

SMARTMATCH

Temat:

Projekt systemu dopasowywania i ujednolicania danych adresowych dla instytucji miejskich

Autorzy:

Andrzej Brzeziński

Spis treści

1. Ogólne Wymagania i Opis Systemu	3
1.1 Opis Systemu	3
1.2 Wymagania Funkcjonalne	3
1.3 Wymagania Niefunkcjonalne	4
2. Cel Projektu	4
3. Analiza Użytkowników	5
3.1 Rodzaje Użytkowników	5
3.2 Cele i Funkcje dla Każdego Użytkownika	5
4. Aktorzy i Przypadki Użycia	7
4.1 Aktorzy	7
4.2 Przypadki Użycia	7
4.3 Wymagania	19
5. Procesy Biznesowe i Diagramy BPMN	20
5.1 Opis Procesów	21
5.1.1 Uruchomienie i sprawdzenie dostępności AI	21
5.1.2 Wybór języka i szablonu	21
5.1.3 Import danych	22
5.1.4 Ustawienia dopasowywania	22
5.1.5 Analiza i Wyniki	23
5.1.6 Zapis / Eksport wyników	24
5.2 Diagram BPMN	25
6. Diagramy Przypadków Użycia	27
7. User Stories	28
8. Diagram Bazy Danych (ERD)	30
9. Architektura i Wybór Technologii	30
9.1 Architektura Systemu	30
9.2 Wybór Technologii	31
10. Plan Implementacji – Wykres Gantta	33
11. Analiza Ryzyka	34
12. Analiza Konkurencji	37
13. Historia zmian	38

1. Ogólne Wymagania i Opis Systemu

1.1 Opis Systemu

Projekt nowoczesnego programu w postaci wtyczki do Microsoft Excel, przeznaczonego do **analizy, dopasowywania i ujednolicania** danych z różnych źródeł. Wtyczka eliminuje ręczne operacje i minimalizuje ryzyko błędów, oferując zaawansowane funkcje wspierane przez **sztuczną inteligencję**. Jej innowacyjność polega na dynamicznym uczeniu się reguł i automatycznym rozwiązywaniu problemów z dopasowywaniem danych. Projekt ten **wyróżnia się** na tle konkurencyjnych rozwiązań, dzięki wykorzystaniu modeli AI do analizy przypadków i dopasowywania danych, które wcześniej były trudne lub niemożliwe do rozwiązania manualnie. Użytkownicy będą mieli możliwość zapamiętania dostosowanego modelu do swoich potrzeb, aby mógł być stosowany w przyszłości.

Wtyczka jest rozwijana z myślą o możliwości rozbudowy w przyszłości o nowe funkcjonalności, w oparciu o feedback użytkowników, zgłaszających swoje potrzeby oraz o statystyki skuteczności działania programu.

W początkowej fazie dystrybucji, program będzie dla użytkowników darmowy, planowane wprowadzenie opłat związanych z korzystaniem modeli AI od ilości analiz użytkownika.

1.2 Wymagania Funkcjonalne

Program będzie realizował funkcje:

Import danych: Wczytywanie danych z plików Excel i CSV, Pobieranie danych z API, Obsługa wielu źródeł danych i ich podgląd w czasie rzeczywistym.

Dopasowywanie danych: Automatyczne dopasowywanie na podstawie modeli AI, Dopasowanie dokładne oraz przybliżone, Normalizacja danych, np. usuwanie spacji, konwersja wielkości liter.

Zarządzanie regułami dopasowywania: Tworzenie, edycja i usuwanie reguł, Automatyczne zapisywanie reguł do przyszłego zastosowania, Eksport i import reguł z wewnętrznej bazy danych.

Uczenie maszynowe: Zarządzanie modelami AI: wybór aktywnego modelu, trenowanie nowych modeli, Rozwiązywanie problematycznych przypadków na podstawie statystyk i feedbacku użytkownika, Alerty dla wyjątków, wymagających interwencji użytkownika.

Analiza wyników: Podgląd dopasowanych, brakujących i problematycznych danych, Dynamiczne generowanie statystyk, Eksport wyników do Excela, CSV lub JSON.

Personalizacja i konfiguracja: Wybór języka, Zapisywanie i wczytywanie szablonów ustawień.

1.3 Wymagania Niefunkcjonalne

Realizacja tego projektu wymaga zastosowania serwera chmurowego do obsługi modeli AI tj. Google Cloud AI Platform, który będzie używał takich technologii jak **Python** i **TensorFlow**. Potrzebna jest również System Zarządzania Bazą Danych zgodny z PostgreSQL oraz Serwer VPS do hostowania bazy. Do przygotowania backendu zostanie użyty **Note.js**, **SQLite** oraz **Office.js**, natomiast do frontentu zostanie użyty **HTML**, **CSS** i **JavaScript**.

Wymagania, które muszą zostać spełnione po stronie użytkownika to posiadanie programu Microsoft Excel w wersji Office 365 lub Excel 2016 i nowsze oraz posiadanie systemu operacyjnego Windows 10/11 (64-bitowy) lub macOS 10.15 Catalina i nowsze.

2. Cel Projektu

Cele projektu - zgodnie z metodologią S.M.A.R.T. (pol. Z.M.O.R.A. - Zdefiniowane, Mierzalne, Ograniczone w czasie, Realne, Akceptowalne)

Celem ogólnym jest zostanie **liderem na rynku narzędzi do analizy danych w Excelu**, oraz stanie się narzędziem pierwszego wyboru dla analityków danych, instytucji publicznych oraz firm prywatnych, poszukujących efektywnego sposobu na ujednolicanie i raportowanie danych

Jako cel bezpośredni przyjęte zostało zdobycie 200 aktywnych użytkowników w ciągu pierwszych trzech miesięcy od wydania wersji beta i osiągnięcie zwrotu kosztów związanych z wdrożeniem i utrzymaniem projektu w ciągu sześciu miesięcy od wprowadzenia opłat związanych z użytkowaniem modeli AI.

Planowane wdrożenie pierwszej wersji wtyczki to pół roku od rozpoczęcia projektu. Aktualizacje wtyczki przewidziane są regularnie **co dwa miesiące**, bazujące na zgłoszeniach użytkowników oraz analizie statystyk działania. Aby dotrzeć do docelowych odbiorców, zostaną wdrożone akcje marketingowe, w tym publikowanie przewodników i filmów instruktażowych. Zakłada się również, że projekt zostanie wykorzystywany przez liczne instytucje typu Urzędy Miasta, co pozwoli na zgromadzenie środków na aktywne akcje promocyjne.

Budżet obejmuje Koszty związane z zakupem Serwera chmurowego do obsługi modeli AI (200zł miesięcznie), Centralnej Bazy Danych na Serwerze VPS (39.99zł miesięcznie). Pracownicy zadeklarowali się wziąć udział w przedsięwzięciu za procent z zysków ze sprzedaży.

Czas realizacji został podzielony na 3 etapy:

- Pierwsza wersja beta: działająca jedynie lokalnie, pół roku od rozpoczęcia projektu
- Wersja stabilna: po zastosowaniu Serwerów chmurowych, trzech miesiącach testów i uwzględnieniu feedbacku użytkowników
- Pełna aplikacja z zaawansowanymi funkcjami po fazach alfa i beta: **rok po rozpoczęciu projektu**.

Szacowane wydatki na pierwszą i drugą fazę projektu(do momentu wprowadzenia opłat za korzystanie z modeli AI) to ~4000zł.

3. Analiza Użytkowników

3.1 Rodzaje Użytkowników

Instytucje publiczne i przedsiębiorstwa współpracujące z urzędami - Organizacje, które wymieniają dane między różnymi systemami, np. przedsiębiorstwa wodociągowe dopasowujące dane o przyłączach do adresów zameldowania przekazywanych przez urzędy miasta.

Analitycy danych - Osoby odpowiedzialne za porządkowanie, analizowanie i raportowanie danych w firmach, instytucjach publicznych czy organizacjach.

Specjaliści ds. raportowania - Użytkownicy, którzy regularnie generują raporty z niespójnych lub niekompletnych danych, np. w arkuszach Excel.

Firmy prywatne - Przedsiębiorstwa z dużymi bazami klientów, produktów lub usług, które muszą synchronizować dane między różnymi systemami (np. CRM, ERP).

Administratorzy baz danych - Specjaliści techniczni, którzy odpowiadają za czystość danych w bazach.

Indywidualni użytkownicy - Osoby zajmujące się analizą danych na potrzeby prywatne lub małych projektów (np. łączenie list mailingowych, porównywanie danych z różnych arkuszy).

3.2 Cele i Funkcje dla Każdego Użytkownika

Rodzaj Użytkowników	Cele	Funkcje
Instytucje publiczne i przedsiębiorstwa współpracujące z urzędami	<p>Automatyczne dopasowywanie danych przekazywanych między różnymi instytucjami i ich raportowanie.</p> <p>Zmniejszenie liczby błędów wynikających z różnic w formatach danych</p>	<ul style="list-style-type: none"> Wykorzystanie AI do analizy i dopasowywania problematycznych danych (np. adresów). Tworzenie reguł dla konkretnych przypadków (np. różne formaty adresów). Zliczanie wartości w kolumnach w przypadku niestandardowych dopasowań wielu rekordów. Możliwość ręcznego zatwierdzania dopasowań w wyjątkowych przypadkach. Możliwość zapisu i wczytywania szablonów dopasowywania. Możliwość definiowania reguł dopasowywania.

		<ul style="list-style-type: none"> • Alerty dla problematycznych przypadków z możliwością ręcznego rozwiązania.
Analitycy danych	<p>Automatyzacja analizy i dopasowywania danych. Zmniejszenie czasu pracy nad ręcznym czyszczeniem i ujednolicaniem danych.</p>	<ul style="list-style-type: none"> • Możliwość definiowania reguł dopasowywania. • Automatyczne dopasowywanie danych wspierane przez modele AI. • Dynamiczne generowanie raportów i wizualizacji.
Specjaliści ds. raportowania	<p>Ujednolicanie danych w celu przygotowania czytelnych raportów. Wykrywanie problematycznych przypadków w danych przed ich raportowaniem.</p>	<ul style="list-style-type: none"> • Eksport wyników dopasowywania do plików Excel, CSV lub JSON. • Zliczanie i grupowanie danych w określonych kolumnach. • Wyróżnianie brakujących lub niedopasowanych rekordów.
Firmy prywatne	<p>Synchronizacja danych między różnymi systemami (np. CRM, ERP). Zapewnienie spójności danych klientów, produktów lub usług.</p>	<ul style="list-style-type: none"> • Automatyczna normalizacja danych, np. usuwanie znaków specjalnych. • Dopasowywanie danych z wielu źródeł w czasie rzeczywistym. • Możliwość definiowania i zapisywania reguł dla powtarzalnych przypadków.
Administratorzy baz danych	<p>Wspomaganie procesu czyszczenia danych. Wykrywanie i naprawa niespójności w bazach danych.</p>	<ul style="list-style-type: none"> • Wizualizacja różnic między danymi w tabelach. • Możliwość ręcznego zatwierdzania dopasowań w wyjątkowych przypadkach. • Statystyki skuteczności dopasowań i liczby wykrytych problemów.
Indywidualni użytkownicy	<p>Szybkie i łatwe porównywanie danych w arkuszach Excel. Wykrywanie niespójności w danych, np. duplikatów lub błędów literowych.</p>	<ul style="list-style-type: none"> • Możliwość zapisu i wczytywania szablonów dopasowywania. • Alerty dla problematycznych przypadków z możliwością ręcznego rozwiązania. • Statystyki podjętych działań, pokazujące różnice i błędy w zbiorach danych

4. Aktorzy i Przypadki Użycia

4.1 Aktorzy

Nazwa	Opis	Rodzaj	Interakcje z Systemem
Użytkownik końcowy	Osoba korzystająca z wtyczki w celu analizy, dopasowywania i ujednolicania danych. Może zarządzać regułami, wczytywać dane i analizować wyniki.	Osoba	<ul style="list-style-type: none"> • Importuje dane z plików lub API. • Wybiera ustawienia dopasowywania i konfiguracje modelu AI. • Przegląda, modyfikuje i zapisuje wyniki analizy. • Tworzy i zarządza regułami dopasowywania. • Eksportuje wyniki w formacie pliku (Excel, CSV, JSON).
Model AI	Komponent systemu odpowiedzialny za analizę danych, dopasowywanie rekordów i dynamiczne uczenie się reguł.	System	<ul style="list-style-type: none"> • Analizuje dane dostarczone przez użytkownika. • Wykrywa wyjątki i sugeruje reguły dopasowywania. • Uczy się na podstawie wprowadzonych przez użytkownika przypadków i zapisanych reguł. • Generuje statystyki skuteczności dopasowywania.
Zewnętrzne źródła danych	Systemy zewnętrzne dostarczające dane do analizy, np. pliki CSV, arkusze Excel lub API.	System	<ul style="list-style-type: none"> • Dostarcza dane wtyczce SmartMatch do analizy i dopasowywania. • Może odbierać wyniki w formacie eksportu, np. do dalszego przetwarzania.

4.2 Przypadki Użycia

UC001: Zmiana języka interfejsu

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik otwiera ekran startowy.
2. Użytkownik klika listę rozwijaną z aktualnie ustawionym językiem interfejsu, znajdującą się w dolnej części.

3. Użytkownik widzi dostępne opcje językowe.
4. Użytkownik wybiera jeden z języków.
5. System automatycznie zmienia język interfejsu na wybrany.

Rozszerzenia:

3A. Brak dodatkowych języków w liście:

3A 1. System wyświetla komunikat informujący o tylko jednym dostępnym języku.

3A 2. Koniec przypadku użycia.

5A. Wybrany język jest już aktywny:

5A 1. System wyświetla komunikat informujący, że wybrany język jest już ustawiony.

5A 2. Koniec przypadku użycia.

UC002: Import danych z arkusza Excel

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik otwiera ekran startowy i klika przycisk „Importuj dane”.
2. System przenosi użytkownika do ekranu importu danych.
3. W sekcji wyboru źródła danych, użytkownik klika przycisk „Wybierz arkusz Excela”.
4. System identyfikuje wszystkie otwarte arkusze w programie Excel i wyświetla komunikat: „Zaznacz zakres komórek w aktywnym arkuszu Excela, który chcesz zaimportować”.
5. Użytkownik przechodzi do otwartego arkusza Excela i za pomocą myszki zaznacza odpowiedni zakres komórek, które mają zostać zaimportowane.
6. Po zaznaczeniu danych użytkownik wraca do wtyczki i klika przycisk zatwierdzający zaznaczenie.
7. System wyświetla podgląd zaimportowanych danych na ekranie importu oraz umożliwia włączenie opcji przetwarzania.
8. Użytkownik zaznacza odpowiednie opcje i klika przycisk „Dalej”.

Rozszerzenia:

4A. Brak otwartych arkuszy w Excelu:

4A 1. System wyświetla komunikat o błędzie.

4A 2. Koniec przypadku użycia.

5A. Użytkownik zaznaczył pusty zakres komórek:

5A 1. System wyświetla komunikat o błędzie.

5A 2. Powrót do kroku 4.

8A. Użytkownik chce wrócić do menu głównego:

8A 1. Użytkownik klika „Wstecz”.

8A 2. System przenosi użytkownika do ekranu startowego.

8A 3. Koniec przypadku użycia.

UC003: Import danych z pliku CSV/Excel

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik otwiera ekran startowy i klika przycisk „Importuj dane”.
2. System przenosi użytkownika do ekranu importu danych.

3. W sekcji wyboru źródła danych, użytkownik klika przycisk „Wgraj plik”.
4. System otwiera okno eksploratora plików.
5. Użytkownik wybiera plik z danymi w formacie CSV lub Excel i klika przycisk „Otwórz”.
6. System wyświetla komunikat „Trwa wczytywanie pliku” oraz ładuje dane.
7. Po wczytaniu danych system wyświetla podgląd wierszy danych oraz opcje przetwarzania.
8. Użytkownik zaznacza wybrane opcje i klika przycisk „Dalej”.

Rozszerzenia:

- 4A. Wybrano plik w nieobsługiwanym formacie:
 - 4A 1. System wyświetla komunikat o błędzie.
 - 4A 2. Powrót do kroku 3.
- 5A. Plik nie zawiera danych:
 - 5A 1. System wyświetla komunikat o błędzie.
 - 5A 2. Powrót do kroku 3.

UC004: Pobieranie danych z API

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik klika przycisk „Importuj dane” na ekranie startowym.
2. System przenosi użytkownika do ekranu importu danych.
3. W sekcji wyboru źródła danych, użytkownik klika przycisk „Pobierz dane z API”.
4. System wyświetla formularz z polami niezbędnymi do uzyskania danych.
5. Użytkownik wprowadza dane w formularzu i klika przycisk „Pobierz”.
6. System nawiązuje połączenie z API i pobiera dane.
7. Po zakończeniu procesu system wyświetla podgląd danych oraz opcje przetwarzania.
8. Użytkownik zaznacza odpowiednie opcje i klika przycisk „Dalej”.

Rozszerzenia:

- 4A. Brak tokenu autoryzacyjnego:
 - 4A 1. System wyświetla komunikat: „Token autoryzacyjny jest wymagany do połączenia z API”.
 - 4A 2. Powrót do kroku 4.
- 6A. Błąd połączenia z API:
 - 6A 1. System wyświetla komunikat z błędem.
 - 6A 2. Powrót do kroku 4.

UC005: Wczytywanie szablonów

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik klika przycisk „Wczytaj szablon” na ekranie startowym.
2. System wyświetla ekran z listą zapisanych szablonów.
3. Użytkownik przegląda listę, wybiera szablon z listy, klika kursorem na odpowiedni i zatwierdza przyciskiem „Załaduj”.
4. System wczytuje konfigurację szablonu, wyświetlając odpowiedni komunikat.
5. System przenosi użytkownika do ekranu ustawień dopasowywania.

Rozszerzenia:

3A. Brak zapisanych szablonów:

- 3A 1. System wyświetla komunikat informujący o braku zapisanych szablonów.
- 3A 2. Koniec przypadku użycia.

UC006: Dodanie nowej reguły

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik klika przycisk „Zarządzaj regułami” w ekranie ustawień dopasowywania.
2. System wyświetla listę reguł z opcją „Dodaj regułę”.
3. Użytkownik klika przycisk „Dodaj regułę”.
4. System wyświetla formularz tworzenia nowej reguły.
5. Użytkownik wypełnia pola formularza i zatwierdza przyciskiem „Zapisz”.
6. System dodaje nową regułę do listy.

Rozszerzenia:

4A. Użytkownik nie wprowadził wszystkich wymaganych danych:

- 4A 1. System wyświetla komunikat: „Proszę wypełnić wszystkie pola”.
- 4A 2. Powrót do kroku 4.

UC007: Edycja istniejącej reguły

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik klika przycisk „Zarządzaj regułami” w ekranie ustawień dopasowywania.
2. System wyświetla listę reguł z opcją „Edytuj” obok każdej z nich.
3. Użytkownik klika przycisk „Edytuj” przy wybranej regule.
4. System otwiera formularz z aktualnymi danymi reguły.
5. Użytkownik zmienia dane i zatwierdza przyciskiem „Zapisz”.
6. System zapisuje zmiany i aktualizuje listę reguł.

Rozszerzenia:

5A. Użytkownik rezygnuje z edycji:

- 5A 1. Użytkownik klika przycisk „Anuluj”.
- 5A 2. System zamyka formularz i wraca do listy reguł.
- 5A 3. Koniec przypadku użycia.

UC008: Usunięcie reguły**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik klika przycisk „Zarządzaj regułami” w ekranie ustawień dopasowywania.
2. System wyświetla listę reguł z opcją „Usuń” obok każdej z nich.
3. Użytkownik klika przycisk „Usuń” przy wybranej regule.
4. System wyświetla komunikat z pytaniem o usunięcie reguły.
5. Użytkownik potwierdza, klikając przycisk „Tak”.
6. System usuwa regułę z listy i wyświetla komunikat: „Reguła została usunięta”.

Rozszerzenia:

5A. Użytkownik anuluje usuwanie:

5A 1. Użytkownik klika „Nie”.

5A 2. System zamyka komunikat i wraca do listy reguł.

5A 3. Koniec przypadku użycia.

UC009: Wybór metody dopasowania - Exact Match**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. System wyświetla dostępne metody dopasowywania.
3. Użytkownik klika przycisk „Exact Match”.
4. System zapisuje wybór i przechodzi do dalszych ustawień dopasowywania bez dodatkowych zmian.

UC010: Wybór metody dopasowania - Fuzzy Match**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. System wyświetla dostępne metody dopasowywania.
3. Użytkownik klika przycisk „Fuzzy Match”.
4. System zapisuje wybór i przechodzi do dalszych ustawień dopasowywania bez dodatkowych zmian.

UC011: Wybór metody dopasowania - AI Match z istniejącym modelem**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. System wyświetla dostępne metody dopasowywania.
3. Użytkownik klika przycisk „Dopasowanie AI”.
4. System wyświetla dodatkowy przycisk „Ustawienia modelu AI”.
5. Użytkownik klika przycisk „Ustawienia modelu AI”.

6. System otwiera ekran szczegółów modeli, wyświetlając listę istniejących modeli z informacjami o modelu.
7. Użytkownik wybiera jeden z modeli i klika przycisk „Zastosuj model”.
8. System zapisuje wybór modelu i wraca do ekranu ustawień dopasowywania.

Rozszerzenia:

6A. Brak istniejących modeli:

6A 1. System wyświetla komunikat: „Nie znaleziono istniejących modeli. Utwórz nowy model”.

6A 2. System przenosi użytkownika do ekranu tworzenia nowego modelu (UC012).

6A 3. Koniec przypadku użycia.

UC012: Wybór metody dopasowania - AI Match z tworzeniem nowego modelu

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. System wyświetla dostępne metody dopasowywania.
3. Użytkownik klika przycisk „Dopasowanie AI”.
4. System wyświetla dodatkowy przycisk „Ustawienia modelu AI”.
5. Użytkownik klika przycisk „Ustawienia modelu AI”.
6. System otwiera ekran szczegółów modeli i wyświetla opcję „Utwórz nowy model”.
7. Użytkownik klika przycisk „Utwórz nowy model”.
8. System wyświetla formularz tworzenia nowego modelu.
9. Użytkownik wypełnia formularz i potwierdza przyciskiem „Zatwierdź”.
10. System inicjalizuje nowy model i wyświetla komunikat o dodaniu nowego modelu.
11. Użytkownik wybiera nowo utworzony model i klika przycisk „Zastosuj model”.
12. System zapisuje wybór modelu i wraca do ekranu ustawień dopasowywania.

Rozszerzenia:

8A. Użytkownik anuluje tworzenie nowego modelu:

8A 1. Użytkownik klika „Anuluj”.

8A 2. System wraca do ekranu szczegółów modeli.

8A 3. Koniec przypadku użycia.

UC013: Trenowanie modelu na nowych danych

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik klika przycisk „Ustawienia modelu AI” w ekranie ustawień dopasowywania.
2. System otwiera ekran szczegółów modeli, wyświetlając listę dostępnych modeli.
3. Użytkownik wybiera model, który chce trenować, i klika przycisk „Trenuj model”.

4. System przenosi użytkownika do ekranu importu danych, gdzie użytkownik wczytuje dane treningowe zgodnie z jednym z wcześniej zdefiniowanych przypadków:
 - Import danych z arkusza Excel (UC002).
 - Import danych z pliku CSV/Excel (UC003).
 - Pobieranie danych z API (UC004).
5. System wyświetla komunikat: „Trwa trenowanie modelu” wraz z paskiem postępu.
6. Po zakończeniu procesu system wyświetla komunikat o pomyślnym przetrenowaniu modelu.
7. System aktualizuje statystyki modelu i przenosi użytkownika z powrotem do ekranu szczegółów modelu.

Rozszerzenia:

- 4A. Brak danych do trenowania:
 - 4A 1. System wyświetla komunikat z błędem.
 - 4A 2. Powrót do kroku 4.
- 5A. Proces trenowania nie powiódł się:
 - 5A 1. System wyświetla komunikat z błędem.
 - 5A 2. Powrót do kroku 4.

UC016: Dynamiczny podgląd wyników

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. Po dokonaniu wyboru kolumn kluczowych i metody dopasowywania, system automatycznie generuje podgląd wyników w czasie rzeczywistym.
3. W sekcji dynamicznego podglądu użytkownik widzi prognozowany podgląd wyników dopasowania danych przy aktualnych ustawieniach.
4. Użytkownik przegląda wyniki i może wprowadzać zmiany w ustawieniach dopasowywania.
5. System automatycznie odświeża podgląd wyników po każdej zmianie.

Rozszerzenia:

- 3A. Brak wyników dopasowania:
 - 3A 1. System wyświetla komunikat o błędzie.
 - 3A 2. Użytkownik wraca do ustawień dopasowywania (UC009-UC011).
 - 3A 3. Koniec przypadku użycia.
- 4A. Użytkownik chce wyłączyć podgląd dynamiczny:
 - 4A 1. Użytkownik klika przycisk „Wyłącz podgląd”.
 - 4A 2. System ukrywa sekcję podglądu i wyświetla komunikat.
 - 4A 3. Koniec przypadku użycia.

UC017: Rozwiązywanie wyjątków w danych**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Po przeprowadzeniu analizy danych, system wyświetla sekcję „Napotkane problemy” w ekranie analizy wyników.
2. Użytkownik klika „Pokaż listę problemów”.
3. System wyświetla listę wyjątków, w tym:
 - Opis problemu (np. „Brak dopasowania dla rekordu”).
 - Proponowaną sugestię rozwiązania.
 - Opcje działań dla użytkownika:
 - „Zaakceptuj sugestię”.
 - „Dodaj regułę”.
 - „Ignoruj przypadek”.
 - „Zastosuj dla podobnych przypadków”.
4. Użytkownik wybiera jedną z dostępnych opcji i zatwierdza.
5. System zapisuje decyzję i aktualizuje wyniki analizy.

Rozszerzenia:

- 3A. Użytkownik rezygnuje z rozwiązywania problemów:
 - 3A 1. Użytkownik klika „Zamknij listę problemów”.
 - 3A 2. System wraca do ekranu analizy wyników.
 - 3A 3. Koniec przypadku użycia.

UC018: Zapis wyników do obecnego arkusza**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik przechodzi do ekranu analizy wyników.
2. Użytkownik klika przycisk „Zapisz do arkusza”.
3. System wyświetla komunikat: „Zaznacz komórkę, w której chcesz wstawić przetworzone dane”.
4. Użytkownik przechodzi do otwartego arkusza w Excelu i za pomocą myszki zaznacza komórkę początkową, gdzie mają zostać wstawione dane.
5. Użytkownik wraca do wtyczki i zatwierdza przyciskiem „Zatwierdź”.
6. System wstawia przetworzone dane do zaznaczonego miejsca w arkuszu i wyświetla komunikat.

Rozszerzenia:

- 4A. Brak aktywnego arkusza w Excelu:
 - 4A 1. System wyświetla komunikat o błędzie.
 - 4A 2. Koniec przypadku użycia.
- 5A. Użytkownik nie zaznaczył komórki:

- 5A 1. System wyświetla komunikat o błędzie.
- 5A 2. Powrót do kroku 3.

UC019: Eksport wyników do pliku CSV/JSON

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik przechodzi do ekranu analizy wyników.
2. Użytkownik klika przycisk „Eksportuj do pliku”.
3. System wyświetla opcje formatu eksportu.
4. Użytkownik wybiera jeden z dostępnych formatów (CSV lub JSON).
5. System otwiera okno dialogowe, umożliwiające użytkownikowi wybór lokalizacji i nazwy pliku.
6. Użytkownik zatwierdza lokalizację i potwierdza przyciskiem „Zapisz”.
7. System generuje plik w wybranym formacie i wyświetla komunikat.

Rozszerzenia:

- 5A. Użytkownik anuluje eksport:
 - 5A 1. Użytkownik klika przycisk „Anuluj”.
 - 5A 2. System zamyka okno dialogowe i wraca do ekranu analizy wyników.
 - 5A 3. Koniec przypadku użycia.
- 6A. Błąd podczas zapisu pliku:
 - 6A 1. System wyświetla komunikat o błędzie.
 - 6A 2. Powrót do kroku 5.

UC020: Zapisywanie szablonu dopasowywania

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. Po skonfigurowaniu opcji ,użytkownik klika przycisk „Zapisz szablon”.
3. System wyświetla okno dialogowe z nazwą i opisem szablonu.
4. Użytkownik wypełnia pola i potwierdza przyciskiem „Zatwierdź”.
5. System zapisuje szablon w bazie wtyczki i wyświetla komunikat.

Rozszerzenia:

- 3A. Użytkownik nie wprowadził nazwy szablonu:
 - 3A 1. System wyświetla komunikat o błędzie.
 - 3A 2. Powrót do kroku 3.
- 5A. Problem z zapisem szablonu:
 - 5A 1. System wyświetla komunikat o błędzie.
 - 5A 2. Powrót do kroku 3.

UC021: Otworzenie poradnika**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik znajduje się na dowolnym ekranie wtyczki.
2. Użytkownik klika przycisk „Poradnik” w górnym rogu interfejsu.
3. System wyświetla samouczek w formie wyskakujących odpowiedzi.
4. Użytkownik klika przyciski „Dalej” lub „Zakończ”, aby odpowiednio przejść przez kolejne kroki samouczka lub go zakończyć.
5. System zamyka poradnik po ukończeniu wszystkich kroków lub na żądanie użytkownika.

Rozszerzenia:

- 3A. Poradnik nie jest dostępny dla bieżącego ekranu:
- 3A 1. System wyświetla komunikat: „Brak poradnika dla tej sekcji”.
 - 3A 2. Koniec przypadku użycia.

UC022: Wybór typu wyników**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. W sekcji „Typ wyników” użytkownik wybiera jedną z dostępnych opcji:
 - „Tylko dopasowane”
 - „Brakujące dane”
 - „Wszystkie dane”.
3. Użytkownik zaznacza wybraną opcję, klikając odpowiedni przycisk radiowy.
4. System zapisuje wybór i wyświetla komunikat.

Rozszerzenia:

- 3A. Użytkownik zmienia zdanie i wybiera inną opcję:
- 3A 1. System aktualizuje wybór bez konieczności dodatkowego potwierdzenia.
 - 3A 2. Koniec przypadku użycia.

UC023: Wybór wartości do zliczenia**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Użytkownik przechodzi do ekranu ustawień dopasowywania.
2. W sekcji „Wartości do zliczenia” użytkownik zaznacza kolumny, które mają zostać zsumowane w wynikach.
3. System wyświetla listę kolumn dostępnych w zaimportowanych danych.
4. Użytkownik zaznacza jedną lub więcej kolumn, klikając odpowiednie pola wyboru.
5. Użytkownik klika przycisk „Zastosuj”.

6. System zapisuje wybór i uwzględnia go w procesie analizy wyników.

Rozszerzenia:

- 4A. Użytkownik nie zaznaczył żadnych kolumn:
 - 4A 1. System wyświetla komunikat o błędzie.
 - 4A 2. Powrót do kroku 3.

UC024: Przejście do analizy wyników po konfiguracji dopasowywania

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Użytkownik kończy konfigurację ustawień dopasowywania.
2. Użytkownik klika przycisk „Analizuj i dopasuj” w ekranie ustawień dopasowywania.
3. System wyświetla komunikat: „Trwa analiza danych i generowanie wyników”.
4. System przetwarza dane zgodnie z wybranymi ustawieniami.
5. Po zakończeniu analizy system przenosi użytkownika do ekranu analizy wyników.

Rozszerzenia:

- 3A. Proces analizy nie powiódł się:
 - 3A 1. System wyświetla komunikat o błędzie.
 - 3A 2. Powrót do ekranu ustawień dopasowywania.

UC025: Podświetlanie różnic w wynikach

Aktorzy: Użytkownik końcowy

Scenariusz główny:

1. Po zakończeniu analizy danych użytkownik zostaje przeniesiony do ekranu analizy wyników.
2. W systemie wyświetlana jest tabela wynikowa.
3. Użytkownik może włączyć/wyłączyć podświetlanie różnic, klikając opcję „Podświetl różnice”.

Rozszerzenia:

- 3A. Użytkownik wyłącza podświetlanie różnic:
 - 3A 1. System usuwa oznaczenia kolorystyczne w tabeli.
 - 3A 2. Koniec przypadku użycia.
- 3B. Użytkownik włącza podświetlanie różnic:
 - 3B 1. System ponownie oznacza różnice w tabeli wyników.
 - 3B 2. Koniec przypadku użycia.

UC026: Przeglądanie raportu zmian**Aktorzy:** Użytkownik końcowy**Scenariusz główny:**

1. Po zakończeniu analizy danych użytkownik zostaje przeniesiony do ekranu analizy wyników.
2. Użytkownik klika przycisk „Pokaż raport zmian”.
3. System wyświetla raport w nowym oknie, zawierający szczegółowe informacje.
4. Użytkownik przegląda raport i klika przycisk „Zamknij”, aby wrócić do ekranu analizy wyników.

Rozszerzenia:

3A. Raport jest pusty:

- 3A 1. System wyświetla komunikat o błędzie.
- 3A 2. Powrót do ekranu analizy wyników.
- 3A 3. Koniec przypadku użycia.

4A. Użytkownik eksportuje raport zmian:

- 4A 1. Użytkownik klika przycisk „Eksportuj raport”.
- 4A 2. System zapisuje raport w formacie Excel, CSV lub JSON zgodnie z wyborem użytkownika.
- 4A 3. Koniec przypadku użycia.

UC027: Automatyczne dostosowanie modelu AI**Aktorzy:** Model AI**Scenariusz główny:**

1. Użytkownik końcowy rozpoczyna proces dopasowywania danych, wybierając metodę „Dopasowywanie AI”.
2. System przesyła dane wejściowe do Modelu AI.
3. Model AI analizuje dane i zwraca wyniki z dopasowaniami oraz ewentualnymi wyjątkami.
4. System zapisuje wyniki w pamięci lokalnej wtyczki.
5. Model AI na podstawie zwróconych wyjątków oraz działań użytkownika dostosowuje swoje parametry, aby zwiększyć skuteczność w przyszłych iteracjach.
6. System zapisuje zaktualizowaną wersję modelu na serwerze w chmurze.

Rozszerzenia:

3A. Model AI nie zwraca wyników w określonym czasie:

- 3A 1. System wyświetla komunikat o błędzie.
- 3A 2. Koniec przypadku użycia.

UC028: Automatyczna aktualizacja danych z zewnętrznego źródła**Aktorzy:** Zewnętrzne źródła danych**Scenariusz główny:**

1. System zapisuje harmonogram pobierania danych (np. co 24 godziny).
2. W określonym czasie system automatycznie łączy się z API zewnętrznego źródła danych.
3. API zwraca zaktualizowane dane.

4. System przeprowadza wstępną analizę i zapisuje dane w lokalnej pamięci wtyczki.
5. Użytkownik jest powiadamiany o zaktualizowanych danych podczas następnego uruchomienia wtyczki.

Rozszerzenia:

2A. API nie odpowiada w wyznaczonym czasie:

2A 1. System loguje błąd w pamięci lokalnej i wyświetla komunikat przy kolejnym uruchomieniu o błędzie”.

2A 2. Koniec przypadku użycia.

3A. Dane z API zawierają błędy:

3A 1. System wyświetla komunikat o błędzie.

3A 2. Koniec przypadku użycia.

4.3 Wymagania

ID	Nazwa wymagania	Opis	UC
W01	Zmiana języka interfejsu	Umożliwia użytkownikowi zmianę języka interfejsu na ekranie startowym.	UC001
W02	Import danych z arkusza Excel	Pozwala użytkownikowi importować dane z wybranego zakresu komórek w aktywnym arkuszu Excela.	UC002
W03	Import danych z pliku CSV/Excel	Pozwala użytkownikowi importować dane z pliku w formacie CSV lub Excel.	UC003
W04	Pobieranie danych z API	Umożliwia użytkownikowi pobieranie danych z zewnętrznego API, korzystając z odpowiedniego formularza.	UC004
W05	Wczytywanie szablonów	Pozwala użytkownikowi wczytać zapisany szablon dopasowywania.	UC005
W06	Modyfikacje reguł	Umożliwia użytkownikowi dodanie nowej, edycję oraz usunięcie reguły dopasowywania danych.	UC006, UC007, UC008
W07	Wybór metody dopasowania	Umożliwia użytkownikowi wybranie metody dokładnego, przybliżonego lub opartego na modelu AI dopasowania danych	UC009, UC010, UC011, UC012
W08	Trenowanie modelu AI	Pozwala użytkownikowi przeprowadzić trenowanie modelu na nowych danych.	UC013
W09	Dynamiczny podgląd wyników	Umożliwia użytkownikowi podgląd wyników dopasowywania w czasie rzeczywistym.	UC016
W10	Rozwiązywanie wyjątków w danych	Pozwala użytkownikowi przeglądać i rozwiązywać problemy napotkane podczas analizy danych.	UC017
W11	Eksport wyników	Umożliwia użytkownikowi zapisanie przetworzonych wyników do wskazanego arkusza Excel lub do pliku CSV/JSON	UC018, UC019

W12	Zapisywanie szablonu dopasowywania	Umożliwia użytkownikowi zapisanie konfiguracji dopasowywania jako nowy szablon.	UC020
W13	Otworzenie poradnika	Umożliwia użytkownikowi otwarcie samouczka w dowolnym momencie pracy z wtyczką.	UC021
W14	Wybór typu wyników	Pozwala użytkownikowi wybrać, które wyniki dopasowywania mają być uwzględnione.	UC022
W15	Wybór wartości do zliczenia	Umożliwia użytkownikowi wskazanie kolumn, które mają być sumowane w wynikach analizy.	UC023
W16	Przejęcie do analizy wyników	Pozwala użytkownikowi rozpocząć analizę danych po zakończeniu konfiguracji ustawień.	UC024
W17	Podświetlanie różnic w wynikach	Umożliwia użytkownikowi włączenie/wyłączenie podświetlania różnic w tabeli wyników.	UC025
W18	Przeglądanie raportu zmian	Pozwala użytkownikowi przeglądać raport zawierający zmiany dokonane podczas analizy danych.	UC026
W19	Automatyczne dostosowanie modelu AI	Umożliwia Modelowi AI automatyczne dostosowanie się na podstawie wyjątków i działań użytkownika.	UC027
W20	Automatyczna aktualizacja danych	Umożliwia systemowi automatyczne aktualizowanie danych z zewnętrznego źródła w określonych odstępach czasu.	UC028

5. Procesy Biznesowe i Diagramy BPMN

Główny Przeptyw programu ma umożliwić użytkownikowi import danych, konfigurację metod dopasowania, zarządzanie regułami oraz przeprowadzenie analizy danych z wyświetleniem wyników i ewentualnym zapisem lub eksportem.

Komponenty:

Użytkownik – Osoba pracująca w Excelu i obsługująca wtyczkę.

Wtyczka SmartMatch – Moduł działający w ramach Excela, odpowiadający za logikę dopasowywania, ekrany konfiguracyjne i komunikaty.

Excel – Arkusz kalkulacyjny, w którym pracuje użytkownik

Chmura/AI – Usługa w chmurze, udostępniająca modele AI do dopasowywania danych oraz trenowania.

5.1 Opis Procesów

5.1.1 Uruchomienie i sprawdzenie dostępności AI

Start (Rozpoczęcie procesu po uruchomieniu wtyczki „SmartMatch” w Excelu)



Sprawdź dostęp do chmury i modeli AI (Wtyczka w tle sprawdza połączenie z serwerem AI w chmurze, aby ustalić, czy można skorzystać z funkcji AI)



Czy chmura dostępna? (Decyzja logiczna: jeśli „Tak” – wtyczka wie, że AI może być używane; jeśli „Nie” – wtyczka nadal działa, ale w trybie ograniczonym (bez AI). W obu przypadkach proces przechodzi dalej do ekranu startowego)



Wyświetl Ekran Startowy (Użytkownik widzi podstawowe menu)

5.1.2 Wybór języka i szablonu

Czy zmienić język? (Użytkownik decyduje, czy chce skorzystać z opcji wielojęzycznego interfejsu (UC001). „Tak” prowadzi do (W2); „Nie” – od razu do wczytania szablonu)



Przełącz interfejs (Wtyczka aktualizuje wszystkie etykiety i komunikaty na wybrany język)



Czy wczytać szablon? (Jeśli w systemie istnieją gotowe konfiguracje (zapisane reguły, modele AI), można je wczytać (UC005). „Tak” → (W3); „Nie” → pytanie o import danych)



Wyświetl listę szablonów (System prezentuje zapisane szablony dopasowywania)



Ustaw reguły i modele z wybranego szablonu (Po wyborze szablonu wtyczka ładuje przypisane reguły, ustawienia i model AI)

5.1.3. Import danych

Czy importować dane? (Decyzja, czy użytkownik rozpocznie wczytywanie danych. „Tak” → ekran importu. „Nie” → powrót do Startu)



Wyświetl Ekran importu (Użytkownik widzi opcje wczytywania)



Wybór importu: Arkusz / Plik / API : Zaznacz zakres komórek (Arkusz) (UC002) ,Import z pliku (UC003) , Pobierz dane z API (UC004).



Powrót do Wyświetl Ekran importu (Po zakończeniu wybranej metody importu wtyczka pokazuje podgląd danych)



Kontynuować czy Anuluj? („Kontynuuj” → przejście do ustawień dopasowywania. „Anuluj” → powrót do Startu)

5.1.4. Ustawienia dopasowywania

Wyświetl Ekran Ustawień Dopasowywania



Czy AI jest dostępne? (Jeśli „Nie” – wtyczka pominie włączenie AI i przechodzi do zarządzania regułami)



Czy włączyć AI? (Użytkownik decyduje, czy skorzystać z dopasowania AI. „Tak” → wchodzi do ustawień modelu)



Ustawienia Modelu AI (Trenowanie modelu AI)(Po zakończeniu konfiguracji AI → wracamy do pytania o zarządzanie regułami)



Zarządzać regułami? („Tak” → edycja, dodawanie, usuwanie reguł. „Nie” → przejście do dynamicznego podglądu)



Zarządzaj regułami (Ekran umożliwia tworzenie reguł dopasowania czy warunków łączenia danych)



Dynamiczny podgląd (Funkcja, która na bieżąco pokazuje przewidywane wyniki dopasowania)



Czy zapisać szablon? (Użytkownik może zdecydować, czy aktualną konfigurację zapisać jako szablon. „Nie” → analiza i dopasowanie)



Zapisz szablon

5.1.5. Analiza i Wyniki

Analiza i dopasowanie (Użytkownik inicjuje proces dopasowania)



Przetwarzanie / Generowanie wyników (System lokalny bądź AI w chmurze analizują dane)



Przetwarzanie dopasowania AI (Jeżeli wybrano AI, dane są przesyłane do chmury, a model generuje dopasowania i zwraca wyniki)



Wyświetl Ekran Analizy Wyników (Prezentacja tabeli wynikowej, statystyk i potencjalnych wyjątków)



Podświetl różnice



Przegląd raportu



Czy są wyjątki do rozwiązania? („Tak” → Rozwiązywanie wyjątków. „Nie” → Zapis wyników)

5.1.6. Zapis / Eksport wyników

Zapis wyników (Użytkownik wybiera, czy wstawić wyniki bezpośrednio do arkusza, czy wyeksportować do pliku)



Wstaw wyniki do arkusza (Po zaznaczeniu komórki startowej w Excelu, dane zostają wklejone)



Wstaw przetworzone dane w arkuszu (Dane pojawiają się w arkuszu)



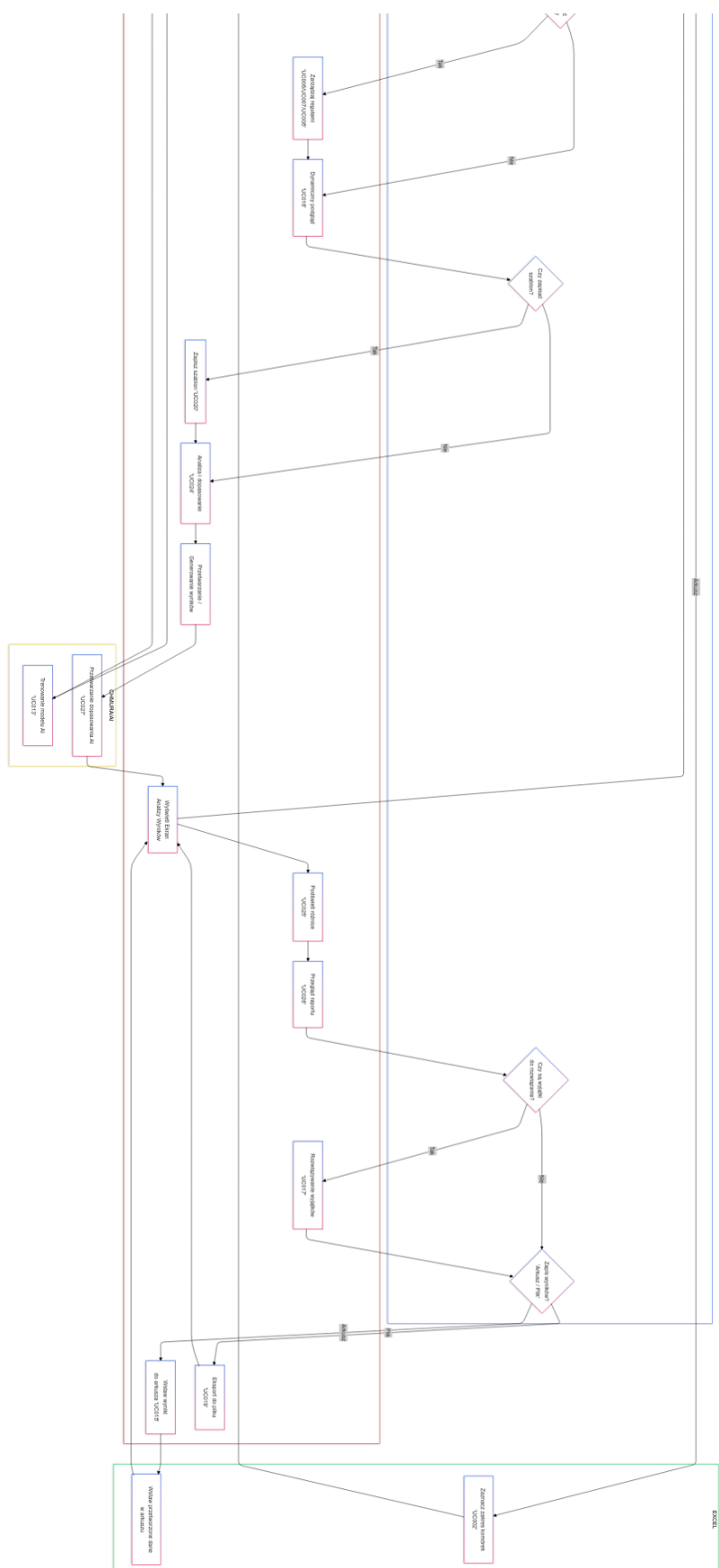
Eksport do pliku



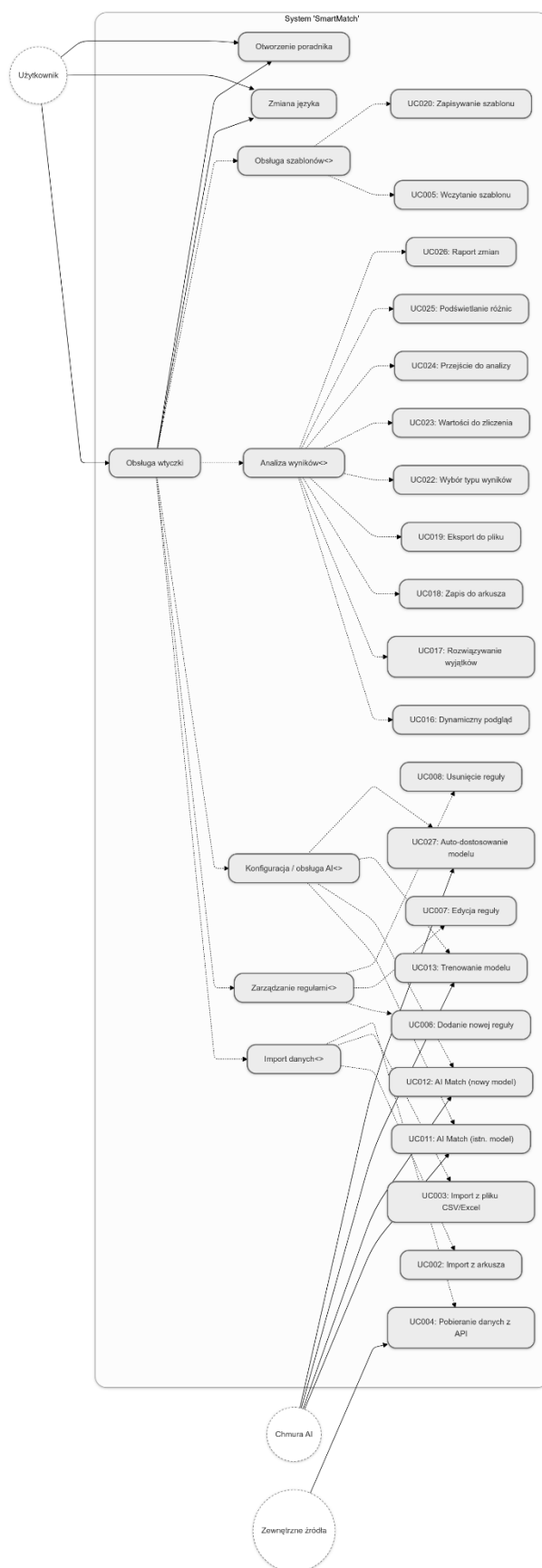
Powrót do Ekranu Wyników



Powrót do Start



6. Diagramy Przypadków Użycia



7. User Stories

1. Zmiana języka interfejsu

[UC001] Jako **użytkownik** chcę **zmienić język interfejsu**, aby móc pracować w języku, w którym najlepiej się czuję.

2. Import danych z arkusza Excel

[UC002] Jako **użytkownik** chcę **zaimportować dane z aktywnego arkusza Excel**, aby szybko wczytać aktualne informacje bez ręcznego kopiowania.

3. Import danych z pliku CSV/Excel

[UC003] Jako **użytkownik** chcę **zaimportować dane z pliku CSV lub Excel**, aby móc pracować z danymi zapisanymi w zewnętrznych plikach.

4. Pobieranie danych z API

[UC004] Jako **użytkownik** chcę **pobrać dane z zewnętrznego API**, aby mieć zawsze aktualne informacje bez ręcznego importu.

5. Wczytywanie szablonów

[UC005] Jako **użytkownik** chcę **wczytać zapisany szablon dopasowywania**, aby zaoszczędzić czas na powtarzalnej konfiguracji.

6. Dodanie nowej reguły

[UC006] Jako **użytkownik** chcę **dodać nową regułę** dopasowywania, aby automatyzować powtarzalne operacje i poprawić jakość łączenia danych.

7. Edycja istniejącej reguły

[UC007] Jako **użytkownik** chcę **edytować istniejącą regułę** dopasowywania, aby dostosować ją do zmieniających się potrzeb.

8. Usunięcie reguły

[UC008] Jako **użytkownik** chcę **usuwać zbędne lub nieaktualne reguły**, aby utrzymać porządek w konfiguracji i uniknąć błędnych dopasowań.

9. Wybór metody dopasowania – Exact Match

[UC009] Jako **użytkownik** chcę **wybrać dokładne dopasowanie**, aby łączyć dane na podstawie identycznych wartości.

10. Wybór metody dopasowania – Fuzzy Match

[UC010] Jako **użytkownik** chcę **wybrać dopasowanie przybliżone**, aby porównywać pola różniące się niewielkimi literówkami lub formatem.

11. Wybór metody dopasowania – AI Match (istniejący model)

[UC011] Jako **użytkownik** chcę **użyć istniejącego modelu AI**, aby wykorzystać dotychczas wypracowane wzorce dopasowania danych.

12. Wybór metody dopasowania – AI Match (tworzenie nowego modelu)

[UC012] Jako **użytkownik** chcę **stworzyć nowy model AI**, aby dopasować dane według aktualnych potrzeb i specyfiki mojego zestawu.

13. Trenowanie modelu na nowych danych

[UC013] Jako **użytkownik** chcę **przetrenować model AI** na nowych danych, aby zwiększyć jego dokładność wraz z rozwojem bazy danych.

14. Dynamiczny podgląd wyników

[UC014] Jako **użytkownik** chcę **oglądać w czasie rzeczywistym zmieniające się wyniki** dopasowywania, aby natychmiast sprawdzać efekty konfiguracji.

15. Rozwiązywanie wyjątków w danych

[UC015] Jako **użytkownik** chcę **przeglądać i poprawiać wyjątki w danych**, aby w razie potrzeby uzupełnić brakujące dopasowania lub zaakceptować sugestie.

16. Zapis wyników do bieżącego arkusza

[UC016] Jako **użytkownik** chcę **zapisać przetworzone dane** w konkretnym miejscu arkusza, aby od razu je wykorzystać w swoim aktualnym skróście Excel.

17. Eksport wyników do pliku CSV/JSON

[UC017] Jako **użytkownik** chcę **zapisać wyniki dopasowywania do pliku CSV lub JSON**, aby użyć ich w innych aplikacjach lub przestać do współpracowników.

18. Zapisywanie szablonu dopasowywania

[UC018] Jako **użytkownik** chcę **zapisać konfigurację dopasowywania** jako szablon, aby móc ją ponownie wczytać przy podobnych projektach.

19. Otworzenie poradnika

[UC019] Jako **użytkownik** chcę **uruchomić poradnik** wtyczki w dowolnym momencie, aby szybko przypomnieć sobie lub nauczyć się obsługi funkcji.

20. Wybór typu wyników

[UC020] Jako **użytkownik** chcę **zdecydować, czy interesują mnie tylko dopasowane wiersze, brakujące dane, czy wszystkie dane**, aby otrzymać właściwy zestaw wyników.

21. Wybór wartości do zliczenia

[UC021] Jako **użytkownik** chcę **wskazać kolumny do sumowania**, aby widzieć łączną wartość dla dopasowanych rekordów.

22. Przejście do analizy wyników

[UC022] Jako **użytkownik** chcę **uruchomić proces dopasowania** i przejść do analizy wyników, aby zobaczyć, jak dane zostały powiązane oraz jakie są rozbieżności.

23. Podświetlanie różnic w wynikach

[UC023] Jako **użytkownik** chcę **włączyć lub wyłączyć oznaczenia wizualne różnic** w tabeli wynikowej, aby łatwiej zidentyfikować miejsca, gdzie dane się nie zgadzają.

24. Przeglądanie raportu zmian

[UC024] Jako **użytkownik** chcę **zobaczyć szczegółowy raport** zmian, aby zrozumieć, co i dlaczego zostało zmodyfikowane w procesie dopasowywania.

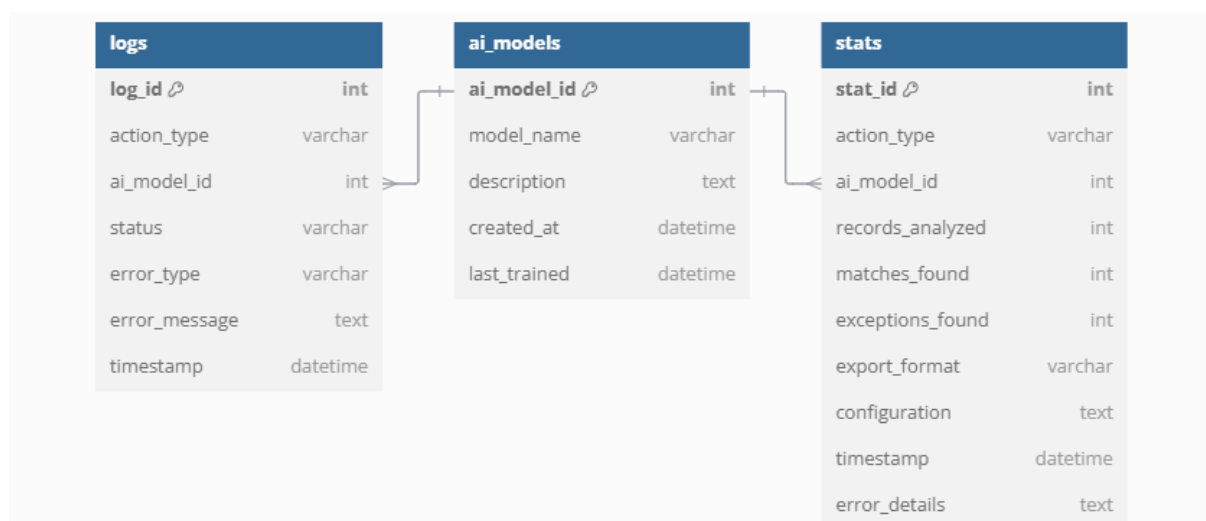
25. Automatyczne dostosowanie modelu AI

[UC025] Jako **model AI** chcę **uczyć się na podstawie wyjątków i akcji użytkownika**, aby w przyszłości skuteczniej dopasowywać dane.

26. Automatyczna aktualizacja danych z zewnętrznego źródła

[UC026] Jako **system** chcę **automatycznie pobierać zaktualizowane dane** z zewnętrznego źródła, aby zapewnić użytkownikowi zawsze najświeższe informacje.

8. Diagram Bazy Danych (ERD)



9. Architektura i Wybór Technologii

9.1 Architektura Systemu

Architektura systemu jest oparta o architekturę klient-serwer, która jest rozszerzona o trzecią warstwę jako chmurę, co finalnie tworzy **Architekturę Trójwarstwową z Chmurą**.

Takie rozwiązanie gwarantuje rozdzielenie odpowiedzialności na Klienta, Backend oraz Chmurę, **zwiększa wydajność** systemu, wykorzystując chmurę do obliczeń AI, oraz gwarantuje działanie podstawowych funkcji w **trybie Offline**.

9. 2 Wybór Technologii

Frontend

- **HTML5:** Język znaczników służący do tworzenia struktury i zawartości stron internetowych. Używany do budowy podstawowego interfejsu użytkownika w aplikacji, aby umożliwić przeglądanie statystyk w przeglądarce, zapewnia kompatybilność z różnymi przeglądarkami oraz wspiera responsywne projektowanie, co zwiększa dostępność aplikacji.
- **CSS:** Kaskadowe arkusze stylów służące do projektowania wyglądu stron internetowych. Odpowiada za estetykę i układ interfejsu użytkownika w raportach i ekranach przeglądania danych w przeglądarce, umożliwiając łatwą modyfikację wyglądu aplikacji, w tym dostosowanie do urządzeń mobilnych.
- **JavaScript:** Język programowania stosowany głównie do obsługi interakcji na stronach internetowych. Używany do dynamicznej aktualizacji interfejsu użytkownika, np. filtrowania logów czy interakcji z raportami. Umożliwia płynne i interaktywne działanie aplikacji, co poprawia doświadczenie użytkownika.
- **Office.js:** Jest oficjalnym SDK Microsoftu dla tworzenia dodatków do Office. Zapewnia pełny dostęp do danych w arkuszu Excel i funkcjonalności programu.

Backend

- **Node.js:** Środowisko wykonawcze JavaScript pozwalające na uruchamianie tego języka po stronie serwera. Odpowiada za backend aplikacji, przetwarzanie żądań, obsługę logów i komunikację z bazą danych. Jest wydajny i umożliwia obsługę dużej liczby równoczesnych połączeń, co jest istotne przy integracji z chmurą
- **SQLite:** Lokalna baza danych, używana do tymczasowego przechowywania konfiguracji i wyników na komputerze użytkownika. Jest lekka, szybka i nie wymaga instalacji dodatkowych serwerów.

Chmura

- **Google Cloud AI Platform:** Platforma chmurowa, oferująca infrastrukturę do trenowania i wdrażania modeli AI. Umożliwia skalowanie obliczeń oraz integrację z innymi narzędziami Google Cloud.
- **Python:** Uniwersalny język programowania, który jest głównym językiem do budowy i trenowania modeli AI. Jego ogromna społeczność i dostęp do bibliotek takich jak TensorFlow sprawiają, że jest to najlepszy wybór.
- **TensorFlow:** Biblioteka do budowy i trenowania modeli AI. Jest zoptymalizowana do obsługi skomplikowanych obliczeń, a jej integracja z Google Cloud AI Platform pozwala na łatwe wdrażanie modeli.

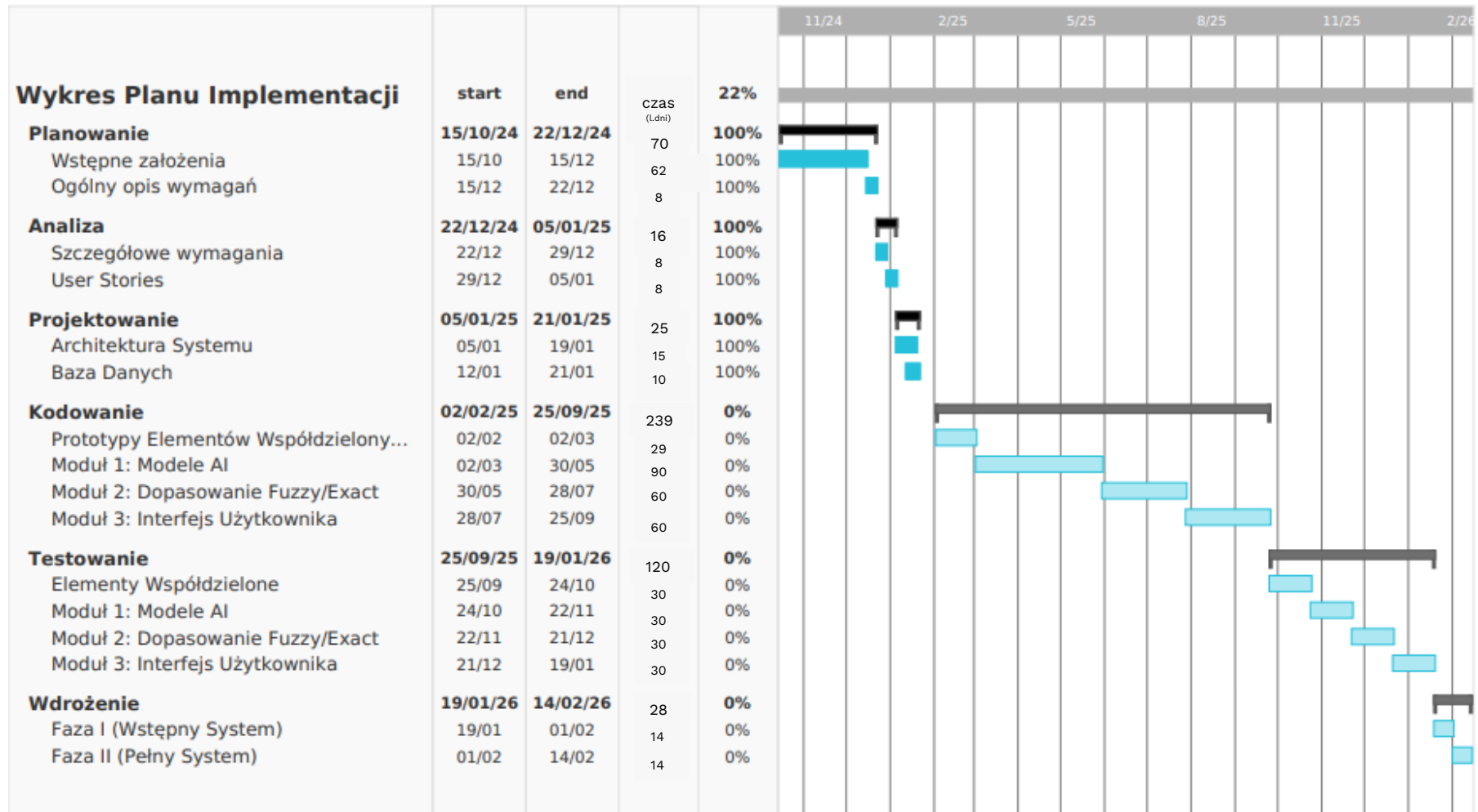
Baza danych

- **PostgreSQL:** jest zaawansowanym systemem zarządzania bazami danych, który zapewnia niezawodność, skalowalność i wysoką wydajność. Jego obsługa JSON i zaawansowane funkcje analityczne czynią go idealnym wyborem do przetwarzania statystyk.

Serwer

- **Serwer VPS:** służy do hostowania bazy danych PostgreSQL oraz backendu Node.js. Dzięki VPS można zarządzać całą infrastrukturą niezależnie od zewnętrznych dostawców. Daje to kontrolę nad dostępnością i bezpieczeństwem danych.

10. Plan Implementacji – Wykres Gantta



11. Analiza Ryzyka

Lp.	Zagrożenie	Wielkość szkody (S)	Prawdopodobieństwo (P)	Ryzyko ($W = S \times P$)	Strategia
1	Brak dostępności usług AI w chmurze	Ciężkie szkody (S=4)– Utrata kluczowych funkcji AI, przestoje w pracy	Mało prawdopodobne (P=2)– Awaria lub brak Internetu raz na kilka lat	W=8 – Dopuszczalna akceptacja ryzyka	<ul style="list-style-type: none"> • Zapewnić tryb off-line w razie niedostępności AI. • Przechowywać kopie kluczowych parametrów lokalnie. • Wyświetlać jasny komunikat użytkownikowi o braku połączenia.
2	Niekompatybilność z różnymi wersjami Excela	Średnie szkody (S=3)– Użytkownicy mogą nie móc uruchomić wtyczki lub uzyskać błędne wyniki	Dość prawdopodobne (P=4)– Zróżnicowane środowisko u różnych klientów	W=12 – Wymaga uwagi i testów wielowersyjnych	<ul style="list-style-type: none"> • Przeprowadzić testy w różnych środowiskach. • Udostępnić dokumentację minimalnych wymagań. • Zapewnić mechanizmy aktualizacji i weryfikacji zgodności
3	Niepoprawny / niepełny import danych	Średnie szkody (S=3)– Potrzeba ręcznej korekty i powtórzenia importu	Możliwe (P=3)– Różne formaty, błędy w plikach, brak kolumn, itp.	W=9 – Dopuszczalna akceptacja ryzyka	<ul style="list-style-type: none"> • Weryfikować formaty w trakcie wczytywania. • Wyświetlać komunikaty o brakujących kolumnach lub nieobsługiwanej formie. • Przechowywać logi błędów i zalecenia naprawcze.
4	Słaba jakość dopasowań	Ciężkie szkody (S=4)– Dużo błędnych wyników, utrata zaufania użytkowników	Możliwe (P=3)– Algorytmy mogą generować fałszywe dopasowania	W=12 – Wymaga ciągłej optymalizacji i nadzoru	<ul style="list-style-type: none"> • Rozwijać i ulepszać algorytmy. • Dodawać reguły i wyjątki.

					<ul style="list-style-type: none"> Raportować skuteczność dopasowania i gromadzić feedback użytkowników.
5	Utrata danych podczas zapisu / eksportu	Ciężkie szkody (S=4)– Konieczność odtwarzania z kopii, brak pełnej historii	Mało prawdopodobne (P=2)– Raz na kilka lat (awaria dysku, przerwane zapisy)	W=8 – Dopuszczalna akceptacja ryzyka	<ul style="list-style-type: none"> Wprowadzić mechanizmy potwierdzania. Jasno informować o lokalizacji zapisu i postępie eksportu. Przetestować scenariusze przerywania zapisu.
6	Błędy w tworzeniu / edycji reguł	Lekkie szkody (S=2)– Błędne wyniki analizy, potrzeba wielokrotnych poprawek	Dość prawdopodobne (P=4)– Użytkownicy mogą tworzyć nieprawidłowe reguły	W=8 – Dopuszczalna akceptacja ryzyka	<ul style="list-style-type: none"> Zapewnić walidację formularza przed zapisaniem reguły. Udostępnić wzorce reguł i przykłady w dokumentacji. Pozwolić na cofnięcie zmian lub wersjonowanie reguł.
7	Brak spójności danych przy aktualizacji (API)	Średnie szkody (S=3)– Konflikty między nowymi a już dopasowanymi rekordami	Możliwe (P=3)– Dane z API mogą się zmieniać w trakcie pracy wtyczki	W=9 – Dopuszczalna akceptacja ryzyka	<ul style="list-style-type: none"> Logować informacje o dacie ostatniej aktualizacji. Umożliwić ręczne zatwierdzanie zmian, jeśli zachodzą konflikty. Przechowywać poprzedni stan danych.
8	Wysoki czas przetwarzania dużych zbiorów danych	Średnie szkody (S=3)– Opóźnienia, frustracja użytkowników, obciążenie systemu	Możliwe (P=3)– Przy dużych plikach lub intensywnym korzystaniu z AI	W=9 – Dopuszczalna akceptacja ryzyka	<ul style="list-style-type: none"> Optymalizować algorytmy. Przesyłać do AI tylko najistotniejsze kolumny.

					<ul style="list-style-type: none"> • Zalecać dzielenie dużych plików na mniejsze porcje.
9	Nieprzyjazny interfejs / brak poradnika	Lekkie szkody (S=2)– Użytkownicy mogą szybko się zniechęcić, potrzebują wsparcia	Możliwe (P=3)– Różne doświadczenie użytkowników, brak intuicyjnego UI	W=6 – Dopuszczalna akceptacja ryzyka	<ul style="list-style-type: none"> • Opracować intuicyjne ekrany i jasne komunikaty. • Dodać moduł „Poradnik”. • Zbierać opinie od użytkowników i stale ulepszać UI.
10	Brak przestrzegania wymagań bezpieczeństwa	Bardzo ciężkie szkody (S=5)– Groźba kar finansowych, utraty reputacji	Mało prawdopodobne (P=2)– Przy założeniu poprawnych procedur i szyfrowania	W=10 – Wymaga zapewnienia zgodności i nadzoru	<ul style="list-style-type: none"> • Szyfrować transmisję i ograniczać gromadzenie danych wrażliwych. • Wprowadzić procedury ochrony danych osobowych. • Monitorować i aktualizować politykę bezpieczeństwa zgodnie z obowiązującymi przepisami.

12. Analiza Konkurencji

Obecnie istnieją konkurencyjne narzędzia na rynku, rozwiązujące podobny problem operacji na danych, jednak w porównaniu z innymi wtyczkami i dodatkami do Excela, projekt ten mocno stawia na **sztuczną inteligencję** i uczenie się oraz skalowalność.

Wtyczka „SmartMatch” wyróżnia się na tle konkurencji w następujący sposób:

- **Power Query**(Narzędzie Microsoft do importu i przekształcania danych) zapewnia Fuzzy Merge, ale **nie** oferuje trenowania AI, dynamicznego rozwiązywania wyjątków, czy obsługi reguł z warunkami.
- **Fuzzy Lookup Add-in**(Bezpłatna wtyczka od Microsoft Research) to rozwiązanie starsze, pozbawione integracji chmurowej i automatycznego uczenia.
- **Ablebits / WinPure**(Komercyjne pakiety dodatków do Excela) dostarczają bogaty zestaw narzędzi do czyszczenia i scalania danych, ale nie oferują w pełni zautomatyzowanego uczenia maszynowego w chmurze.
- **OpenRefine**(Zewnętrzne, otwarte źródłowe narzędzie do oczyszczania i przekształcania danych) oferuje silne lokalne narzędzia do czyszczenia i klastrowania, ale **nie** jest wtyczką Excel i nie korzysta z chmury AI.

13. Historia zmian

Data	Autor	Podsumowanie zmian	Wersja	Znaczenie Zmian
30-12-2024	Andrzej Brzeziński	Pierwsze wydanie zamysłu aplikacji	0.1	KRYTYCZNE
03-01-2025	Andrzej Brzeziński	Dodanie wymagań funkcjonalnych, przypadków użycia, aktorów	0.2	KRYTYCZNE
08-01-2025	Andrzej Brzeziński	Dodanie UserStories i diagramów przypadków użycia	0.3	WAŻNE
19-01-2025	Andrzej Brzeziński	Dodanie diagramu ERD, wykresu Gantt'a, analizy ryzyka i konkurencji	1.0	KRYTYCZNE