

Zobacz WSZYSTKIE dane

Kluczowe korzyści:

Zakoduj wiedzę organizacji

- Zapisywanie i współdzielenie ścieżek analizowania danych innym użytkownikom.
- Szybki, wielowymiarowy scoring do nadawania priorytetów dowolnym obiektom i zdarzeniom.
- Graficzne eksplorowanie danych z wykorzystaniem analizy połączeń, map, histogramu, wizualnych zapytań i analizy szeregów czasowych.
- Łatwa i bezpieczna współpraca użytkowników w ramach zakresu ich uprawnień.

Wgląd we WSZYSTKIE dane

- Natychmiastowy dostęp do wszystkich danych zintegrowanych w jednym miejscu.
- Dopasowanie i przeszukiwanie dziesiątków / setek źródeł danych.
- Szybkie analizy milionów / miliardów rekordów (big data, enterprise data).
- Szybkie dodawanie zbiorów danych, łączenie i synchronizowanie wewnętrznych i zewnętrznych źródeł nie wymaga umiejętności technicznych.

Praktyczne rozwiązanie

- System oferowany jako gotowy produkt „out-of-the-box”.
- Rozwiązanie w rozsądnej cenie
- Wykorzystanie sprzętu typu commodity
- Wyniki w ciągu kilku dni / godzin, a nawet minut
- Nie wymaga usług konsultingowych przed wdrożeniem
- Intuicyjne wizualizacje i graficzne widoki
- Obsługa nie wymaga dedykowanych umiejętności.
- Stałe koszty utrzymania i szybki zwrot z inwestycji
- Obsługa we własnym zakresie po wdrożeniu (self-service)

Wyzwanie

Duże organizacje tj. banki czy firmy ubezpieczeniowe mierzą się z problemem analiz dużej liczby danych rozproszonych w wielu źródłach o niejednorodnej strukturze.

Dostęp do danych, ich normalizacja, uspojnianie i analiza to proces, który na ogół jest bardzo czasochłonny i problematyczny i stanowi nie lada wyzwanie dla każdej osoby pracującej z danymi.

Rozwiązanie

Platforma analityczna DataWalk pozwalająca na łączenie wielu dużych, złożonych zbiorów danych, zarówno wewnętrznych jak i zewnętrznych, w jednym środowisku analitycznym umożliwiającym ich szybką, wizualną analizę.

Interaktywne analizy w DataWalk wykonywane są natychmiast, bez pomocy IT i bez znajomości SQL-a. W przeciwieństwie do tradycyjnych systemów analitycznych, opartych w głównej mierze na usługach (konsultingowych, wdrożeniowych), DataWalk dostarcza pierwsze wyniki w ciągu zaledwie kilku dni od rozpoczęcia projektu.

Uwolnienie wiedzy

Dzięki funkcjonalności Szybkich Analiz, każdy kto posiada odpowiednie uprawnienia, może jednym kliknięciem zobaczyć efekty analiz i eksploracji danych stworzonych przez najlepszych specjalistów. Następnie takie analizy można dowolnie edytować bez konieczności przechodzenia całej ścieżki analitycznej od początku.

Dzięki temu użytkownicy biznesowi mogą w pełni korzystać z możliwości programu DataWalk i wykorzystywać wiedzę ekspertów i analityków dziedzinowych.

Identyfikowanie właściwych danych

W DataWalk wystarczy kilka prostych kroków, aby zredukować miliardy rekordów do konkretnych, interesujących, z punktu widzenia użytkownika, danych. Umożliwia to DataWalk Universe Viewer, który jest interaktywnym interfejsem do zwinnego eksplorowania danych, dającym wgląd w pełen obraz relacji pomiędzy nimi (np. szkoda-przedmiot umowy-umowa-klient-produkt-reklamacje).

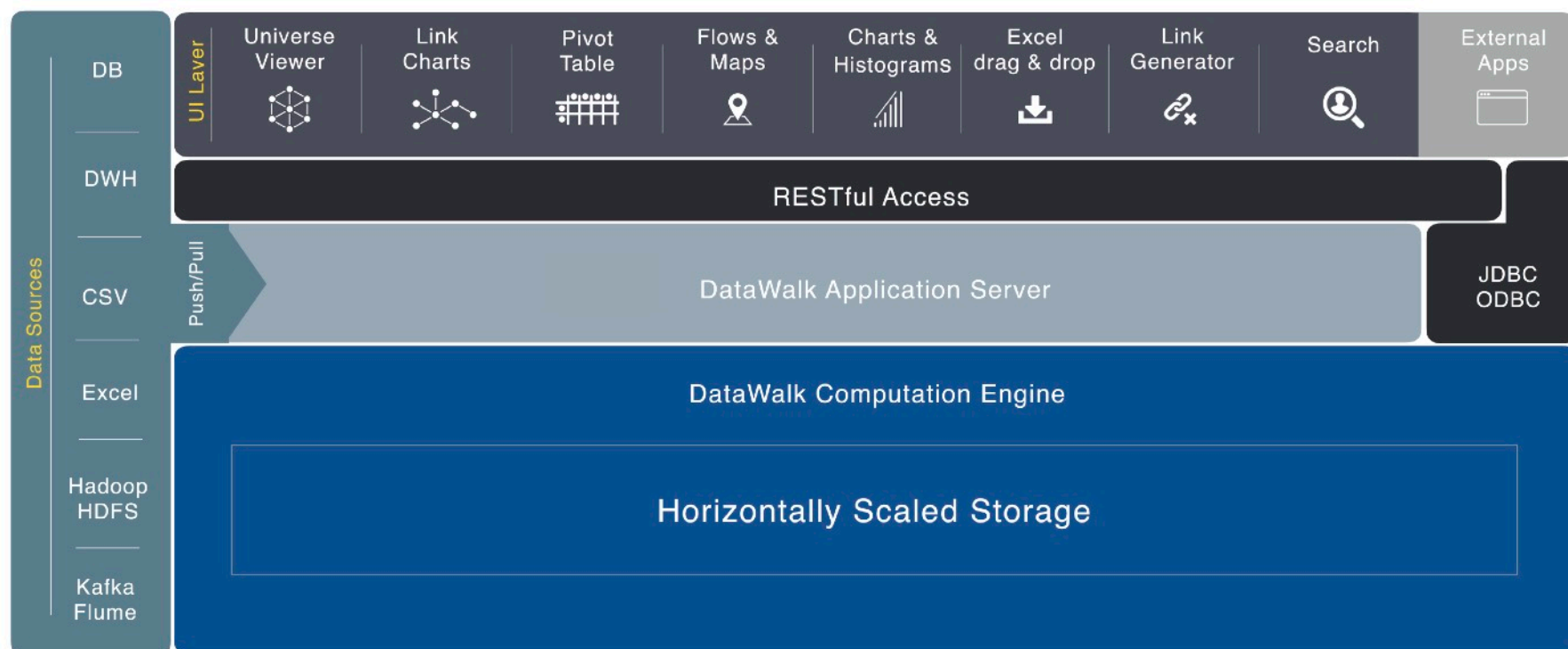
DataWalk Universe Viewer umożliwia łatwe przechodzenie od kluczowych wskaźników efektywności (KPI) do szczegółowych danych.

Szybki, wielowymiarowy scoring

DataWalk wykracza poza tradycyjne metody scoringowe. Wyniki są generowane natychmiast na podstawie dowolnej liczby zaawansowanych analiz, a ocenie może podlegać dowolny obiekt (np. ludzie, nieruchomości, działania, podatnik, faktura, wpłata, lokalizacje, itp.).

Przykładowo, można szybko wygenerować ocenę ryzyka mającą na celu wyłapanie potencjalnych oszustów (lub zagrożeń) na podstawie transakcji, powiązań między osobami, numerami identyfikacyjnymi lub dowolnymi zdarzeniami w czasie.

Proces budowy scoringu nie wymaga tworzenia osobnych hurtowni dziedzinowych tzw. data martów i może być realizowany metodą ekspercką lub w połączeniu z algorytmami statystycznymi.



Rys. 1: Schemat systemu DataWalk

Obiekty wytypowane przez scoring mogą podlegać dalszej inwestygacji w DataWalk, jak również mogą być automatycznie wyeksportowane do zewnętrznych systemów.

Łatwa i bezpieczna współpraca

DataWalk to system stworzony po to, aby ułatwić współpracę wewnątrz i na zewnątrz organizacji. DataWalk został zaprojektowany na bazie szczegółowego schematu uprawnień, co oznacza, że użytkownik widzi tylko te dane i powiązane z nimi analizy, do których zostały nadane mu uprawnienia.

Potężne narzędzie analityczne

DataWalk zawiera cały zestaw w pełni zintegrowanych narzędzi analitycznych, zaprojektowanych tak, aby zapewnić lepsze wyniki i głębszy wgląd we wszystkie dostępne dane.

Jednym z unikatowych narzędzi DataWalk jest Universe Viewer, który zapewnia interaktywną wizualizację danych oraz możliwość tworzenia zarówno podstawowych, jak i złożonych zapytań analitycznych.

Oprócz tego, w skład DataWalk wchodzi:

- Histogramy
- Analizy połączeń
- Analizy geoprzestrzenne
- Analizy szeregów czasowych
- Tabele przestawne
- Interaktywne wykresy

Łatwe synchronizowanie i przeszukiwanie wielu źródeł danych

DataWalk został zaprojektowany z myślą o obsłudze środowisk składających się z wielu różnorodnych źródeł z dużym wolumenem danych. Dlatego z DataWalk różna struktura lub zróżnicowane opisy danych w wielu systemach źródłowych przestają być wyzwaniem dla analityków.

DataWalk pozwala elastycznie mapować dane na biznesową, logiczną strukturę danową, identyfikując najkrótsze możliwe połączenia pomiędzy zbiorami danych i błyskawicznie dostarczając interaktywne analizy, nawet dla miliardów rekordów.

Łączenie i wykorzystanie danych z dowolnych źródeł

DataWalk pozwala łatwo łączyć dane z baz wewnętrznych i zewnętrznych.

Możliwy jest także import i podłączenie do środowiska DataWalk dowolnych arkuszy Excel, a następnie analizowanie ich w trybie prywatnym. Ma to szczególne znaczenie, gdy badane są specyficzne sytuacje i do standardowo wykorzystywanych danych dodawane są nowe źródła w celu przetestowania dowolnej hipotezy biznesowej lub śledczej. Co istotne, współdzielenie wyników z takich analiz w organizacji jest niewiarygodnie łatwe.

Optymalne rozwiązanie

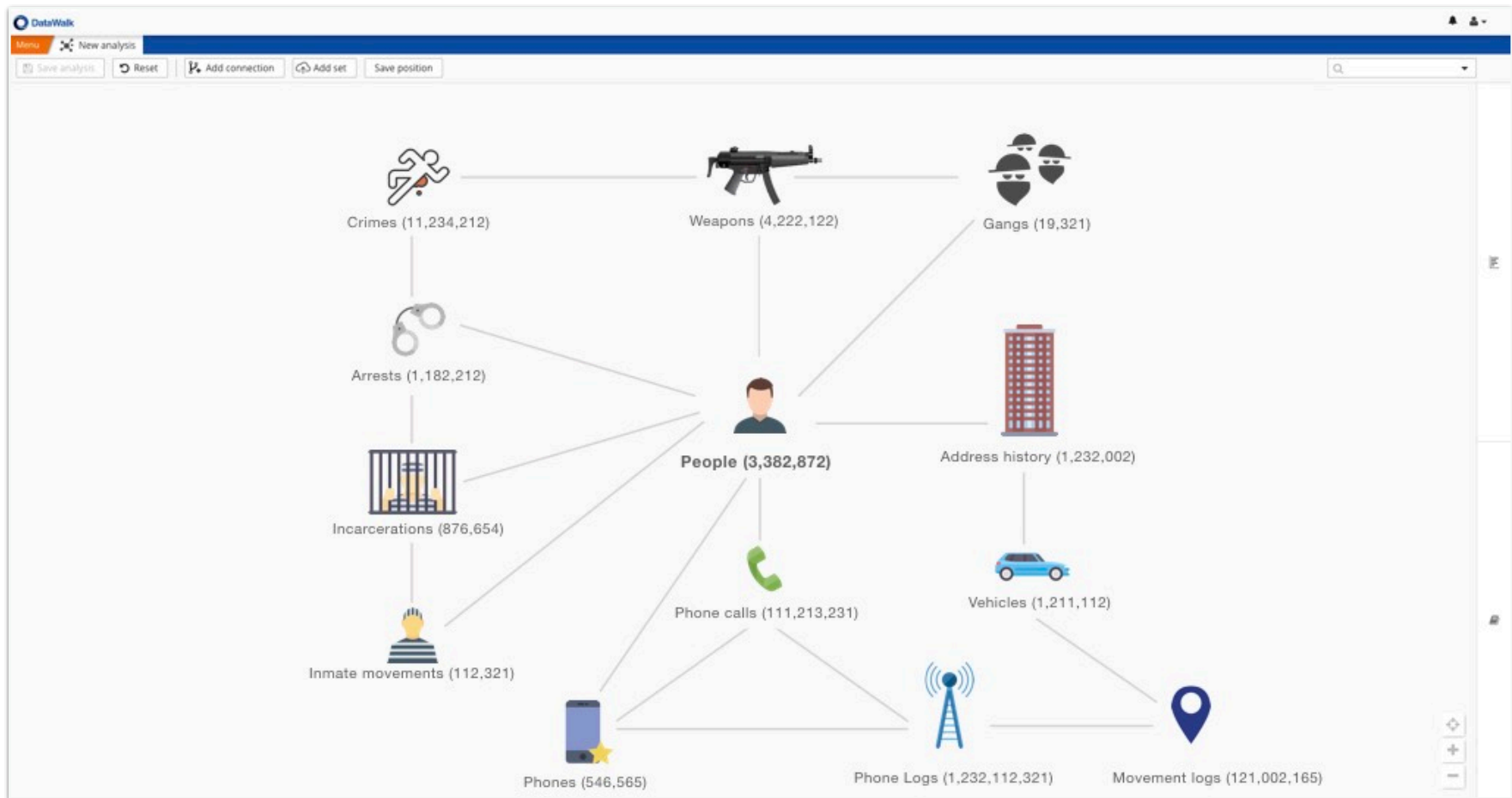
DataWalk to skalowalna platforma działająca na serwerach typu commodity, co jest ekonomicznym rozwiązaniem w porównaniu do innych skalowalnych systemów analitycznych klasy Enterprise. Oprogramowanie DataWalk może być wdrożone stacjonarnie, na miejscu u klienta lub w chmurze.

Bez konsultingu przedwdrożeniowego i przy minimalnych nakładach po stronie IT

Wdrożenie DataWalk nie wymaga długich, profesjonalnych usług konsultingowych oraz zatrudniania osób o dedykowanych, specjalistycznych umiejętnościach, co znacząco wpływa na obniżenie kosztów uruchomienia i utrzymania systemu.

Po wdrożeniu, DataWalk zazwyczaj wymaga minimalnego wsparcia IT.

Użytkownicy mogą natychmiast uzyskać dostęp do wszystkich danych, do których są upoważnieni. W przeciwieństwie do alternatywnych systemów, zaawansowani administratorzy mogą modyfikować analityczną strukturę danych przy użyciu prostych, graficznych interfejsów, bez konieczności zlecenia dodatkowych prac technicznych.



Rys. 2: Universe Viewer – wizualne środowisko analityczne

Pierwsze wyniki w kilka godzin lub dni

W przeciwieństwie do alternatywnych rozwiązań, które dostarczają wyniki po kilku miesiącach, a nierzadko nawet latach, DataWalk daje pełen obraz analizowanych danych już w ciągu pierwszych kilku dni, a czasami nawet kilku godzin - dotyczy to zaawansowanych, wieloźródłowych projektów w środowiskach danych klasy Enterprise.

Dodatkowo konfiguracja DataWalk zazwyczaj nie wymaga umiejętności technicznych i dostępna jest z poziomu interfejsu użytkownika biznesowego.

Jednolita, w pełni zintegrowana platforma analityczna

DataWalk to w pełni zintegrowana platforma, oferowana jako gotowy produkt. Oznacza to, że nie ma potrzeby łączenia wielu elementów lub modułów, aby otrzymać kompletne rozwiązanie. DataWalk jest dostarczany wraz z gotową do użycia analityką, wieloma regułami eksperckimi i systemem do przechowywania danych.

Intuicyjne wizualizacje i graficzne prezentowanie danych

DataWalk jest platformą analityczną, która nie wymaga specjalistycznej wiedzy w zakresie SQL-a, jakiegokolwiek języka programowania ani umiejętności pisania skryptów. Zbiory z danymi przedstawione są w formie ikon, a relacje pomiędzy nimi za pomocą linii. „Chodzenie” po danych i ich analiza odbywa się za pomocą intuicyjnego, interaktywnego interfejsu, dzięki czemu nawet najbardziej złożone zapytania mogą być wykonywane przez użytkowników o profilu biznesowym, a prezentowany wynik w formie tabeli, wykresu czy historii wykonanych kroków jest łatwy do interpretacji.

Dostęp do wiedzy eksperckiej

Naszym celem jest osiągnięcie sukcesu zespołów, z którymi pracujemy. Dlatego przy każdym wdrożeniu zapewniamy konsultantów o szerokiej wiedzy technicznej i specjalistycznej. Dzięki temu osiągamy maksymalne wyniki, rozwiązując najtrudniejsze wyzwania w zakresie inwestygowania i eksplorowania złożonych zbiorów danych.

Zwinność i łatwa adaptacja

Wiele tradycyjnych systemów jest sztywnych i wymaga miesięcy profesjonalnych usług, aby np. dodać nowe źródło danych czy zmienić fizyczną strukturę środowiska analitycznego. DataWalk działa zupełnie inaczej.

Dodawanie i integrowanie nowego źródła może być wykonywane ciągu kilku minut lub max. godzin, a struktura danych (np. rodzaje i warunki połączeń między danymi) może zostać zmodyfikowana za pomocą kilku kliknięć i predefiniowanych reguł i algorytmów.

Wiarygodne wyniki nawet na brudnych danych

Tradycyjne narzędzia analityczne nie działają dobrze na środowiskach, w których występują błędy, braki lub niespójności w danych, lub co sprawia, że ich wdrożenie musi być poprzedzone długimi i kosztownymi projektami czyszczenia danych.

DataWalk - przykładowe wyniki

DataWalk - przykładowe wyniki				
6	4B	4TB	40B	1
źródła danych	rekordów	danych	połączeń	node
Rodzaj działania	Tradycyjne podejście	z DataWalk		
wdrożenie	wiele miesięcy / lata	tygodnie		
zmiana w logicznym modelu danych	1-2 tygodnie	minuty / godziny		
dodanie nowego źródła danych	1 miesiąc / 1 tydzień	1-2 dni, a nawet kilka godzin		
stworzenie nowej reguły alarmowej	kilka dni	minuty / godziny		
stworzenie nowej, złożonej analizy z nowym źródłem danych	godziny / dni	minuty / godziny		
złożone zapytania - czas obliczania	godziny	60x szybciej		
konieczność programowania?	tak, godziny	nie		
SQL / pisanie skryptów?	tak, godziny	nie		

Tabla 1: Przykładowe wyniki

DataWalk pozwala szybko wskazać niespójności, profilować oraz łączyć dane i na bazie złożonych reguł wyodrębniać z duplikatów unikatowe rekordy, bez angażowania działów IT lub osób odpowiedzialnych za utrzymanie systemów źródłowych. Co najważniejsze, same dane źródłowe nie zostają naruszone w tym procesie i nadal można je łatwo zidentyfikować w systemach źródłowych.

Unikatowa technologia DataWalk

System DataWalk jest zbudowany na bazie zestawu opatentowanych oraz będących w trakcie procedury patentowej technologii.

Zamiast złożonych JOIN'ów, DataWalk wykorzystuje do łączenia i analizowania danych tzw. linki (czyli najkrótsze możliwe relacje pomiędzy danymi).

Pozwala to utrzymać liniową wydajność w miarę zwiększania złożoności obliczeniowej. zaawansowanych zapytań analitycznych.

Linki są generowane tylko raz, a złożone zapytania / analizy dzielone są na mniejsze operacje liniowe, co znacząco zwiększa wydajność obliczeniową dla

W DataWalk aby eksplorować dane oddalone (np. klient, polisa, szkoda, dane geolokalizacyjne z aplikacji mobilnej) nie trzeba przeprowadzać ich denormalizacji. Dzięki temu analizy dużych środowisk – z wieloma miliardami rekordów – możliwe są na serwerach typu commodity o standardowej mocy obliczeniowej.

Przykładowe aplikacje

- Przeciwdziałanie nadużyciom
- Analizy procesu likwidacji szkód
- Analizy ryzyka
- Analizy procesu sprzedaży
- Windykacja
- Przeciwdziałanie praniu brudnych pieniędzy
- Inwestygacje śledcze (służby mundurowe i wywiadowcze)

Specyfikacja

Architektura systemu	<ul style="list-style-type: none">• Współdzielony system dla wielu użytkowników• Skalowalne, współdzielone, raz wdrażane repozytorium danych• Dwie warstwy: UI i silnik obliczeniowy (Computation Engine)
Universe Viewer i histogramy	<ul style="list-style-type: none">• Proste, jednorodne, wizualne środowisko do importowania, łączenia i analizy danych z wielu dużych zbiorów• Użytkownicy o uprawnieniach administratora mogą łatwo zmodyfikować strukturę środowiska analitycznego, za pomocą zaledwie kilku kliknięć.• Użytkownicy mogą natychmiast dodać plik Excel do Univer Viewer'a i połączyć ten plik z pozostałymi danymi, korzystając z prywatnej przestrzeni analitycznej.• Każdy użytkownik może natychmiast generować histogramy dla dowolnych zestawów danych.• Histogramy mogą być wykorzystywane do profilowania i / lub filtrowania danych.• Breadcrumbs'y (historia kroków w analizie) umożliwiają łatwą wizualizację, modyfikację i ponowne wykorzystanie analiz.• Użytkownicy mogą wykonywać zapytania ad hoc i eksplorować dane, przechodząc od jednych, dowolnych danych do drugich• Użytkownicy nietechniczni mogą łatwo wykonywać złożone zapytania za pomocą wizualnego interfejsu, filtrując i łącząc różne zestawy danych i ścieżki między nimi.• Połączenia między zbiorami danych mogą opierać się na dowolnej kombinacji porównań liczbowych, przybliżonego/bezpośredniego dopasowania tekstu (np. miasto = miasto), godziny, daty, odległości geograficznej, odległości Levenshteina i Soundex.• Zestawy danych można podłączyć wykorzystując klucz główny, klucz obcy, a także za pomocą dowolnego warunku logicznego czy biznesowego, kiedy nie istnieje bezpośredni warunek łączący.
Szybkie analizy	<ul style="list-style-type: none">• Łatwy dostęp do wiedzy poprzez ponowne wykorzystanie zapisanych ścieżek analizy przeprowadzonych w Universe Viewer.
Analiza połączeń	<ul style="list-style-type: none">• Umożliwia analizę sieci, geo-przestrzenną i analizę szeregów czasowych• Wiele układów, w tym strukturalnych, promieniowych i hierarchicznych• Analiza sieci społecznych, w tym otoczenie, odległość, page rank i Eigenvector• Identyfikacja najkrótszej drogi między dwoma dowolnymi obiektami w systemie.• Łatwa wizualizacja przepływów pieniędzy i innych przedmiotów.
Inne możliwości wizualizacji i analizy	<ul style="list-style-type: none">• Wyszukiwanie dowolnych obiektów• Wykresy pozwalające przechodzić do szczegółowych danych• Tabele przestawne (uniwersalny OLAP)• Wyszukiwanie tekstowe• Sekwencje zdarzeń i obiektów• Wielowymiarowy scoring• Możliwość tworzenia własnych kolumn obliczeniowych na danych tabelarycznych• Podstawowe statystyki danych tabelarycznych (min, max, suma, średnia)• Dedykowane dossier przedstawia wszystkie dane o obiekcie na jednym
Skalowalność	<ul style="list-style-type: none">• Łatwo łączy dziesiątki / setki źródeł danych.• Horyzontalnie analiza wielu milionów / miliardów obiektów• Skalowalna architektura: skalowalność systemu na serwerach typu commodity

Źródła danych	<ul style="list-style-type: none"> • Dowolna relacyjna baza danych • Pliki Microsoft Excel • Pliki CSV • Pliki XML • Hadoop HDFS
Współpraca	<ul style="list-style-type: none"> • Pojedynczy widok danych udostępniona wszystkim użytkownikom systemu • Dokumentacja "historii" – zrzuty z ekranu z komentarzami. • Łatwe udostępnianie analiz. • Eksport wyników do szablonów Microsoft Excel. • Udostępnianie danych i analiz dla innych narzędzi analitycznych za pomocą RESTful access, JDBC i ODBC.
Bezpieczeństwo	<ul style="list-style-type: none"> • Rozdrobione uprawnienia oparte na predykatkach • Pełna ścieżka audytu
Alerty	<ul style="list-style-type: none"> • Można importować alerty z zewnętrznego systemu alertowego lub mogą być generowane za pośrednictwem DataWalk. • Alerty można wyświetlać na pulpicie nawigacyjnym Alerts Dashboard, co ułatwia ich sprawdzanie • Alerty konfigurowane przez użytkownika mogą być ustawione dla dowolnego elementu danych, w tym na własnych obliczeniach.
Raportowanie	<ul style="list-style-type: none"> • Automatyczne generowanie raportów • Eksport raportów do PDF lub pliku Microsoft Excel • Możliwość automatycznej dystrybucji zgodne ze zdefiniowaną ścieżką
Wspierane przeglądarki	<ul style="list-style-type: none"> • Chrome C38+ (rekomendowana, umożliwia najlepsze działanie platformy DataWalk) • Firefox 33+ i nowsze wersje • Microsoft IE11 • Inne przeglądarki – na żądanie klienta
Wdrożenie	<ul style="list-style-type: none"> • Oprogramowanie działające na sprzęcie typu commodity • Błyskawiczne wdrożenie wersji podstawowej oprogramowania • Wdrożenie stacjonarne lub w chmurze