

## Krok 14: Rozwój repozytorium metadanych

Tu omówiono następujące tematy:

- \* Rzeczy do rozważenia przy tworzeniu repozytorium metadanych
- \* Wiele potencjalnych źródeł metadanych i różnica między aktywnymi i pasywnymi repozytoriami metadanych
- \* Konieczność opracowania dwóch typów procesów interfejsu repozytorium metadanych: procesu interfejsu narzędzia i procesu interfejsu dostępu
- \* Cztery typy testowania, które mają zastosowanie do tworzenia repozytorium metadanych: testowanie jednostkowe, testowanie integracyjne, testowanie regresyjne i testowanie akceptacyjne
- \* Sześć czynności do wykonania, zanim repozytorium metadanych przejdzie do produkcji
- \* Jak zorganizować katalog dla repozytorium metadanych?
- \* Krótkie opisy działań związanych z tworzeniem repozytorium metadanych, rezultatów wynikających z tych działań oraz zaangażowanych ról
- \* Ryzyko niewykonania kroku 14

Rzeczy do rozważenia

Wsparcie produktu Repozytorium Metadanych

- \* Jakie interfejsy są wymagane dla produktu repozytorium metadanych?
- \* Czy będziemy musieli skontaktować się z dostawcą repozytorium metadanych w celu uzyskania pomocy przy pisaniu tych interfejsów?
- \* Czy spodziewamy się problemów z wyodrębnieniem metadanych z narzędzi, które je tworzą (np. narzędzie do oprogramowania komputerowego [CASE], narzędzie do ekstrakcji/transformacji/ładowania [ETL] lub narzędzie do przetwarzania analitycznego online [OLAP])? Czy wiemy, czy inne firmy miały jakieś problemy?
- \* Czy będziemy musieli skontaktować się z dostawcami narzędzi i poprosić ich o pomoc?
- \* Czy możemy ulepszyć raporty metadanych, które są dostarczane z produktem repozytorium metadanych?
- \* Czy produkt repozytorium metadanych ma kontekstową?

funkcja pomocy online?

Niestandardowe repozytorium metadanych

- \* Czy projekt naszego repozytorium metadanych jest wystarczająco elastyczny, aby można go było rozszerzyć w przyszłości?
- \* Czy musimy budować katalog repozytorium metadanych jako część interfejsu dostępu dla ludzi biznesu i techników?
- \* Czy musimy zbudować kontekstową funkcję pomocy online dla?

repozytorium metadanych?

\* Czy musimy pisać raporty metadanych? Jak je rozprowadzimy? Na papierze? W intranecie?

\* Czy metadane będą zintegrowane z zapytaniami i raportami BI?

aplikacja? Jaki rodzaj procesu musimy zbudować, aby to wspierać?

\* Jak możemy zapewnić, że metadane w repozytorium metadanych

nie traci synchronizacji z metadanymi w innych narzędziach i systemie zarządzania bazą danych (DBMS)?

Rekrutacja

\* Jak bardzo możemy zmodularyzować kodowanie i testowanie? Czy mamy wystarczająco dużo programistów i testerów, aby przyspieszyć prace programistyczne?

\* Czy ci sami programiści będą pracować nad funkcją pomocy online repozytorium metadanych? Czy możemy równolegle rozwijać funkcję pomocy?

\* Czy otrzymamy wsparcie od administratora bazy danych w pełnym lub niepełnym wymiarze godzin? Czy ten sam administrator bazy danych będzie obsługiwał docelowe bazy danych BI?

\* Kto będzie nadal prowadził repozytorium metadanych? Czy mamy pełnoetatowego administratora metadanych? Czy wystarczy jedna osoba?

\* Czy są jakieś specjalne potrzeby szkoleniowe, które musimy wziąć pod uwagę? Dla programistów? Dla ludzi biznesu?

Przygotowanie do produkcji

\* Czy produkcyjne repozytorium metadanych zostanie zainstalowane na dedykowanym serwerze produkcyjnym? Kto zainstaluje i przetestuje serwer?

\* Czy musimy tworzyć regularnie zaplanowane repozytorium metadanych?

programy w harmonogramie zadań? Czy personel operacyjny będzie je prowadził i monitorował?

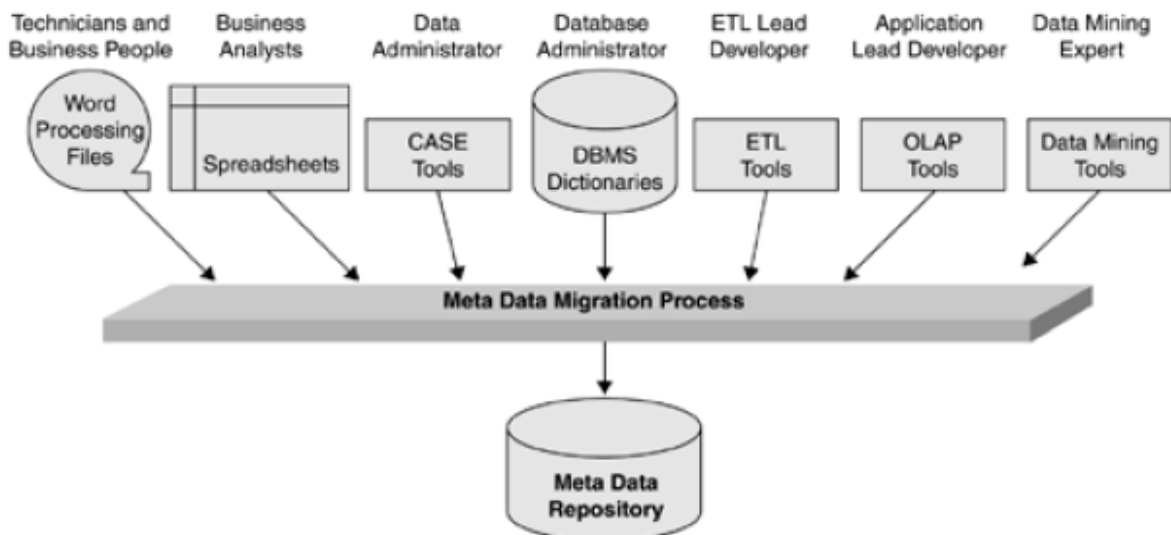
\* Czy musimy napisać procedury operacyjne dla personelu operacyjnego?

\* Czy musimy napisać poradnik dla pracowników help desku i dla ludzi biznesu?

Aby sprawniej poruszać się po środowisku wspomaganie decyzji BI, ludzie biznesu muszą mieć dostęp do repozytorium metadanych. Są tylko dwie opcje: licencja (kupno) repozytorium metadanych lub zbudowanie jednego. Po zaimplementowaniu repozytorium metadanych musi być ono utrzymywane i rozwijane w czasie. Musi być również wypełniany i aktualizowany podczas każdego cyklu procesu ETL statystykami obciążenia, sumami uzgadniania, metrykami niezawodności danych, licznikami odrzuceń danych i przyczynami odrzucenia danych.

### **Zapełnianie repozytorium metadanych**

Zapełnianie repozytorium metadanych zwykle nie jest pracą ręczną. Repozytorium metadanych otrzymuje większość swoich metadanych z wielu różnych źródeł metadanych. Te źródła metadanych są kontrolowane przez osoby inne niż administrator metadanych, jak pokazano na rysunku 14.1.



Należy opracować proces migracji metadanych, aby wyodrębnić metadane z tych źródeł, powiązać (połączyć) powiązane komponenty metadanych i zapisać w repozytorium metadanych. Poniżej krótko omówiono różne źródła metadanych.

\* Pliki edytora tekstu mogą być podręcznikami, procedurami i innymi mniej formalnymi dokumentami, które zawierają definicje danych i reguły biznesowe. Osadzone w tych regułach biznesowych mogą być zasady dotyczące danych, reguły przetwarzania, domeny danych (w podręcznikach tłumaczenia kodu) oraz różne uwagi dotyczące historii i własności danych lub procesów wykonywanych na danych. Niektóre pliki edytorów tekstu zawierają również cenną dokumentację techniczną opisującą dane i reguły procesów wymuszane przez programy.

Bądź ostrożny z plikami do edycji tekstu. Pliki te są rzadko utrzymywane, a informacje w nich zawarte mogą być nieaktualne i nie mają już zastosowania.

\* Arkusze kalkulacyjne zawierają obliczenia i makra, które są wykonywane na danych w prywatnych arkuszach kalkulacyjnych analityków biznesowych po pobraniu przez nich danych z różnych systemów operacyjnych. Te obliczenia i makra mogą być źródłem reguł transformacji, reguł czyszczenia, wprowadzeń, agregacji i podsumowań.

\* Narzędzia CASE zawierają nazwy, definicje, rozmiary, długości, relacje, informacje o kardynalności, reguły integralności referencyjnej oraz uwagi dotyczące danych, które zostały zamodelowane dla systemu operacyjnego lub aplikacji BI. Ponadto narzędzia CASE zwykle mogą przechowywać nazwy techniczne tabel i kolumn, a także klucze podstawowe i klucze obce. Niektóre z bardziej wyrafinowanych narzędzi CASE mają moduły, które zawierają metadane dla komponentów procesu, takich jak programy, ekrany i układy raportów.

\* Wewnętrzne słowniki DBMS są integralną częścią wszystkich DBMS, ponieważ słowniki kontrolują strukturę bazy danych. W relacyjnych bazach danych są one zwykle nazywane SYSTABLES i przechowują nazwy, definicje, rozmiary, długości, relacje i woluminy struktur baz danych, takich jak grupy pamięci masowej, obszary tabel, tabele, kolumny, klucze podstawowe, klucze obce i indeksy.

\* Narzędzia ETL nie działałyby bez instrukcji (metadane techniczne) dla wymaganych przekształceń. Wewnętrzne słowniki narzędzi ETL przechowują mapowanie źródło-cel oraz wszystkie algorytmy transformacji, które są stosowane do danych źródłowych podczas procesu ETL.

\* Narzędzia OLAP przechowują specyfikacje dotyczące wyprowadzania danych (obliczeń), agregacji i podsumowań w swoich wewnętrznych katalogach. Te specyfikacje umożliwiają narzędziowi OLAP wykonywanie funkcji drążenia i rozwijania. Niektóre produkty OLAP mają możliwość drążenia do innej bazy danych w tym samym DBMS lub nawet do innej bazy danych w innym DBMS w celu wyodrębnienia szczegółowych danych dla zapytania.

\* Narzędzia do eksploracji danych przechowują opisy modeli analitycznych danych, na podstawie których wykonywane są operacje eksploracji danych.

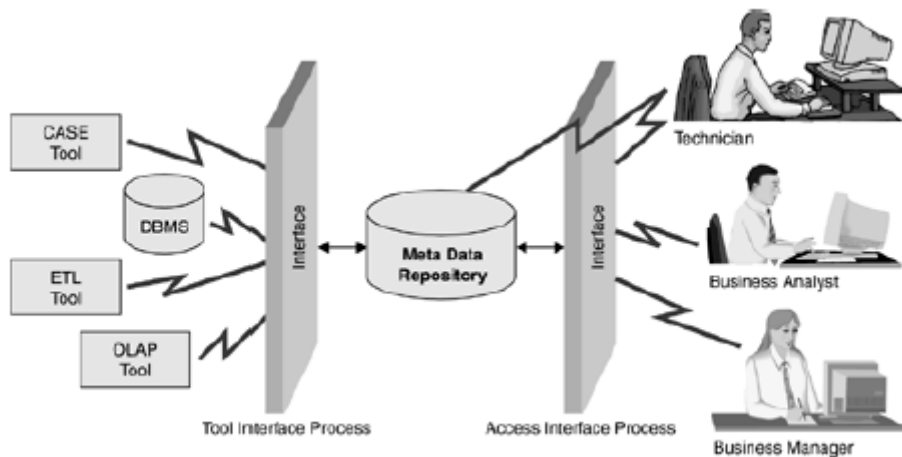
Jeśli jakiegokolwiek metadane zawarte w tych źródłach metadanych mają ulec zmianie, administrator metadanych musi zostać powiadomiony przed zmianą. Następnie będzie musiał określić, czy repozytorium metadanych może pomieścić tę zmianę, czy też repozytorium metadanych musi zostać zmodyfikowane lub ulepszone. Dlatego w celu utrzymania zdrowego i użytecznego repozytorium metadanych, administrator metadanych musi współpracować z zespołem ETL, zespołem OLAP, ekspertem od eksploracji danych, administratorem danych, administratorem bazy danych i ludźmi biznesowymi w projektach BI. Ponadto, co ważniejsze, administrator metadanych musi mieć pełną współpracę z osobami z systemów operacyjnych, które utrzymują operacyjne pliki źródłowe i źródłowe bazy danych.

Zmiany wprowadzone do operacyjnych systemów źródłowych często nie są przekazywane administratorowi metadanych na czas, aby wprowadzić niezbędne zmiany w repozytorium metadanych. Ponieważ zmiany te wpływają również na proces ETL, a w niektórych przypadkach na struktury docelowych baz danych BI, ma to również wpływ na zespół ETL. To załamanie w komunikacji między osobami zajmującymi się systemami operacyjnymi a personelem wspierającym podejmowanie decyzji BI może powodować poważne opóźnienia.

Idealnie, repozytoria metadanych powinny być aktywnymi repozytoriami, podobnymi do słowników DBMS. W aktywnym repozytorium metadanych zmiany byłyby wprowadzane tylko do repozytorium metadanych, a repozytorium metadanych propagowałoby zmiany do odpowiedniego narzędzia docelowego lub DBMS. Jednak obecnie dostępne na rynku produkty meta-repozytoriów danych są nadal repozytoriami pasywnymi. Oznacza to, że zmiany muszą być dokonywane zarówno w repozytorium metadanych, jak i w odpowiednim narzędziu docelowym lub DBMS, a zmiany te muszą być synchronizowane ręcznie lub z programami.

### **Procesy interfejsu repozytorium metadanych**

Każde repozytorium metadanych musi mieć dwa interfejsy, aby funkcjonować: interfejs narzędzia do akceptowania metadanych z innych narzędzi oraz interfejs dostępu do interakcji z ludźmi biznesu i technikami. Jeśli repozytorium metadanych jest budowane we własnym zakresie, należy opracować procesy zapewniające oba typy interfejsów (rysunek 14.2). Jeśli produkt repozytorium metadanych jest licencjonowany, interfejs dostępu zostanie dostarczony jako część produktu, ale proces interfejsu narzędzia musi być nadal rozwijany.



### Proces interfejsu narzędzia

Większość metadanych biznesowych i metadanych technicznych nie jest wprowadzana bezpośrednio do repozytorium metadanych, ale jest pobierana z narzędzi, które je przechwytyują i wykorzystują do obsługi własnej funkcjonalności. W celu wymiany tych metadanych repozytorium metadanych i narzędzia, z których pochodzą metadane, muszą komunikować się ze sobą przy użyciu wspólnych standardów metadanych. Niestety, te narzędzia nie współdziałają ze sobą zbyt dobrze poza podstawowymi możliwościami importu i eksportu, a wspólne standardy metadanych nie zostały jeszcze uzgodnione i ratyfikowane przez dostawców repozytorium metadanych. Dlatego konieczne jest opracowanie procesu interfejsu narzędzia z kilkoma programami interfejsu, aby wyodrębnić metadane z różnych narzędzi. Dostawcy repozytoriów metadanych uznają brak integracji narzędzi za problem, a standardy metadanych są dyskutowane przez dwie autorytatywne grupy: Meta Data Coalition (MDC, pod wpływem Microsoft) i Object Management Group (OMG, pod wpływem Oracle). W tej chwili nie jest jasne, które standardy metadanych zostaną przyjęte jako standardy branżowe.

### Proces interfejsu dostępu

Repozytorium metadanych musi mieć również wspólny interfejs dostępu dla osób, które potrzebują do niego dostępu. Ludzie biznesu i technicy będą chcieli pobierać informacje (metadane) o danych biznesowych w docelowych bazach danych BI, statystyki obciążenia z procesów ETL oraz inne metadane związane z procesem. Ponieważ ludzie biznesu i technicy mogą mieć różne wymagania, niektóre ekrany interfejsu dostępu i programy mogą wymagać nieco odmiennego zaprojektowania dla tych dwóch okręgów wyborczych. Technicy i „zaawansowani użytkownicy” mogą również uzyskać upoważnienie do bezpośredniego dostępu do repozytorium metadanych bez przechodzenia przez interfejs dostępu. Jeśli masz tylko jedno scentralizowane repozytorium metadanych, budowanie wspólnego procesu interfejsu dostępu wydaje się oczywistym wymogiem i stosunkowo łatwe do wykonania. Ale jeśli wdrażasz zdecentralizowane rozwiązanie repozytorium metadanych, a nawet rozwiązanie rozproszone, budowanie wspólnego procesu interfejsu dostępu do wszystkich repozytoriów metadanych lub narzędzi może nie być tak oczywistym wymogiem. Zdecydowanie o wiele trudniej jest zbudować proces wspólnego interfejsu dostępu dla zdecentralizowanego lub rozproszonego repozytorium metadanych - z pomocą lub bez pomocy Extensible Markup Language (XML). Jednak bez tego wspólnego interfejsu dostępu każdy, kto chce uzyskać dostęp do metadanych z wielu repozytoriów metadanych lub narzędzi, musiałby nauczyć się korzystać z innego interfejsu dostępu dla każdego repozytorium metadanych i narzędzia. Aby uniknąć konieczności korzystania z wielu interfejsów dostępu, ludzie biznesu mogą nalegać na powielanie wspólnych metadanych we

własnych „preferowanych” repozytoriach. Może to szybko prowadzić do nadmiarowych silosów repozytoriów metadanych, czego właśnie należy unikać.

### **Testowanie repozytorium metadanych**

Tworzenie repozytorium metadanych może być równie skomplikowane, jak tworzenie dowolnej innej aplikacji. Dlatego należy przestrzegać tych samych ustrukturyzowanych wytycznych dotyczących rozwoju, w szczególności wytycznych dotyczących testowania. Podczas gdy większość menedżerów biznesowych i menedżerów IT (IT) wymaga jakości swoich aplikacji, często nie chcą poświęcać dużo czasu na testowanie. To nie działa! Ponadto wielu kierowników projektów stale nie docenia, ile czasu zajmie testowanie, i nie przeznaczają wystarczającej ilości czasu na testowanie w planie projektu. Co gorsza, pominięcie kamieni milowych powoduje dalsze skrócenie krótkiego czasu zaplanowanego na testy. Chodzi o to, że musisz zaplanować wystarczająco dużo czasu na testowanie repozytorium metadanych i odzwierciedlić to w planie projektu. Ponadto zaplanuj dostępność wystarczającej liczby testerów w odpowiednim czasie. Na szczęście testowanie jest jedną z niewielu czynności, które można łatwo zmodularyzować, a dodanie większej liczby testerów przyspieszy prace testowe.

Doświadczenie pokazuje, że na każdy dzień kodowania należy zaplanować trzy dni testowania.

Spośród różnych typów testów omówionych w kroku 11, opracowywanie ETL (testy jednostkowe, testy integracyjne, testy regresywne, testy wydajnościowe, testy zapewniania jakości [QA] i testy akceptacyjne), tylko cztery typy testów są zwykle wymagane dla metadanych repozytorium, jak pokazano w tabeli.

Testy jednostkowe: czy kod kompiluje się bez błędów?

Testy integracyjne: czy produkty spełniają wymagania?

Testy regresji : czy zmiany w programie spowodowały uszkodzenie tego lub innego programu?

Testy akceptacyjne : Czy wszystkie aspekty elementu wynikowego są akceptowalne?

\* Testowanie integracyjne, czasami nazywane testowaniem systemowym, sprawdza, czy pierwotne cele aplikacji metadanych zostały spełnione oraz czy wszystkie funkcje repozytorium metadanych działają zgodnie z oczekiwaniami. Testy integracyjne powinny obejmować początkową populację repozytorium metadanych, okresowe lub sporadyczne jego aktualizacje, proces interfejsu dostępu, proces interfejsu narzędzia, zaplanowane raporty, standardowe zapytania i funkcję pomocy online. Ten test jest wykonywany zgodnie z planem testów (opisanym w kroku 11, ETL Development), a przypadki testowe są uruchamiane przez testerów, a nie przez twórców kodu. (Programiści mogą testować kod innych programistów, ale nie własny).

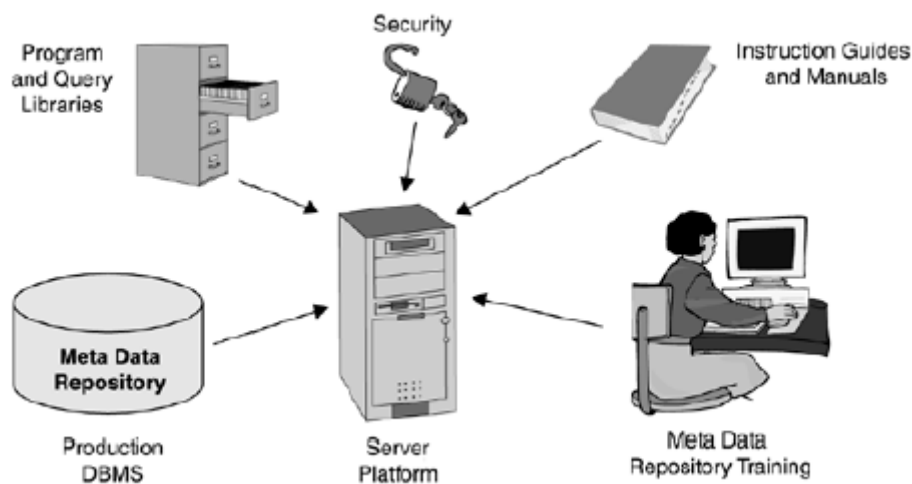
\* Testy regresji są przeprowadzane na różnych wersjach tego samego oprogramowania, które mogą być tworzone w trakcie cyklu rozwoju w ramach jednego projektu lub gdy to samo oprogramowanie jest aktualizowane w różnych kolejnych projektach. Za każdym razem, gdy zmiany są wprowadzane w dowolnym miejscu istniejących programów, należy przeprowadzić pełny cykl testu regresji, aby upewnić się, że zmiany nie naruszyły przypadkowo części logiki programistycznej w programie lub gdzieś w całym strumieniu programów. Testowanie regresyjne jest formalną czynnością testową i musi być wykonywane wraz z planem testów. Ponieważ kompletne repozytorium metadanych

prawdopodobnie nie zostanie zbudowane od razu, ale z czasem, testowanie regresji jest ważną częścią wszystkich projektów, które ulepszają repozytorium metadanych.

\* Testy akceptacyjne są nawet ważniejsze niż inne formy testowania. Ten test określa, czy repozytorium metadanych (nowe lub ulepszone) jest gotowe do wdrożenia. Ten końcowy proces testowania repozytorium metadanych zwykle łączy testy QA i testy akceptacyjne. W testowanie zaangażowani są ludzie biznesu i technicy, którzy będą używać i wspierać repozytorium metadanych. Podobnie jak w przypadku testów integracyjnych i testów regresyjnych, test ten jest wykonywany z planem testów, predefiniowanymi przypadkami testowymi i oczekiwanymi wynikami testów; wszystkie przebiegi testowe są udokumentowane w dzienniku testów.

### Przygotowanie do wdrożenia repozytorium metadanych

Przeniesienie repozytorium metadanych do środowiska produkcyjnego wymaga przygotowania, jak pokazano na rysunku. To przygotowanie powinno rozpocząć się wcześniej, a nie po zakończeniu całego kodowania i testowania. Początkujące zespoły ds. repozytoriów metadanych często nie doceniają czasu potrzebnego na przygotowanie podstawowych aspektów środowiska produkcyjnego opisanych poniżej.



\* Platforma serwerowa: repozytorium metadanych powinno znajdować się na serwerze produkcyjnym, a nie na serwerze deweloperskim. W związku z tym należy zainstalować i przetestować platformę serwera produkcyjnego, która obejmuje komponenty sprzętowe, system operacyjny, narzędzia monitorujące i łączność sieciową.

\* Produkcyjny DBMS: Jeśli repozytorium metadanych jest instalowane na nowym serwerze produkcyjnym, instancja DBMS musi zostać utworzona, a jej parametry muszą być ustawione i przetestowane w systemie operacyjnym. Jeśli licencjonowany jest produkt repozytorium metadanych, wszystkie komponenty produktu, w tym baza danych repozytorium metadanych, muszą być zainstalowane i przetestowane na serwerze produkcyjnym.

\* Biblioteki programów i zapytań: Wszystkie programy do migracji metadanych (w tym programy interfejsu narzędzi) i wszystkie programy aplikacji metadanych (w tym programy interfejsu dostępu, funkcja pomocy online, raporty i zapytania) będą rezydować w kontrolowanej wersji biblioteka. Produkt do zarządzania biblioteką musi zostać zainstalowany i przetestowany przed przeniesieniem programów repozytorium metadanych do środowiska produkcyjnego.

\* Bezpieczeństwo: Serwer produkcyjny, produkt DBMS, baza danych repozytorium metadanych i wszystkie programy muszą mieć zaimplementowane odpowiednie poziomy bezpieczeństwa. Poziomy bezpieczeństwa w środowisku produkcyjnym są znacznie bardziej rygorystyczne niż w środowisku programistycznym. Deweloperzy, którzy mogliby dowolnie zmieniać struktury bazy danych repozytorium metadanych i ich zawartość w środowisku programistycznym, nie powinni mieć takich samych uprawnień w środowisku produkcyjnym.

\* Przewodniki po instrukcjach i podręczniki: Gdy repozytorium metadanych zostanie uruchomione, niektóre programy metadanych zostaną zaplanowane do automatycznego uruchamiania w każdym cyklu procesu ETL, aby przechwycić statystyki obciążenia i metryki jakości danych oraz załadować je do repozytorium metadanych. Niektóre raporty lub zapytania mogą być również umieszczane w harmonogramie. Personel operacyjny będzie monitorować zaplanowane uruchomienia i powiadomić administratora metadanych, jeśli coś pójdzie nie tak. Należy dla nich przygotować procedury operacyjne, wymieniając zaplanowane zadania, kolejność, w jakiej zadania powinny być uruchamiane i co zrobić, jeśli zadanie się nie powiedzie. Pracownicy działu pomocy będą mentorować i wspierać ludzi biznesu w ich pytaniach dotyczących metadanych. Dlatego pomoc personelu recepcji będzie potrzebować przewodnika, do którego może się odwołać w przypadku pytań, które muszą zbadać. Ten przewodnik referencyjny można również przekazać ludziom biznesowym, ponieważ zawiera przydatne wskazówki dotyczące lokalizacji określonych składników metadanych i miejsc, w których można uzyskać dodatkową pomoc.

\* Szkolenie dotyczące repozytorium metadanych: ludzie biznesu i technicy muszą być przeszkoleni w różny sposób w zakresie korzystania z repozytorium metadanych za pośrednictwem interfejsu dostępu lub bezpośrednio i interaktywnie. Analitycy biznesowi muszą wiedzieć, jak uzyskać dostęp do repozytorium metadanych, aby pomóc im wybrać dane BI do zapytań ad hoc. Technicy muszą wiedzieć, jak używać metadanych, aby pomóc im w utrzymaniu aplikacji BI i jak pobierać metadane, aby dostarczać je jako integralną część raportów i zapytań aplikacji BI.

Najlepszą praktyką jest zapewnienie „wystarczającego” szkolenia „w samą porę”. Pierwsza sesja szkoleniowa, nie dłuższa niż jeden dzień, powinna stanowić omówienie repozytorium metadanych, jego organizacji, sposobu wydobycia niektórych komponentów metadanych oraz wprowadzenie do jednej lub dwóch podstawowych funkcji. Powiedz uczestnikom, że pierwsza sesja szkoleniowa jest tylko wstępem do repozytorium metadanych i będą potrzebować dodatkowego szkolenia. Po jednym lub dwóch tygodniach praktycznej praktyki podczas wykonywania obowiązków zawodowych uczestnicy powinni wrócić na kolejną jedno- lub dwudniową sesję szkoleniową, aby poznać bardziej zaawansowane funkcje repozytorium metadanych. Wykorzystaj sesję szkoleniową jako okazję do przedstawienia ludzi biznesu pracownikom help desku, którzy będą ich wspierać. Wspomnij ludzi biznesu, że pracownicy działu pomocy technicznej będą ich mentorować, gdy będą biegli w poruszaniu się po repozytorium metadanych i całym środowisku wspomagania decyzji BI. Zachęć ludzi biznesu do stworzenia własnej sieci. W tej sieci mogliby sobie nawzajem pomagać nie tylko w korzystaniu z repozytorium metadanych, ale także w poznawaniu danych biznesowych i funkcji aplikacji środowiska wspomagania decyzji BI jako całości.

### **Katalog repozytorium metadanych**

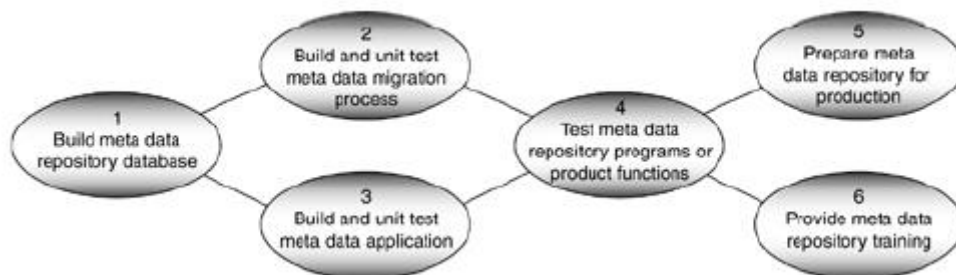
Należy zastanowić się nad najlepszym sposobem organizowania i prezentowania zawartości repozytorium metadanych ludziom biznesu i technikom. Zawartość repozytorium metadanych może być zorganizowana w katalogu, który służy jako mapa ułatwiająca nawigację po repozytorium metadanych. Na przykład zawartość ogólnego katalogu repozytorium metadanych może być podzielona na trzy główne grupy: katalog biznesowy, nawigator informacji i katalog techniczny. Te trzy



grupy mogą zawierać listę składników metadanych zawartych w każdej grupie lub mogą przedstawiać warstwę podgrup niższego poziomu. Katalog repozytorium metadanych może być zaprojektowany tak fantazyjny jak mapa witryny sieci Web lub tak prosty, jak podstawowa macierz kontekstowa. Katalog mógłby również zostać rozszerzony i włączony jako część narzędzia pomocy online środowiska wspomagania decyzji BI jako całości, aby służyć jako potencjalny punkt wejścia do repozytorium metadanych. Jeśli tworzony jest katalog repozytorium metadanych, jego wykorzystanie powinno zostać omówione podczas sesji szkoleniowych.

### Działania związane z tworzeniem repozytorium metadanych

Czynności związane z tworzeniem repozytorium metadanych nie muszą być wykonywane liniowo. Rysunek wskazuje, jakie czynności można wykonywać jednocześnie. Poniższa lista zawiera krótki opis czynności związanych z krokiem 14, tworzenie repozytorium metadanych.



1. Zbuduj bazę danych repozytorium metadanych. Jeśli budujesz repozytorium metadanych, niezależnie od tego, czy jest ono oparte na projekcie relacji encji, czy na projekcie zorientowanym obiektowo, wygeneruj dane języka definicji (DDL) i uruchom go, aby utworzyć struktury bazy danych repozytorium metadanych. Ponadto wygeneruj język kontroli danych (DCL) i uruchom go w celu ustanowienia uprawnień do tworzenia, odczytu, aktualizacji i usuwania (CRUD) w bazie danych repozytorium metadanych. Jeśli licencjonujesz produkt z repozytorium metadanych, zainstaluj i przetestuj wszystkie komponenty produktu, w szczególności bazę danych repozytorium metadanych. Skonfiguruj uprawnienia CRUD w produkcie repozytorium metadanych, aby umożliwić wykonanie procesu migracji metadanych i raportów oraz aby umożliwić bezpośredni dostęp do repozytorium metadanych.

2. Zbuduj i przetestuj proces migracji metadanych. Po utworzeniu bazy danych repozytorium metadanych należy opracować proces migracji metadanych, w tym proces interfejsu narzędzi i programy do transformacji metadanych, które przygotowują wyodrębnione metadane dla repozytorium metadanych. Jeśli licencjonowano produkt z repozytorium metadanych i jeśli funkcja importowania produktu jest używana do zapełniania repozytorium metadanych, przetestuj je, aby sprawdzić, czy działa zgodnie z oczekiwaniami.

3. Zbuduj i przetestuj aplikację metadanych. Jeśli budujesz repozytorium metadanych, musisz również opracować funkcje aplikacji metadanych, w tym proces interfejsu dostępu i funkcję pomocy online, a także raporty i zapytania dotyczące metadanych. Jeśli repozytorium metadanych jest produktem licencjonowanym, należy przetestować jego funkcje aplikacyjne (interfejsy, raporty, zapytania). Jeśli konieczne jest wzbogacenie produktu o dodatkowe funkcjonalności, napisz i przetestuj dodatkowy kod.

4. Przetestuj programy repozytorium metadanych lub funkcje produktu. Testuj wszystkie programy repozytorium metadanych lub funkcje produktu od początku do końca poprzez formalne testy integracyjne lub regresyjne. Każdy element procesu migracji metadanych, a także każdy element

aplikacji metadanych musi być intensywnie testowany. Wykonywanie testów integracyjnych lub testów regresyjnych z formalnym planem testów; uruchom przygotowane przypadki testowe, zapisz rzeczywiste wyniki testów w dzienniku testów i porównaj je z oczekiwanymi wynikami. Gdy programy repozytorium metadanych lub funkcje produktu zostaną dokładnie zintegrowane lub przetestowane regresywnie, ludzie biznesu i technicy mogą przeprowadzić kombinację testów kontroli jakości i akceptacji.

5. Przygotuj repozytorium metadanych do produkcji. Zainstaluj i przetestuj platformę serwerową dla produkcyjnego repozytorium metadanych. Utwórz DDL i DCL dla produkcyjnej bazy danych repozytorium metadanych. Napisz procedury operacyjne dla personelu operacyjnego wraz z instrukcjami dotyczącymi regularnego uruchamiania programów repozytoriów metadanych. Napisz także przewodnik dla pracowników działu pomocy i dla ludzi biznesu z instrukcjami, jak korzystać z repozytorium metadanych. Ustanów inne procedury, takie jak monitorowanie wydajności bazy danych i wykorzystanie metadanych.

6. Zapewnij szkolenie dotyczące repozytorium metadanych. Ponieważ aplikacja metadanych może być tak skomplikowana, jak każda aplikacja biznesowa, szkolenie jest ważnym aspektem. Ludzie biznesu i personel łącznikowy, tacy jak „zaawansowani użytkownicy” i pracownicy pomocy technicznej, muszą zostać przeszkoleni w zakresie korzystania z bazy danych repozytorium metadanych, funkcji pomocy online, raportów i zapytań. Opracuj i przedstaw wewnętrzne sesje szkoleniowe lub zaplanuj szkolenia za pośrednictwem dostawcy repozytorium metadanych.

### **Rezultaty wynikające z tych działań**

1. Fizyczna baza danych repozytorium metadanych. To jest fizyczna baza danych repozytorium metadanych. Jego tabele, kolumny, klucze podstawowe, klucze obce i autoryzacja CRUD są zdefiniowane w DBMS z instrukcjami Structured Query Language (SQL) zawierającymi instrukcje DDL i DCL.

2. Plan testów repozytorium metadanych. Plan testów powinien określać cel każdego testu repozytorium metadanych i przedstawiać harmonogram uruchamiania testów we wcześniej określonej kolejności. Powinien również opisywać przypadki testowe, w tym kryteria wejściowe i oczekiwane wyniki wyjściowe. Dziennik testów powinien dokumentować, kiedy testy zostały uruchomione, kto je przeprowadził i jakie były wyniki testów.

3. Programy do repozytoriów metadanych. Wszystkie programy do migracji metadanych, programy interfejsu dostępu, programy interfejsu narzędzi, programy raportów, skrypty zapytań i programy funkcji pomocy online dla repozytorium metadanych powinny być zakodowane i przetestowane. Jeśli używany jest produkt repozytorium metadanych, należy przetestować wszystkie funkcje produktu repozytorium metadanych.

4. Biblioteka programu meta repozytorium danych. Wszystkie programy i skrypty repozytorium metadanych powinny znajdować się w bibliotece programu repozytorium metadanych. Cały proces migracji metadanych, a także funkcje aplikacji metadanych, powinny zostać przetestowane pod kątem integracji lub regresji oraz kontroli jakości/akceptacji.

5. Dokumentacja produkcyjna repozytorium metadanych. Dokumentacja produkcyjna repozytorium metadanych obejmuje:

- Procedury operacyjne dla personelu operacyjnego obejmujące wszystkie zaplanowane zadania repozytorium metadanych

- Przewodnik po repozytorium metadanych dla działu pomocy i ludzi biznesu z instrukcjami korzystania z repozytorium metadanych

6. Materiały szkoleniowe do repozytorium metadanych. Materiały szkoleniowe do szkolenia w zakresie wewnętrznego repozytorium metadanych powinny zawierać slajdy prezentacji, notatki instruktora, zeszyty ćwiczeń dla uczniów, ćwiczenia i ich rozwiązania oraz wszelkie dodatkowe stosowne materiały informacyjne.

### **Rezultaty wynikające z tych działań**

1. Fizyczna baza danych repozytorium metadanych. To jest fizyczna baza danych repozytorium metadanych. Jego tabele, kolumny, klucze podstawowe, klucze obce i autoryzacja CRUD są zdefiniowane w DBMS z instrukcjami Structured Query Language (SQL) zawierającymi instrukcje DDL i DCL.

2. Plan testów repozytorium metadanych. Plan testów powinien określać cel każdego testu repozytorium metadanych i przedstawiać harmonogram uruchamiania testów we wcześniej określonej kolejności. Powinien również opisywać przypadki testowe, w tym kryteria wejściowe i oczekiwane wyniki wyjściowe. Dziennik testów powinien dokumentować, kiedy testy zostały uruchomione, kto je przeprowadził i jakie były wyniki testów.

3. Programy do repozytoriów metadanych. Wszystkie programy do migracji metadanych, programy interfejsu dostępu, programy interfejsu narzędzi, programy raportów, skrypty zapytań i programy funkcji pomocy online dla repozytorium metadanych powinny być zakodowane i przetestowane. Jeśli używany jest produkt repozytorium metadanych, należy przetestować wszystkie funkcje produktu repozytorium metadanych.

4. Biblioteka programu meta repozytorium danych. Wszystkie programy i skrypty repozytorium metadanych powinny znajdować się w bibliotece programu repozytorium metadanych. Cały proces migracji metadanych, a także funkcje aplikacji metadanych, powinny zostać przetestowane pod kątem integracji lub regresji oraz kontroli jakości/akceptacji.

5. Dokumentacja produkcyjna repozytorium metadanych. Dokumentacja produkcyjna repozytorium metadanych obejmuje:

- Procedury operacyjne dla personelu operacyjnego obejmujące wszystkie zaplanowane zadania repozytorium metadanych

- Przewodnik po repozytorium metadanych dla działu pomocy i ludzi biznesu z instrukcjami korzystania z repozytorium metadanych

6. Materiały szkoleniowe do repozytorium metadanych. Materiały szkoleniowe do szkolenia w zakresie wewnętrznego repozytorium metadanych powinny zawierać slajdy prezentacji, notatki instruktora, zeszyty ćwiczeń dla uczniów, ćwiczenia i ich rozwiązania oraz wszelkie dodatkowe stosowne materiały informacyjne.

### **Role zaangażowane w te działania**

\* Przedstawiciel firmy. Przedstawiciel biznesowy powinien uczestniczyć w testach akceptacyjnych funkcji aplikacji metadanych w taki sam sposób, w jaki uczestniczy w testach akceptacyjnych funkcji ETL oraz funkcji dostępu i analizy aplikacji BI. Przedstawiciel biznesowy musi również wziąć udział w szkoleniu dotyczącym repozytorium metadanych.

\* Administrator bazy danych. Administrator bazy danych musi utworzyć struktury bazy danych (tabele, kolumny, indeksy itd.) dla repozytorium metadanych. Przyznaje uprawnienia technikom, biznesmenom, narzędziom i programom do uzyskiwania dostępu do bazy danych repozytorium metadanych i zapisywania w niej. Pomaga również administratorowi metadanych w procesie migracji metadanych oraz wezwań dostępu do bazy danych dla funkcji aplikacji metadanych.

\* Administrator metadanych. Administrator metadanych musi zainstalować produkt repozytorium metadanych, jeśli jest on licencjonowany. Jeśli repozytorium metadanych jest budowane od podstaw, administrator metadanych musi nadzorować prace programistyczne. On lub ona musi zainstalować i przetestować serwer produkcyjny dla repozytorium metadanych. Koordynuje również działania związane z tworzeniem repozytorium metadanych z administratorem bazy danych, programistami repozytorium metadanych i testerami.

\* Twórcy repozytorium danych meta. Deweloperzy repozytorium metadanych powinni napisać kod dla procesu migracji metadanych, w tym procesu interfejsu narzędzia, oraz dla metadanych aplikacji, w tym proces interfejsu dostępu i funkcja pomocy online.

\* Testerzy. Ponieważ programiści nigdy nie powinni testować własnego kodu, niezależni testerzy (inni programiści) powinni przeprowadzać testy integracji i regresji.

#### **Ryzyko niewykonania kroku 14**

Bez repozytorium metadanych musiałbyś opracować skomplikowaną niestandardową aplikację metadanych, aby wyodrębnić metadane ze wszystkich narzędzi i DBMS w celu tworzenia raportów metadanych. Byłoby to zbyt trudne, zbyt czasochłonne, zbyt intensywne kodowanie, zbyt zawiłe i zbyt frustrujące. Inne podejścia do metadanych, takie jak rozszerzenie wykorzystania narzędzia CASE, mogą działać jako środki tymczasowe, ale nie jako rozwiązania długoterminowe. Narzędzia CASE nie są przystosowane do akceptowania metadanych z narzędzi ETL lub narzędzi OLAP, podobnie jak narzędzia ETL nie są przystosowane do akceptowania metadanych z narzędzi CASE i tak dalej. Jedynym rozwiązaniem jest repozytorium metadanych.