* Latitude, Longitude
* Location
* Summary
* Success
* Suicide
* Attactype
* Target, Target subtype
* Motive
* Weapon type, Weapon subtype, Weapon detail
* Hostiges
* Loss
* System wyświetla zakres dat (dzień-miesiąc-rok) o atakach terrorystycznych.
* Użytkownik wybiera zakres dat z którego chce utworzyć raport.
* System pyta użytkownika czy chce dodatkowo posortować dane-rekordy.
* Użytkownik nie wybiera sortowania danych-rekordów.
* System wyświetla zakładkę z jednostką, w której ma być wykonany raport (wartość liczbowa, lub procenty).
* Użytkownik ustala jednostkę w tym przypadku procenty w której generowany jest raport ataków.
* System wyświetla poglądowy raport utworzony przez użytkownika.
* Użytkownik sprawdza czy przykładowy raport spełnia jego oczekiwania.
* System wyświetla okienko z akceptacją lub odrzuceniem raportu.
* Użytkownik akceptuje utworzony raport.
* Użytkownik z zakładki plik klika utwórz raport.

**Scenariusz alternatywny:**

* Użytkownik zaznaczając wybiera z listy wyboru interesujące go dane:
* Country
* Region, State, City
* Latitude, Longitude
* Location
* Summary
* Success
* Suicide
* Attactype
* Target, Target subtype
* Motive
* Weapon type, Weapon subtype, Weapon detail
* Hostiges
* Loss
* System wyświetla zakres dat (dzień-miesiąc-rok) o atakach terrorystycznych.
* Użytkownik wybiera zakres dat z którego chce utworzyć raport.
* System wyświetla komunikat o błędzie, zwracając wiadomość, że w tym okresie nie wystąpił żaden atak terrorystyczny.

1. **Przypadek użycia:**

* Przeglądanie danych ze względu na różne parametry:

**Aktorzy:**

* Użytkownik, System

**Scenariusz Główny:**

* Użytkownik otwiera hurtownię danych.
* Użytkownik wybiera z panelu widoku opcję przeglądania danych.
* System wyświetla formularz do wyboru.
* Użytkownik wybiera dane, przeklikując cały formularz i zaznaczając interesujące go rekordy.
* Użytkownik po wyborze klika akceptuj na końcu formularza.
* System wyświetla zaznaczone przez użytkownika rekordy.

1. **Przypadek użycia:**

* Archiwizacja danych i raportów:

**Aktorzy:**

* Użytkownik, System

**Scenariusz Główny:**

* Użytkownik otwiera hurtownię danych.
* Użytkownik wybiera z panelu widoku opcję archiwizowania danych.
* System wyświetla zakres danych do zaznaczenia, które użytkownik chce zarchiwizować
* Użytkownik wybiera dane zaznaczając je po przez kliknięcie w interesujące go rekordy.
* System pyta użytkownika czy chce zarchiwizować zaznaczone dane.
* Użytkownik akceptuje wybór.
* System pyta użytkownika w jakim formacie chce zapisać dane.
* Użytkownik wybiera rozszerzenie obsługiwane przez aplikację.
* System generuje raport historyczny.

# 5. Aspekty techniczne

## 5.1. Warstwa zasilania – interfejsy wymiany danych

### 

### 5.1.1. Szczegóły dotyczące warstwy zasilania interfejsów

Będziemy korzystać z dwóch stron do pobierania interesujących nas danych. Jedna jest uzupełniana raz do roku, jednak jest obszerniejsza w szczegółowe dane. Druga jest uzupełniana codziennie jednak ma bardziej ogólne dane.

Do pobierania danych będziemy używać web scrapingu, w tym przypadku wykorzystamy darmową wtyczkę do przeglądarki Chrome – Web Scraper.

### 5.1.2. Szczegóły bazy danych z Wikipedii

Będziemy korzystać ze strony <https://en.wikipedia.org/wiki/List_of_terrorist_incidents#1970%E2%80%93present>, podstrony Wikipedii. Są na niej dane incydentów terrorystycznych, które nie zostały przeprowadzone przez państwo lub jego siły. Zabójstwa są wymienione na liście zamordowanych ludzi.

Definicje terroryzmu są różne, dlatego wymienione tu incydenty są ograniczone do tych, które są godne uwagi i opisane jako „terroryzm” na podstawie konsensusu wiarygodnych źródeł.

### 5.1.3. Szczegóły bazy danych Global Terrorism Database

Dodatkowo połączymy dane z Wikipedii z danymi z darmowej bazy danych ze strony <https://www.start.umd.edu/gtd/>. Global Terrorism Database (GTD) ™ to najobszerniejsza na świecie niesklasyfikowana baza danych ataków terrorystycznych. Narodowe konsorcjum do spraw badań nad terroryzmem i odpowiedziami na terroryzm (START) udostępnia GTD ™ za pośrednictwem tej strony, aby zwiększyć zrozumienie przemocy terrorystycznej, aby można ją było łatwiej studiować i pokonać. GTD ™ jest produkowany przez oddany zespół naukowców i pracowników technicznych, wspieranych przez absolwentów naukowców. Zarządzanie bazą danych zapewnia rada doradcza, w skład której wchodzą jedni z najbardziej szanowanych naukowców w dziedzinie badań nad terroryzmem.

GTD ™ to baza danych o otwartym kodzie źródłowym, która zawiera informacje o krajowych i międzynarodowych wydarzeniach terrorystycznych na całym świecie od 1970 r., A teraz obejmuje ponad 190 000 wydarzeń. Dla każdego wydarzenia dostępny jest szeroki zakres informacji, w tym data i miejsce incydentu, zastosowana broń, charakter celu, liczba ofiar oraz - jeśli to możliwe - odpowiedzialna grupa lub osoba.

Aktualne źródło, zawierające informacje o incydentach o podłożu terrorystycznym, które miały miejsce we wszystkich krajach świata, na przestrzeni lat 1970 -2016.  Zawiera dużo szczegółowych informacji na temat każdego rekordu, takich jak motyw ataku, rodzaj użytej broni, cel, liczba ofiar itd.

Baza danych GTD jest uzupełniana raz do roku, jednak zawiera dużo dodatkowych informacji takich jak na przykład: szczegółowy cel ataku, motyw, typ użytej broni, szczegóły u użytej broni, zakładnicy, straty.

Na dzień dzisiejszy posiadamy bazę danych GTD do końca 2018 roku, która będzie uzupełniana raz w roku. W przeciwieństwie do bazy danych pobieranej z Wikipedii aktualizowanej codziennie, która będzie naszą główną bazą danych.

### 5.1.4. Porównanie obydwu źródeł danych

Z Wikipedii pobieramy dane:

1. data,
2. typ ataku,
3. martwi (plus ile martwych należało do organizacji terrorystycznej),
4. poszkodowani,
5. lokalizacja,
6. opis,
7. sprawca,
8. czy atak należy do bardziej ogólnej akcji i jakiej.

Z GTD pobieramy dane:

1. id,
2. rok, miesiąc, dzień,
3. państwo,
4. region, stan, Miasto,
5. dokładne dane geograficzne,
6. lokacja,
7. podsumowanie,
8. sukces,
9. samobójstwo,
10. typ ataku,
11. cel,
12. motyw,
13. szczegóły na temat użytej broni,
14. zakładnicy,
15. straty.

### 5.1.5. Web scraping

Do pobierania danych będziemy korzystać z Web Scrapera, darmowego rozszerzenia przeglądarkowego Chrome. Jest to narzędzie do ekstrakcji danych z witryny internetowej. Bez oprogramowania do pobrania, nie jest potrzebne Python / PHP / JS.

Korzystając z tego rozszerzenia, możesz stworzyć plan (mapę witryny), w jaki sposób należy przeglądać witrynę internetową i co należy wyodrębnić. Za pomocą tych map witryn Skrobak internetowy będzie odpowiednio nawigować w witrynie i wyodrębniać wszystkie dane. Zeskrobane dane można później wyeksportować jako CSV.

## 5.2. Metodyka zespolenia rozwoju i eksploatacji oraz zapewnienia jakości

W naszym projekcie korzystamy z DevOps – metodyka zespolenia rozwoju (ang. development) i eksploatacji (ang. operations) oraz zapewnienia jakości (ang. quality assurance). Metodyka ta kładzie nacisk na ścisłą współpracę i komunikację profesjonalistów z zakresu utrzymania IT (administratorów) oraz specjalistów od rozwoju oprogramowania (programistów). Uwzględnia współzależność rozwoju i utrzymania IT. Skraca czas wdrożenia funkcji w oprogramowaniu.

Metoda rozwoju oprogramowania DevOps jest wskazana dla firm, w których częstotliwość edycji jest stosunkowo wysoka. W naszym projekcie w zależności od wydarzeń ze świata może dojść do realizacji jednego wdrożenia dziennie.

### 5.2.1. Dostępność infrastruktury w chmurze i jej wirtualizacja

### 5.2.2. Narzędzia automatyzacji i zarządzania konfiguracją w centrum danych

### 5.2.3. Planning

## 5.3 Warstwa prezentacji i analizy

### 5.3.1 Raporty dynamiczne (OLAP, SQL)

# 6. Szczegóły dotyczące wdrożenia modułu „Hurtownia danych” w YYY

## 6.1 OLAP: Zagrożone lokalizacje

## 6.2 OLAP: Motyw a cel ataku

## OLAP: Rodzaj broni i straty

# 7. Czas i koszty wdrożenia

System budowany w cyklu przyrostowym wymaga dobrze zdefiniowanej architektury, inaczej niemożliwe jest zachowanie kontroli nad jego rozwojem. Chaotyczna rozbudowa hurtowni danych, brak punktu odniesienia umożliwiającego zaplanowanie kolejnych przyrostów i modyfikacji sprawia, że adaptacja systemu do zmian biznesowych staje się coraz trudniejsza i bardziej kosztowna.

W przypadku hurtowni danych definicja architektury obejmuje dokumentację strategii biznesowych, korporacyjny model danych, dokumentację systemów źródłowych, model systemu zasilania i technologii wykorzystywanej do implementacji i wdrożenia systemu. Model technologii jest szczególnie istotny w przypadku zastosowań korporacyjnych, wymagających integracji rozwiązań różnych dostawców - od środowiska raportowania i analiz interakcyjnych, (OLAP) poprzez bazę danych dla korporacyjnej hurtowni danych, do systemu transformacji i czyszczenia danych źródłowych.

W dłuższej perspektywie zdefiniowanie architektury zwiększa wydajność procesu budowy hurtowni dzięki temu, że w poszczególnych iteracjach zminimalizowany jest obszar prac koncepcyjnych na korzyść implementacji funkcjonalności rozwiązania.

W zależności od wielkości przedsięwzięcia, wdrożenie hurtowni danych może trwać od kilku tygodni do kilkunastu miesięcy. Warto zauważyć, że większość przedsiębiorstw, po wdrożeniu systemu widząc szereg korzyści i zalet z zaimplementowanego rozwiązania, zaczyna dostrzegać nowe możliwości rozwoju i decyduje się na rozszerzenie systemu o kolejne obszary i funkcjonalności.

## 7.1. Czas wdrożenia i licencje

W przypadku hurtowni danych Global Terrorism DW szacujemy, że wdrożenie jej zajmie 2 miesiące (8 tygodni). W tym czasie firma Bunny Hop będzie korzystała z narzędzi:

|  |  |  |  |
| --- | --- | --- | --- |
| **Nazwa** | **Czas użytkowania (w tyg.)** | **Cena (PLN)** | **Cel** |
| Microsoft Azure | 8 | - | zestaw usług do przetwarzania w chmurze |
| SQL Server | 8 | 55658,28 | zarządzanie relacyjnymi bazami danych |
| Integration Services | 8 | - | do budowy przepływów, integrujących dane z różnych systemów |
| Data Quality Services | 8 | - | do badania jakość danych i budowy automatycznych mechanizmów czyszczących |
| Data Master Services | 8 | - | dzięki któremu użytkownicy mogą z poziomu dedykowanego interfejsu webowego, czy też Excela, zarządzać tzw. master danymi |
| SQL Server Analysis Services | 4 | - | narzędzie do tworzenia i zarządzania hurtowniami danych |
| SQL Server Reporting Services | 4 | - | tworzenie prostych raportów |
| Power BI Premium | 4 | 19501,48 | zaawansowana analiza, obsługa danych big data oraz raportowanie w środowisku lokalnym i chmurze |
| Microsoft Office 2019 Professional Plus | 8 | 1249,00 | oficjalne dokumenty firmowe, Word, Excel |
| Visual Studio Professional | 8 | 175,69 (msc) | środowisko Visual Studio Professional IDE |
| **Suma cen licencji:** | | ***76,760.14*** | |

## 7.2. Wynagrodzenie zespołu

Menadżer działu IT: 22000 zł brutto/msc

Szef wdrożeń: 20000 zł brutto/msc

Główny Analityk: 18000 zł brutto/msc

Deweloperzy: 30000 zł brutto/msc

**Wynagrodzenie zespołu za cały okres pracy przy projekcie (2 miesiące)**: 180,000 zł brutto