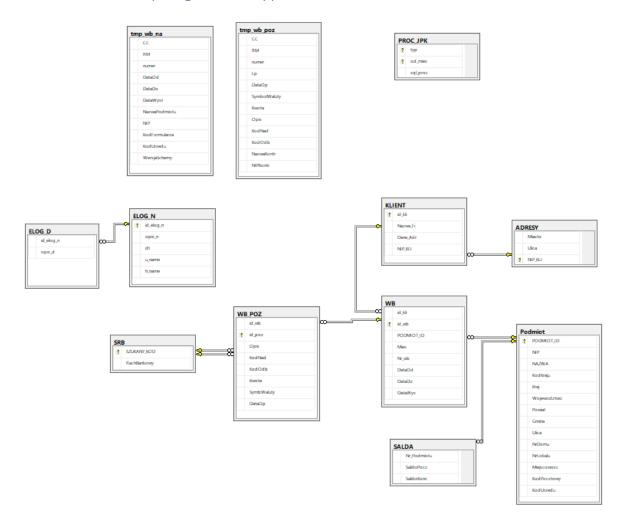
Sprawozdanie końcowe z projektu

Karol Sekściński 319093

03.04.2024

Struktura tabel (Diagram bazy)



Opis tabel

Tabele tmp_wb_na i tmp_wb_poz nie są powiązane z innymi tabelami. Służą tylko i wyłącznie do transferu danych z plików wejściowych do tych tabel. Zawsze podczas transferu są wstępnie czyszczone (usuwane są z nich wszystkie wiersze) a następnie zawartość plików wejściowych jest transportowana do odpowiednich tabel.

Tabela PROC_JPK odpowiada za rozwiązanie problemu obowiązywanie różnych formatów JPK_WB dla różnych miesięcy. Jest to tabela pomocnicza, która umożliwia działanie procedury 'RUN_JPK'.

Tabele ELOG_N i ELOG_D odpowiadają za logi błędów. ELOG_N zawiera informacje o nagłówku błędu (komu i gdzie wydarzył się błąd), natomiast ELOG_D odpowiada za detale błędu.

Tabela SRB jest słownikiem rachunków bankowych, w danych wejściowych znajdują się odpowiednie kody, które odpowiadają rachunkom bankowym zarówno dla nadawcy określonego przelewu jak i odbiorcy określonego przelewu.

Tabela ADRESY jest to tabela pomocnicza będąca książką adresową dla klientów, zawiera informacje o tym w jakim mieście oraz na jakiej ulicy znajduje się odbiorca przelewu.

Tabela SALDA jest to tabela przechowująca informacje o saldzie początkowym i końcowym dla danego podmiotu.

Tabela KLIENT zawiera informacje o klientach, których dotyczą transakcje wyciągu bankowego.

Tabela Podmiot zawiera informacje o podmiocie, który wykonuje określone przelewy zawarte w wyciągu bankowym takie jak dokładny adres podmiotu.

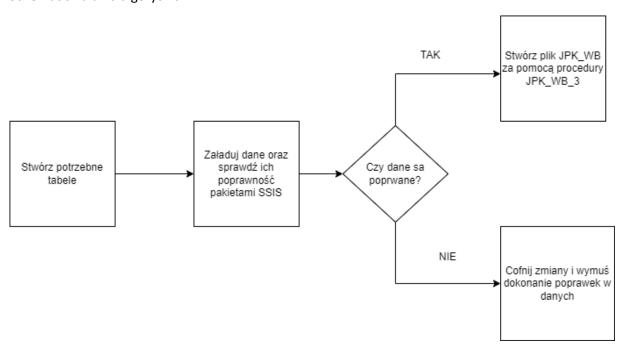
Tabela WB jest tabela docelowa dla danych wejściowych, które najpierw trafiają do tabeli tmp_wb_na.

Tabela WB_POZ jest tabela docelowa dla danych wejściowych, które najpierw trafiają do tabeli tmp_wb_poz.

Opis algorytmu

System realizuje transfer, przetworzenie i walidacje otrzymanych danych oraz generowanie raportu w formacie XML określonym przez MF w Polsce. System został wykonany w postaci skryptu MS SQL "zaliczenie.sql" oraz pakietów SSIS WB_NA oraz WB_POZ, całość załączona w załączniku.

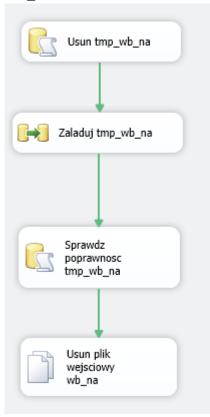
Schemat działania algorytmu



Opis pakietu

System składa się z dwóch pakietów SSIS wygenerowanych za pomocą narzędzi "Import and Export Data (64-bit) oraz SQL Server Data Tools 2008. Pakiety to: WB_NA oraz WB_POZ, odpowiadających za wczytanie odpowiednio nagłówków (zawierających ogólne informacje o dokumencie) oraz pozycji wyciągu (zawierających szczegółowe informacje o każdej transakcji)

WB_NA



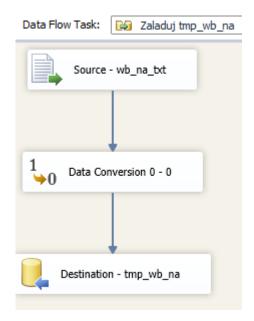
Pierwszy box (Usun tmp wb na) usuwa dane z tabeli tymczasowej: DELETE FROM [dbo].[tmp wb na].

Drugi box (Zaladuj tmp_wb_na) przesyła dane z pliku źródłowego do tabeli w bazie. Detale są widoczne poniżej.

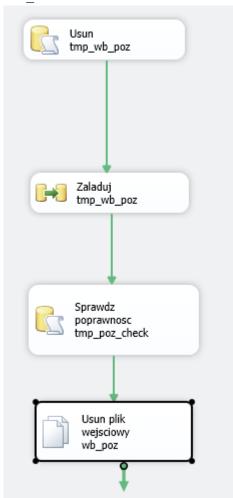
Trzeci box (Sprawdz poprawność tmp_wb_na) wywołuje procedurę bazodanowa, która waliduje poprawność danych nagłówkowych: *EXEC tmp_na_check*

Czwarty box (Usun plik wejściowy wb_na) kasuje plik wejściowy z danymi źródłowymi.

Szczegóły box'u nr 2 (Zaladuj tmp_wb_na), pokazujące, ze najpierw następuje odczytanie pliku w formacie Text Tab Delimited UTF-8, następnie następuje konwersja danych na odpowiednie typy danych odpowiadające tym z tabeli tmp_wb_na a następnie dana są importowane do tabeli wymienionej wcześniej.



WB_POZ



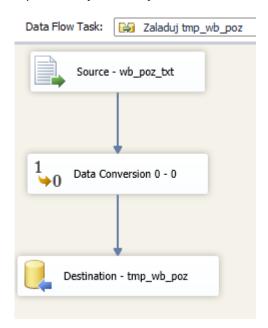
Pierwszy box (Usun tmp_wb_poz) usuwa dane z tabeli tymczasowej: DELETE FROM [dbo].[tmp_wb_poz].

Drugi box (Zaladuj tmp_wb_poz) przesyła dane z pliku źródłowego do tabeli w bazie. Detale są widoczne poniżej.

Trzeci box (Sprawdz poprawnosc tmp_wb_na) wywołuje procedurę bazodanowa, która waliduje poprawność danych pozycji wyciagu: *EXEC tmp_poz_check*

Czwarty box (Usun plik wejsciowy wb_na) kasuje plik wejściowy z danymi źródłowymi.

Szczegóły box'u nr 2 (Zaladuj tmp_wb_poz), pokazujące, ze najpierw następuje odczytanie pliku w formacie Text Tab Delimited UTF-8, następnie następuje konwersja danych na odpowiednie typy danych odpowiadające tym z tabeli tmp_wb_poz a następnie dana są importowane do tabeli wymienionej wcześniej.



Analiza walidacji danych

Typy danych

wb_na.txt

Dane wejściowe powinny być w formacie Text Tab Delimited UTF-8. W pierwszym wierszu pliku powinny znajdować się kolumny takie jak:

- CC Dopuszczalna długość (10) Company Code Identyfikator Podmiotu np. PL00
- RM Dopuszczalna długość (6) Rok Miesiąc Identyfikator Roku i Miesiąca np. 202403
- numer Dopuszczalna długość (5) Numer nagłówka np. XX1
- DataOd Dopuszczalna długość (10) Data od której obowiązuje dany nagłówek np. 01.03.2024
- DataDo Dopuszczalna długość (10) Data do której obowiązuje dany nagłówek np.
 31 03 2024
- DataWyst Dopuszczalna długość (10) Data wystawienia danego nagłówka np. 03.04.2024
- NazwaPodmiotu Dopuszczalna długość(100) Nazwa klienta dla którego podmiot wykonał transakcje np. PW
- NIP Numer NIP Klienta Dopuszczalna długość (20) np. PL55555555555
- KodFormularza Dopuszczalna długość (20) np. JPK_WB
- KodUrzedu Dopuszczalna długość (20) np. 0001
- WersjaSchemy Dopuszczalna długość (20) np. 1.0

wb poz.txt

Dane wejściowe powinny być w formacie Text Tab Delimited UTF-8. W pierwszym wierszu pliku powinny znajdować się kolumny takie jak:

- ➤ Lp Dopuszczalna długość (10) Liczba Porządkowa Dopuszczalna długość(10) np. 1
- CC Dopuszczalna długość (10) Company Code Identyfikator Podmiotu np. PL00
- RM Dopuszczalna długość (6) Rok Miesiąc Identyfikator Roku i Miesiąca np. 202403
- DataOp Dopuszczalna długość (10) Data Operacji np. 04.03.2024
- SymbolWaluty Dopuszczalna długość (10) np. PLN
- Kwota Dopuszczalna długość (20) Kwota transakcji np. 100
- Opis Dopuszczalna długość (100) Opis transakcji np. Gruszki
- KodNad Dopuszczalna długość (20) Kod Nadawcy (musi znajdować się w systemie w tabeli SRB) odpowiadający konkretnemu rachunkowi bankowemu np. ABC123
- KodOdb Dopuszczalna długość (20) Kod Odbiorcy (musi znajdować się w systemie w tabeli SRB) odpowiadający konkretnemu rachunkowi bankowemu np. DEF456
- NazwaKontr Dopuszczalna długość (100) Nazwa Kontrahenta, na rzecz, którego dokonywana jest transakcja np. PW
- NIPKontr Dopuszczalna długość (20) Numer NIP Kontrahenta, na rzecz, którego dokonywana jest transakcja np. PL555555555

Błędy i wyjątki

Lista potencjalnych błędów i wyjątków

- a) tmp_na_check
 - Plik nagłówkowy jest pusty
 - o Jest więcej niż jeden podmiot w pliku
 - o Jest więcej niż jeden miesiąc w pliku
 - o Podany Podmiot nie znajduje się w słowniku podmiotów (Tabela Podmiot)
 - o Podany miesiąc i rok jeszcze się nie zakończył
 - o Powtarzające się numery wyciągów
- b) tmp_poz_check
 - Błędy w nagłówkach
 - Nie wszystkie nagłówki maja swoje pozycje w pliku wb_poz.txt
 - Są nagłówki bez pozycji
 - o Kodu nadawcy lub odbiorcy nie ma w słowniku (Tabela SRB)

W razie wystąpienia błędu lub wyjątku w którejś z powyższych procedur, nie nastąpi zapisanie danych do tabel docelowych. Dane zostaną tylko w tabelach tymczasowych odpowiednio tmp_wb_na oraz tmp_wb_poz.

Analiza potrzebnych konwersji

Wszystkie dane w plikach wejściowych są w formie tekstu. Podczas działania pakietów, format danych zmienia się odpowiednio dla formatów kolumn w docelowych kolumnach

- Tekst -> daty
- Tekst -> Money
- Tekst -> Tekst

Analiza integralności i jakości dla finalnego raportu

Według broszury informacyjnej dla plików JPK_WB:

https://www.gov.pl/attachment/9f318b30-efd9-446a-8d5b-6ce204381bfd

Elementy składające się na JPK_WB:

Nagłówek – nagłówek JPK_WB

Nagłówek JPK_WB składa się z następujących pól: "KodFormularza", "WariantFormularza", "CelZlozenia", "DataWytworzeniaJPK", "DataOd", "DataDo", "DomyslnyKodWaluty", "KodUrzędu". Określają one właściwości JPK WB.

Podmiot1 – podmiot JPK WB

Struktura podmiotu JPK_WB składa się z pól: "IdentyfikatorPodmiotu", "AdresPodmiotu". Zawierają one dane identyfikujące podmiot, który składa JPK_WB.

Na dane identyfikujące podmiot składają się następujące pola: "NIP", "PelnaNazwa".

Do pól, które wchodzą w skład adresu podmiotu należą: "KodKraju", "Wojewodztwo", "Powiat", "Gmina", "Ulica", "NrDomu", "NrLokalu", "Miejscowosc", "KodPocztowy".

NumerRachunku – numer IBAN rachunku którego dotyczy wyciąg

Dane dotyczące numeru rachunku bankowego obejmują pole: "NumerRachunku".

Salda – salda początkowe i końcowe wyciągu

W skład danych elementu Salda wchodzą następujące pola: "SaldoPoczatkowe", "SaldoKoncowe".

WyciagWiersz – szczegółowe wiersze (zapisy) wyciągu bankowego

Dane dotyczące WyciagWiersz obejmują pola: "NumerWiersza", "DataOperacji", "NazwaPodmiotu", "OpisOperacji", "KwotaOperacji", "SaldoOperacji".

WyciagCtrl – sumy kontrolne dla tabeli WyciagWiersz

Sumy kontrolne dla tabeli WyciągWiersz składają się z pól: "LiczbaWierszy", "SumaObciazen", "SumaUznan".

Wszystkie powyższe elementy zawierają się w końcowym wynikowym pliku XML powstającym po wykonaniu procedury JPK_WB_3.

Opis procedur

Nazwa procedury oraz argumenty wejściowe	Opis działania
rmv_table @tab_name	Procedura, która usuwa tabele o zadanej nazwie. Procedura pomocnicza ułatwiająca testowanie i sprawdzanie poprawności działania skryptu.
create_empty_proc @proc_name	Procedura, która tworzy pusta procedurę o zadanej nazwie. Procedura pomocnicza ułatwiająca wielokrotne wykonywanie skryptu.
create_empty_fun @fun_name	Procedura, która tworzy pusta funkcje o zadanej nazwie. Procedura pomocnicza ułatwiająca tworzenie funkcji.
txt2M @txt	Funkcja, która dokonuje konwersji typu danej z tekstowego na money

txt2D @txt	Funkcja, która dokonuje konwersji typu danej z tekstowego na datetime
tmp_na_check	Procedura służy do walidacji nagłówków wyciągów bankowych zawartych w tabeli tymczasowej tmp_wb_na. Procedura wykonuje serie sprawdzeń walidacyjnych i rejestruje wszelkie znalezione błędy w tabeli błędów. Jeśli zostanie wykryty błąd, procedura zwraca kod błędu '-1', w przeciwnym razie zwraca kod sukcesu '0'.
tmp_poz_check	Procedura ta jest rozszerzeniem procedury tmp_na_check i służy do dodatkowej walidacji oraz tworzenia wyciągów bankowych i ich pozycji na podstawie danych zawartych w tabelach tymczasowych 'tmp_wb_na' oraz 'tmp_wb_poz'. Jeśli podczas tworzenia wyciągów i ich pozycji wystąpią błędy, procedura wycofa wszelkie zmiany (dokona 'ROLLBACK') i zwróci kod błędu '-1', w przeciwnym razie zakończy się sukcesem i zwróci kod sukcesu '0'.
DANE_ADR_KLI @miasto, @ulica	Procedura łączy nazwę miasta i nazwę ulicy w jedna ciągłą wartość typu 'NVARCHAR'
jpk_wb0 @rm	Procedura odpowiedzialna za generowanie danych w odpowiednich formatach dla okresu od 01.2019
jpk_wb1 @rm	Procedura odpowiedzialna za generowanie danych w odpowiednich formatach dla okresu od 01.2020
jpk_wb2 @rm	Procedura odpowiedzialna za generowanie danych w odpowiednich formatach dla okresu od 01.2021
jpk_wb3 @rm	Procedura odpowiedzialna za generowanie danych w odpowiednich formatach dla okresu od 01.2022
RUN_JPK @typ, @rm	Procedura szuka w tabeli 'PROC_JPK' informacji o najnowszym formacie raportu, który obowiązuje w podanym miesiącu dla danego typu. Dzięki temu mechanizmowi możemy uruchomić odpowiednia procedurę do generowania raportu zgodnie z formatem obowiązującym w danym miesiącu i dla danego typu raportu.
SAFT_RMV_PREFIX @msg	Funkcja usuwa kod kraju z numeru NIP
SAFT_INI_DATA @ym	Procedura tworzy dwie daty – pierwszy dzień podanego miesiąca oraz ostatni dzień podanego miesiąca
SAFT_CLEAR_TEXT @msg	Funkcja czyści pole tekstowe z wrażliwych znaków takich jak np. '>', '<', ':'
SAFT_DATE @d	Funkcja zamienia otrzymana datę na format dopuszczalny w plikach JPK
SAFT_GET_AMT @ant	Funkcja zamienia otrzymana kwotę na format dopuszczalny w plikach JPK
JPK_WB_3 @ym, @CC	Procedura odpowiedzialna za utworzenie pliku XML JPK_WB dla konkretnego miesiąca oraz konkretnego podmiotu

Testy poprawności

Testy poprawności polegają na sprawdzeniu czy system poprawnie zareaguje na błędy lub wyjątki zawarte w analizie walidacji danych -> Błędy i wyjątki

Walidacja nagłówków

Plik nagłówkowy jest pusty

- Sprawdz poprawnosc tmp_wb_na
 - Validation has started
 - Validation is completed
 - Validation has started
 - Validation is completed
 - Start 23:10:11
 - Validation has started
 - Validation is completed
 - (9) Error: Executing the query "EXEC tmp_na_check" failed with the following error: "Blad w procedurze: Wyciag Bankowy: WB_n_process / Plik nagłowkowy jest pusty.".
 - parameters not set correctly, or connection not established correctly.
 - Task Sprawdz poprawnosc tmp_wb_na failed
 - Finished, 23:10:12, Elapsed time: 00:00:00.109

Jest więcej niż jeden podmiot w pliku

- Sprawdz poprawność tmp wb na
 - A Validation has started
 - Validation is completed
 - Validation has started
 - Validation is completed
 - Start, 23:11:27
 - Validation has started
 Validation is completed

 - 🕦 Error: Executing the query "EXEC tmp_na_check" failed with the following error: "Blad w procedurze: Wyciag Bankowy: WB_n_process / Więcej jak jeden podmiot w pliku.".
 - (1) correctly, parameters not set correctly, or connection not established correctly
 - Task Sprawdz poprawnosc tmp_wb_na failed
 - Finished, 23:11:27, Elapsed time: 00:00:00.110

Jest więcej niż jeden miesiąc w pliku

- Sprawdz poprawnosc tmp_wb_na

 - ⇒ Start, 23:16:37

 ✓ Validation has started
 - Validation is completed
 - ⑤ Error: Executing the query "EXEC tmp_na_check" failed with the following error: "Blad w procedurze: Wyciag Bankowy: WB_n_process / Więcej jak jeden miesiąc w pliku.".
 - correctly, parameters not set correctly, or connection not established correctly.
 - Task Sprawdz poprawnosc tmp_wb_na failed
 - Finished, 23:16:37, Elapsed time: 00:00:00.141

Podany Podmiot nie znajduje się w słowniku podmiotów (Tabela Podmiot)

- Sprawdz poprawnosc tmp_wb_na

 - ∇ Validation has started
 Validation is completed

 - Start. 23:05:46
 - Validation has started
 - Validation is completed
 - 🕟 Error: Executing the query "EXEC tmp_na_check" failed with the following error: "Blad w procedurze: Wyciag Bankowy: WB_n_process / Podmiotu o tym kodzie nie ma w słowniku.". () correctly, parameters not set correctly, or connection not established correctly
- Task Sprawdz poprawnosc tmp_wb_na failed
- Finished, 23:05:46, Elapsed time: 00:00:00.125

Podany miesiąc i rok jeszcze się nie zakończył

- Sprawdz poprawnosc tmp_wb_na
 - Validation has started
 - Validation is completed
 - ∇alidation has started
 ∇alidation is completed
 - Start, 23:13:14

 - Validation has started
 Validation is completed
 - ⑤ Error: Executing the query "EXEC tmp_na_check" failed with the following error: "Blad w procedurze: Wyciag Bankowy: WB_n_process / Można raportować TYLKO miesiące które MINĘŁY". set correctly, parameters not set correctly, or connection not established correctly.
 - Task Sprawdz poprawnosc tmp_wb_na failed
 - Finished, 23:13:14, Elapsed time: 00:00:00.110

Powtarzające się numery wyciągów

- Sprawdz poprawnosc tmp_wb_na
- ∀ Validation has started
 ∀ Validation is completed
- Validation has started
 Validation is completed
- Start. 23:17:24
- A Validation has started
- Validation is completed
- 🗜 Error: Executing the query "EXEC tmp_na_check" failed with the following error: "Blad w procedurze: Wyciag Bankowy: WB_n_process / Mamy powtarzające się numery wyciagow
- correctly, parameters not set correctly, or connection not established correctly
- Task Sprawdz poprawnosc tmp_wb_na failed
- Finished, 23:17:25, Elapsed time: 00:00:00.110

Poprawny plik z nagłówkami do wyciągu bankowego

- Sprawdz poprawnosc tmp_wb_na
 - Validation has started
 - Validation is completed
 - Validation has started
 - Validation is completed
 - Start, 23:40:04
 - Validation has started
 - Validation is completed
 - Progress: Executing query "EXEC tmp_na_check". 100 percent complete
 - Finished, 23:40:05, Elapsed time: 00:00:00.063

Walidacja pozycji

Błędy w nagłówkach

- Sprawdz poprawnosc tmp_poz_check
 - Validation has started
 Validation is completed

 - A Validation has started
 - Validation is completed
 - Start, 23:18:21
 - Validation has started Validation is completed
 - 🗜 Error: Executing the query "EXEC tmp_poz_check" failed with the following error: "Blad w procedurze: Wyciag Bankowy: WB_n_process / Mamy powtarzające się numery wyciagow".
 - Correctly, parameters not set correctly, or connection not established correctly.
 - Task Sprawdz poprawnosc tmp_poz_check failed
 - Finished, 23:18:21, Elapsed time: 00:00:00.265

Są nagłówki bez pozycji

- Sprawdz poprawnosc tmp_poz_check
 - Validation has started
 - Validation is completed
 - Validation has started
 - Validation is completed
 - → Start, 23:28:18
 - Validation has started
 - Validation is completed
 - 🗜 Error: Executing the query "EXEC tmp_poz_check" failed with the following error: "Blad w procedurze: tmp_poz_check / Mamy nagłowki bez pozycji". 🛚
 - (i) correctly, or connection not established correctly.
 - Task Sprawdz poprawnosc tmp_poz_check failed
 - Finished, 23:28:18, Elapsed time: 00:00:00.094

Kodu nadawcy lub odbiorcy nie ma w słowniku (Tabela SRB)

- Sprawdz poprawnosc tmp_poz_check
 - Validation has started
 - Validation is completed
 - Validation has started Validation is completed
 - Start, 23:37:20
 - Validation has started
 - Validation is completed
 - (1) Error: Executing the query "EXEC tmp_poz_check" failed with the following error: "Blad w procedurze: tmp_poz_check / Mamy rachunek bankowy ktorego nie ma w slowniku". (i) correctly, parameters not set correctly, or connection not established correctly.
 - Task Sprawdz poprawnosc tmp_poz_check failed
 - Finished, 23:37:20, Elapsed time: 00:00:00.218

Poprawny plik z pozycjami do wyciągu bankowego

- Sprawdz poprawnosc tmp_poz_check
 - Validation has started
 - Validation is completed
 - Validation has started
 - Validation is completed
 - Start, 23:40:40
 - Validation has started
 - Validation is completed
 - Progress: Executing query "EXEC tmp_poz_check". 100 percent complete
 - Finished, 23:40:41, Elapsed time: 00:00:00.515

Wnioski

System zbudowany w oparciu o skrypty MS SQL i pakiety SSIS spełnia swoje zadanie, jakim jest transfer, przetworzenie i walidacja danych z plików tekstowych do formatu XML JPK_WB. System jest dobrze zaprojektowany i posiada szereg funkcji ułatwiających jego działanie, takich jak:

- Procedury walidacyjne: tmp_na_check oraz tmp_poz_check zapewniają kontrolę poprawności danych na etapie importu.
- Mechanizm wyboru procedury: RUN_JPK dobiera odpowiednią procedurę do generowania raportu JPK WB w zależności od miesiąca i typu raportu.
- Funkcje pomocnicze: txt2M, txt2D itp. ułatwiają konwersję typów danych.
- Słowniki: SRB, Podmiot i Adresy zapewniają spójność i poprawność danych.

Mocne strony:

- Skalowalność: System może obsługiwać dużą ilość danych.
- Automatyzacja: Proces importu, walidacji i generowania raportu jest zautomatyzowany.
- Dokumentacja: Dokumentacja zawiera szczegółowy opis działania systemu.

Słabe strony:

- Zależność od MS SQL: System wymaga dostępu do MS SQL Server.
- Brak interfejsu użytkownika: System nie posiada interfejsu graficznego, co utrudnia jego obsługę dla mniej doświadczonych użytkowników.

Dalsze możliwości rozwoju

Możliwości rozwoju:

- Interfejs użytkownika: Stworzenie interfejsu graficznego ułatwiającego obsługę systemu.
- Analiza danych: Dodanie funkcji analizy danych, np. generowanie raportów i statystyk.
- Integracja z innymi systemami: Integracja z systemami bankowymi i księgowymi.

• Obsługa innych formatów: Dodanie obsługi innych formatów plików wejściowych i wyjściowych.

Krytyczna analiza wniosków

Należy wziąć pod uwagę, że system jest prototypem i wymagałby dalszych prac rozwojowych. System nie został przetestowany na dużych zestawach danych liczących setki lub tysiące pozycji.