

WEBCON

AI Proxy Self-hosted

2026.2.260.63.

historia zmian wersji 2026.2.260.63.

1. Nowości

- Dodano w API v5 etapowy przepływ generowania w AI Process Builder: osobne endpointy do utworzenia sesji, uruchomienia pojedynczego narzędzia oraz wygenerowania komentarza, dzięki czemu klient może wyświetlać postęp i ponawiać pojedynczy krok. Dotychczasowy endpoint nadal działa, ale został oznaczony jako przestarzały.
- Dodano możliwość wyboru modelu (`modelId`) w endpointach AI Process Builder. Model wskazany podczas tworzenia wątku jest używany we wszystkich kolejnych wywołaniach w ramach tego wątku.
- Dodano możliwość wyboru modelu (`modelId`) w endpointach analizy sentymentu, podsumowań i promptów wykorzystywanych przez reguły biznesowe.
- Dodano obsługę atrybutów typu Lista pozycji w AI Process Builder, w tym kolumn, ich typów oraz walidacji.
- Dodano obsługę matrycy atrybutów w AI Process Builder, w tym kolumn List pozycji oraz uprawnień do kolumn.
- Dodano deterministyczne parsowanie diagramów BPMN w AI Process Builder: struktura procesu jest odczytywana bez udziału modelu, natomiast model odpowiada za scalanie i nazewnictwo kroków.
- Agent AI może korzystać z danych zawartych w plikach statycznych dołączonych do procesu.
- Dodano opis i kategorię pliku do kontekstu załączników przekazywanego do modelu.
- Dodano Helm chart umożliwiający uruchamianie AI Proxy Self-hosted w klastrze Kubernetes, w tym k3s i AKS.

2. Ulepszenia

- AI Proxy Self-hosted obsługuje uwierzytelnianie JWT z użyciem ClientId/ClientSecret jako alternatywę dla mTLS oraz osobne certyfikaty dla każdej roli (HTTPS, mTLS, JWT).
- Helm chart obsługuje teraz:
 - przełączanie między mTLS a JWT,
 - osobne certyfikaty dla każdej roli,
 - sekrety przekazywane z Kubernetes Secrets zamiast wpisywania ich bezpośrednio w konfiguracji,
 - komunikację wyłącznie po HTTPS,
 - walidację wartości przed wdrożeniem.
- AI Concierge ustawia wartości pól wyboru wyłącznie wtedy, gdy wynikają one bezpośrednio ze źródła, zamiast określać je na podstawie luźnych skojarzeń.
- AI Concierge i Agent AI podczas wypełniania formularza uwzględniają instrukcje projektanta zawarte w opisach pól.
- Tłumaczenia AI są teraz ponawiane i przełączane na zapasowego dostawcę osobno dla każdej paczki. Paczka odrzucona z kodem HTTP 429 jest od razu przekazywana do kolejnego dostawcy, natomiast w przypadku innych błędów jest ponawiana maksymalnie 3 razy, dzięki czemu pojedyncza nieudana paczka nie jest już pomijana w wyniku.
- Wyłączono myślenie modelu dla podstawowych operacji AI (sentiment, podsumowanie, prompt, tłumaczenie, transkrypcja audio i analiza plików), co znacząco skraca czas odpowiedzi modeli z włączonym myśleniem.
- Myślenie modeli Gemini jest teraz ograniczone do 2048 tokenów. Dzięki temu czas przetwarzania jest bardziej przewidywalny, również podczas generowania procesu w AI Process Builder z pliku BPMN.
- Podczas generowania metadanych AI Process Builder z plików BPMN do modelu wysyłana jest zwięzła transkrypcja procesu zamiast surowego XML, co znacząco zmniejsza zużycie tokenów i skraca czas odpowiedzi.
- Generator obiegu AI Process Builder nie pozostawia już pustych faz, których nie było w opisie aplikacji.
- Generator obiegu AI Process Builder umieszcza negatywne kroki końcowe za krokami pozytywnymi.

3. Naprawa błędów

- Tokeny myślenia modeli Gemini są teraz uwzględniane w zużyciu. Dostawca rozlicza je jako tokeny wyjściowe, dlatego raportowane zużycie może być wyższe niż wcześniej.
- Naprawiono błędy Agenta AI i AI Concierge występujące podczas korzystania z modeli Azure AI Foundry (HTTP 404 „Response not found”).
- Naprawiono konwersję BPMN:
 - bramki z decyzją użytkownika były konwertowane na osobne kroki zamiast na ścieżki,
 - ścieżki powrotne do scalonych kroków były pomijane.
- Naprawiono pozostawianie pustych pól Tak/Nie przez AI Concierge i Agenta AI, gdy model zwracał wartości takie jak Tak, yes czy 1.
- Naprawiono integrację AI Concierge z LiteLLM:
 - odpowiedzi z pustymi licznikami tokenów były odrzucane,
 - załączniki inline były przekazywane niepoprawnie.

WEBCON

AI Proxy Self-hosted

2026.2.135.35.

historia zmian wersji 2026.2.135.35.

1. Nowości

- Dodano webowy interfejs administracyjny do konfiguracji dostawców AI, połączeń i modeli. Interfejs jest zawsze dostępny w trybie Self-hosted.
- Dodano konfigurowalny limit czasu połączenia z dostawcą AI (`TimeoutSeconds`). Po jego przekroczeniu wywołanie jest przerywane, a żądanie przekazywane do kolejnego skonfigurowanego modelu.

2. Ulepszenia

- Zaktualizowano integrację z Gemini, zastępując dotychczasową bibliotekę pakietem Google Gen AI SDK.
- Zaktualizowano domyślną konfigurację modeli: dodano model Gemini 2.5 Pro oraz określono daty wygaśnięcia modeli Gemini planowanych do wycofania.
- Usprawniono odwzorowywanie przez asystenta projektowania procesów (Designer Desk) obiegów i modeli danych opisanych w załączonych plikach. Asystent wiernie odwzorowuje załączoną definicję zamiast dodawać nieopisane pola, a pliki przetwarza tylko raz, co zmniejsza zużycie tokenów.
- Ulepszono prompty asystenta projektowania procesów i przełączono go na model Gemini 2.5 Flash.
- Dodano skrypty i dokumentację umożliwiające lokalne budowanie i uruchamianie AI Proxy w trybie Self-hosted z użyciem Dockera. Materiały są dostępne w repozytorium [WEBCON-BPS/WebCon-AiProxy-GettingStarted](#).

3. Naprawa błędów

- Naprawiono przelączenie na kolejnego dostawcę, gdy model stale zwracał błąd przekroczenia limitu żądań (429). Po wyczerpaniu ponowień żądanie jest teraz przekazywane do kolejnego modelu zamiast kończyć się błędem.
- Naprawiono generowanie przez asystenta projektowania procesów niechcianych elementów obiegu lub pól formularza, gdy zadanie dotyczyło tylko jednego z tych obszarów. Narzędzia projektowe są teraz dobierane na podstawie treści zadania.
- Naprawiono niespójne odwołania między fazami, krokami, ścieżkami i źródłami danych w komendach projektowania procesu po przenumerowaniu identyfikatorów oraz błąd unikalności identyfikatorów występujący podczas dodawania elementów.
- Naprawiono powielanie przez asystenta projektowania procesów kroku startowego i faz zamiast aktualizowania istniejących elementów oraz poprawiono generowane nazwy ścieżek.
- Naprawiono błąd parsowania kolekcji `PickersToSet` w komendach projektowania procesu.
- Naprawiono błąd tworzenia kolekcji podczas wykonywania tłumaczeń AI z użyciem modelu gpt-4.1-mini.
- Naprawiono błąd pobierania tokena AI Proxy spowodowany brakiem wartości limitu w bazie danych.
- Naprawiono wiązanie i wczytywanie konfiguracji modeli podczas komunikacji z modelami AI za pośrednictwem proxy działającego w trybie Self-hosted.
- Naprawiono błąd „Parameter cannot be null” występujący w zadaniu czyszczenia sesji wykonywanym na koniec dnia.

WEBCON

AI Proxy Self-hosted

2026.2.64.17.

historia zmian wersji 2026.2.64.17.

1. Nowości

- Wprowadzono obsługę oznaczania modeli jako wygasłe, dzięki czemu AI Proxy może poprawnie rozpoznawać i obsługiwać modele, które nie powinny być już używane.
- Dodano możliwość wyboru modelu w endpointach v4 dla AI Concierge, tłumaczeń, transkrypcji oraz sesji Agentów AI.
- Dodano zwracanie informacji o użytym modelu w wyniku operacji AI.
- Udostępniono endpoint pozwalający pobierać dostępne modele wraz z ich identyfikatorami.
- Dodano obsługę `Chat with Tools` dla modeli OpenAI i Gemini.
- Dodano connector zgodny z OpenAI API, umożliwiający integrację m.in. z LiteLLM, LM Studio, OpenRouter i podobnymi dostawcami.
- Dodano obsługę przesyłania plików inline dla OpenAI, Azure, Gemini oraz connectora zgodnego z OpenAI API.
- Dodano możliwość uwierzytelniania w Azure za pomocą `DefaultAzureCredential` jako alternatywy dla klucza API.
- Dodano obsługę OIDC w usługach korzystających z klienta OpenAI.

2. Ulepszenia

- Usprawniono wczytywanie i walidację konfiguracji AI Proxy, w tym obsługę identyfikatorów modeli oraz konfiguracji dostawców.
- Zmieniono nazwę typu `MethodType` na grupujące `AiTaskType` w pliku `aiconfiguration.json`, aby lepiej odzwierciedlała przeznaczenie konfigurowanej operacji AI.
- Zmieniono nazwę węzła `Authorization` na `OIDcAuthorization`, doprecyzowując konfigurację autoryzacji OIDC.
- Ujednolicono działanie connectora zgodnego z OpenAI API z obsługą OpenAI, w tym mechanizmy czyszczenia danych oraz obsługi plików.
- Usprawniono przekazywanie i obsługę kluczy API w usługach korzystających z OpenAI.
- Dodano logowanie informacji o modelu, na którym wykonano zapytanie AI.
- Usprawniono obsługę anulowania operacji w `Chat with Tools`.
- Ulepszono odpowiedzi AI Concierge podczas nawigowania oraz obsługę treści zwracanych przez narzędzia.

3. Naprawa błędów

- Naprawiono błąd przesyłania plików do Azure przez przejście na klienta OpenAI.
- Naprawiono obsługę załączników .txt w rozmowach Gemini chat.
- Naprawiono uwzględnianie plików w connectorze OpenAI.
- Naprawiono parsowanie komend AI Concierge w connectorze zgodnym z OpenAI API.
- Naprawiono wyjątek występujący przy braku klucza API w `OpenAiService`.
- Naprawiono obsługę endpointów v4 w przypadku braku parametru `method`.
- Naprawiono obsługę promptów w wybranych scenariuszach AI Concierge.
- Naprawiono nazwę transkrypcji w enumie.

WEBCON

AI Proxy Self-hosted

2026.1.27.108.

historia zmian wersji 2026.1.27.108.

1. Nowości

- Dodano endpoint Concierge v3 z obsługą tooli nawigacyjnych, dzięki czemu **AI Concierge** jest dostępny w całym Portalu, może wspierać poruszanie się po nim oraz sugerować przejścia do odpowiednich miejsc. Sama nawigacja jest wykonywana po kliknięciu przycisku przejścia. Rozszerzono także obsługę o komendę `//goto`, która wymusza użycie mechanizmu nawigacji. Ta zmiana wymaga do działania systemu WEBCON w wersji 2026.1.6 lub wyższej.

WEBCON

AI Proxy Self-hosted

2026.1.24.106

historia zmian wersji 2026.1.24.106

1. Nowości

- Dodano obsługę connectora Azure AI Foundry w AiProxy wraz z logiką AIConnectora. Doprecyzowano także wymagania konfiguracji zasobów i dostępu sieciowego po stronie Azure potrzebne do uruchomienia oraz lokalnego testowania integracji.
- Zmieniono sposób numeracji wersji AI Proxy Self-hosted. Numer wersji jest teraz wyliczany automatycznie i wskazuje zgodność z daną wersją WEBCON.

WEBCON

AI Proxy Self-hosted

1.0.0.235

historia zmian wersji 1.0.0.235

1. Nowości

- Opublikowano **AI Proxy Self-hosted** – nowy komponent, który umożliwia klientom korzystanie z funkcjonalności AI w WEBCON w oparciu o własną infrastrukturę, z całkowitym pominięciem infrastruktury WEBCON. Rozwiązanie jest dostępne w licencji subskrypcyjnej Enterprise oraz w licencji wieczystej WEBCON Agentic AI Framework.

Udostępniono możliwość instalacji AI Proxy Self-hosted w środowisku klienta (on-premise lub prywatna chmura) wraz z opcją podpięcia i konfiguracji własnych modeli AI – hostowanych w Google Vertex, OpenAI lub Azure AI Foundry. Dzięki temu klienci zyskują pełną kontrolę nad danymi i infrastrukturą wykorzystywaną do przetwarzania AI, co ułatwia spełnienie wymogów bezpieczeństwa, polityk prywatności oraz regulacji branżowych.

Wprowadzono dedykowany tryb AI Proxy Self-hosted oraz nowy węzeł konfiguracyjny, dodany w plikach konfiguracyjnych `appsettings.json` wykorzystywanych do uruchamiania AI Proxy, dla tego trybu, obejmujący:

- możliwość włączenia użycia AzureKeyVault do przechowywania danych wrażliwych,
- zmienną przechowującą ścieżkę dostępu do certyfikatu .pem lub .pfx oraz hasło.

Przygotowano także alternatywną implementację klas, która w tym trybie umożliwia zastąpienie Azure Key Vault zmiennymi środowiskowymi.

Zaprojektowano i udostępniono nowy, ujednolicony model konfiguracji dostawców AI oparty o priorytety. Umożliwia on definiowanie wielu dostawców, automatyczne przełączanie na dostawców o niższym priorytecie oraz przypisywanie różnych modeli AI do poszczególnych metod (modele tekstowe, obrazowe i audio). Konfiguracja jest wykonywana w pliku `aiconfiguration.json`, który jest wczytywany podczas uruchamiania aplikacji AI Proxy.

Dodano możliwość wyboru trybu działania aplikacji w pliku konfiguracyjnym oraz obsługę lokalnych certyfikatów (.pem/.pfx) wykorzystywanych do podpisywania i weryfikacji tokenów dostępowych – bez konieczności komunikacji z zewnętrznymi usługami autoryzacyjnymi (w tym z LAS) czy bazami danych. W trybie AI Proxy Self-hosted ograniczono zależność od zewnętrznych komponentów infrastrukturalnych: aplikacja może działać bez połączeń do usług licencyjnych, baz danych oraz magazynów danych w chmurze, a wymagane mechanizmy są dobierane automatycznie na podstawie konfiguracji.