

X Konferencja PLOUG
Kościelisko
Październik 2004

Hurtownie Danych – Współczesność

Paweł Chomicz

Betacom S.A.

Abstrakt

Referat porusza zagadnienie ewolucji zagadnienia hurtowni danych. Jakie miejsce zajmuje obecnie w przedsiębiorstwach. Czym jest i kto jej tak naprawdę potrzebuje. Jakie wyzwania stoją obecnie przed hurtowniami danych.

Czym jest hurtownia danych i jakie są korzyści z jej wdrożenia

W obecnych czasach podstawą działania i konkurencyjności firm jest szybkie reagowanie na zewnętrzne i wewnętrzne czynniki wpływające na sferę ekonomiczną, finansową czy gospodarczą. Nie można tego osiągnąć bez wiarygodnej informacji. Jednak jak wykazuje praktyka czasami nadmiar informacji może być równie niebezpieczny jak jej brak, dlatego ważne jest umiejętne segregowanie zgromadzonych danych, by na ich podstawie można było wyciągać użyteczne wnioski.

W każdym wykorzystywanym w przedsiębiorstwie systemie informatycznym mamy możliwość tworzenia różnorodnych raportów. Każdy z tych raportów pozwala wybrać pewne dane transakcyjne według wybranego kryterium i posegregować je według pewnego klucza. Bez względu na to, jak dużo takich raportów istnieje, często zdarza się, że żaden z nich nie zawiera dokładnie tej informacji, którą potrzebujemy. Czasem by uzyskać potrzebne dane należy stworzyć kilka, bądź kilkanaście raportów, często opartych o dane pochodzące z wielu różnych systemów. Zdarza się także, że pewnych informacji w ogóle nie da się uzyskać. Nawet, jeżeli dysponujemy narzędziami wspomagającymi tworzenie raportów, sposób ułożenia danych w bazach systemu czy też systemów, może skutecznie uniemożliwić uzyskanie pożądanego zestawienia.

Innym istotnym problemem jest taka budowa wzorca, w której dane są zależne od dużego zbioru elementów ich opisujących np. czas, lokalizacja, typy. Bazy relacyjne są bardzo trudne do raportowania w tym kontekście. Jeszcze trudniejsza jest zmiana perspektywy spojrzenia na dane, ponieważ wiąże się to z utworzeniem odrębnych raportów dla każdej z nich. Jeżeli np. chcemy zbadać wartość sprzedaży w zależności od czasu lub lokalizacji, zazwyczaj wiąże się to z utworzeniem dwóch odrębnych raportów.

Kolejną trudnością jest analizowanie informacji z raportów. Po wydrukowaniu stu stron zestawienia trudno odnaleźć pozycje, które rzeczywiście nas interesują, np. ze względu na ich wysokie bądź niskie wartości. Jeśli natomiast wydrukujemy zbiorcze zestawienie, musimy generować kolejne raporty, które na większym poziomie szczegółowości ukażą nam składniki pewnej wybranej z raportu syntetycznej pozycji. Generowanie raportów pochłania wiele czasu i obciąża system komputerowy. Problem potęguje się, gdy do uzyskania jednej wartości, konieczne jest utworzenie kilku różnych zestawień. Taka metoda uzyskiwania informacji ma ponadto, inne istotne wady: czas tworzenia raportów jest długi, a używanie danych transakcyjnych jako źródła informacji nie jest efektywne, gdyż nie są one usystematyzowane i "oczyszczone", czyli pozbawione błędów.

Następna kwestia to zarządzanie danymi historycznymi. Systemy transakcyjne zazwyczaj są zaprojektowane do wydajnego przetwarzania niewielkich porcji danych takich jak: wystawienie faktury czy obsługa zamówienia a nie do generowania zbiorczych analiz, poza tym przechowują tylko aktualne informacje o większości obiektów. Dane historyczne przeważnie przechowuje się na oddzielnych nośnikach i nie udostępnia na bieżąco, co utrudnia wykonywanie raportów porównujących aktualne wskaźniki z wartościami z lat ubiegłych. Zdarza się również, że nastąpiła zmiana systemu w przedsiębiorstwie i odczytanie archiwalnych danych jest wręcz niemożliwe.

Odpowiedzią na dzisiejsze potrzeby zarządzania jest technologia wielowymiarowej struktury danych, której celem jest wspomaganie zarządzania poprzez dostarczanie właściwych informacji, właściwym ludziom, we właściwym czasie przy niskim koszcie.

Technologia ta jest niezależna od platformy sprzętowej, systemu operacyjnego czy bazy danych. Informacje potrzebne do analiz, które mogą pochodzić z wielu różnych systemów informatycznych wykorzystywanych w przedsiębiorstwie są gromadzone w jednym miejscu, dzięki czemu łatwe jest ich utrzymywanie. Zysk z wdrożenia tego typu systemu polega z jednej strony na redukcji kosztu związanych dotychczasowym raportowaniem oraz oszczędności i eliminacja zagrożeń związanych z pełniejszym i szybszym dostępem do informacji. W grupie kosztów możemy wyróżnić: koszty związane z wdrożeniem i koszty eksploatacji. Te pierwsze zwykle są duże, ale jednorazowe, i można je znacznie ograniczyć tworząc system adekwatny do potrzeb przedsiębiorstwa.

Jeżeli chodzi o koszty eksploatacji to są one wyraźnie niższe ze względu na pełną automatyzację procesu pobierania, czyszczenia i ładowania danych.

Korzystając z tej technologii można budować aplikacje, dzięki którym możliwa będzie analiza wszystkich obszarów działalności przedsiębiorstwa - analiza finansów, sprzedaży, klientów, marketingu, dystrybucji, itd. Pozwoli to nie tylko skutecznie mierzyć biznes, ale również skutecznie go planować.

Najlepszym rozwiązaniem tych problemów jest utworzenie korporacyjnej hurtowni danych, w której gromadzone dane pochodzą z wszystkich systemów transakcyjnych firmy. Jest to jednak proces bardzo złożony, przez co wzrasta ryzyko niepowodzenia przedsięwzięcia. Dodatkowy minus stanowi fakt, że jest to rozwiązanie bardzo drogie i czasochłonne. Alternatywą jest utworzenie systemu sprawozdawczego opartego o wielowymiarowe struktury danych tzw. Data Marts, poświęcone wybranym zagadnieniom tematycznym przedsiębiorstwa. Najważniejszymi cechami takiego systemu jest: intuicyjność utworzonego modelu, proste tworzenie raportów i zapytań „ad-hoc”. Dodatkową zaletą takiego rozwiązania jest krótki czas wdrożenia i szybkie rezultaty, dzięki czemu użytkownicy końcowi uzyskują dostęp do systemu umożliwiającego raportowanie i analizę wybranych obszarów biznesowych.

Obecnie dynamika rozwoju firm jest tak duża, że klasyczne metody pozyskiwania informacji stają się zbyt wolne i nieoptymalne. Pomimo ogromu danych przechowywanych w różnych systemach informatycznych poziom zadowolenia z dostarczanych informacji nie jest zbyt wysoki. Systemy te nie posiadają skutecznych mechanizmów pozwalających na dokonywanie wielowymiarowych analiz prowadzonego biznesu, a tym samym nie są w stanie efektywnie dostarczyć informacji istotnie wspomagających zarządzanie.

Natomiast dzięki wielowymiarowej strukturze mamy możliwość wykonywania dowolnych analiz i raportów będących źródłem wyczerpujących informacji, które pozwalają na ocenę obserwowanych zjawisk i stanowią podstawę do podejmowania odpowiednich i szybkich decyzji biznesowych. Dane zostają zamienione w wiedzę, która pozwala na lepsze zrozumienie otoczenia, w którym funkcjonuje przedsiębiorstwo, obniżyć koszty i ryzyko działalności przy jednoczesnym zwiększeniu zysku, a przede wszystkim zbudować i utrzymać długotrwałą przewagę konkurencyjną.

Koncepcja wielowymiarowej struktury danych pozwala na osiągnięcie niżej wymienione korzyści:

- Otrzymywanie na czas aktualnej informacji,
- Możliwość tworzenia przekrojowych analiz danych,
- Szybsze dostrzeżenie pojawiających się problemów i możliwości ich rozwiązania,
- Możliwość "zagłębiania się" w dane, drażnienia ich od ogólnego spojrzenia na firmę jako całość do szczegółowej analizy wybranych, interesujących składników,
- Możliwość prowadzenia prognoz – analiz, co by było gdyby, co będzie za miesiąc, rok,
- Użytkownicy mają okazję wykazania się w twórczej pracy nad analizą danych, a nie tylko nad ich ewidencją,
- Dane zebrane w systemach transakcyjnych stają się dostępne do sporządzania dowolnych raportów,
- Zwiększenie czytelności danych, gdyż o ich umiejscowieniu decyduje tematyka a nie np. pochodzenie czy przeznaczenie,
- Integracja danych, dbanie o ich spójność i zgodność z przyjętymi standardami. Napływające z różnych źródeł dane są, w razie potrzeby przekształcane poprzez ujednoczenie formatów, jednostek miar, itp.
- Zasilanie danymi odbywa się automatycznie, co zmniejsza ryzyko błędów w danych.

Tworzenie raportów porównujących wartości bieżące ze wskaźnikami z poprzednich okresów dzięki przechowywaniu danych historycznych,

Najlepszym wydaje się rozwiązanie polegające na utworzeniu warstwy pośredniej w postaci struktury wielowymiarowej pomiędzy systemami transakcyjnymi a warstwą raportową. Nie spowoduje to konieczności rezygnacji z dotychczasowych raportów a jedynie poszerzy obecną funkcjonalność systemu przy jednoczesnej poprawie jakości systemu raportowego. Nowy system pozwoli na eksplorację informacji i generowanie analiz na dowolnym poziomie agregacji danych. Możliwość dogłębnej analizy i wielowymiarowej interpretacji (w dowolnym kontekście) danych dają przewagę w podejmowaniu szybszych decyzji, dostrzeganiu korelacji i czynników niewidocznych przy tradycyjnych metodach analiz takich jak statyczne raporty i arkusze kalkulacyjne.

Dodatковым produktem będzie pełna dokumentacja obecnych źródeł danych, przepływów danych i raportów wraz z oceną jakości danych.

Następna faza rozwoju oraz propozycje innowacji

Hurtownie danych stały się architektoniczną odpowiedzią na chaos różnorodnych systemów i niezgodnych danych. Głównym celem było uzyskanie integracji i integralności na polu całych przedsiębiorstw. Naturalny rozwój hurtowni danych przebiegał w różnych kierunkach. Większość z nich radykalnie oddaliła się od idei integracji. Wiele z nich skończyło się sukcesem, niektóre przebiegały burzliwie.

Data Mart, Operational Data Store, Departmental Warehouse, Shared Data Network, Corporate Information Factory i wiele innych terminów było przedstawiane jako różne warianty rozwiązań z zakresu hurtowni danych. Spory pomiędzy konsultantami a środowiskiem akademickim doprowadziły do zawirowań w naturalnym kierunku rozwoju.

Pomimo niektórych błędnych perspektyw zasadność głównej koncepcji jest niezaprzeczalna: funkcjonalny zbiór zgodnych danych i rozpowszechnianie ich jest niezbędne dla przetrwania przedsiębiorstwa. Bez względu na rozmiar przedsiębiorstwa oraz niezliczoną ilość źródeł generowania danych, integracja z partnerami i kontrola łańcucha wartości (value chain) jest obowiązkową praktyką biznesu. Zbieranie i analiza danych marketingowych w celu poznania klienta na potrzeby lepszej obsługi nie jest już wyłącznie sloganem marketingowym. Jest cechą charakterystyczną sukcesu.

Wybuchowy wzrost na rynku pakietów analitycznych jest znakiem osiągnięcia przez hurtownie danych pewnej dojrzałości, lecz nie w sensie cyklu życia, ale przejścia do innego spojrzenia na osiągnięcie krytycznych celów przedsiębiorstwa.

Wyzwania przed współczesnymi hurtowniami danych

Możliwość wzrostu zwrotu z istniejących inwestycji

Należy zastanowić się nad tym, czego nauczyliśmy się podczas pracy z hurtowniami. Spójrzmy nie tylko na to, co zbudowaliśmy, lecz również jak dobrze to działa. Pierwszym dobrym krokiem jest dostrojenie środowiska. Naszą naturalną tendencją jest skupianie się na elementach, które mogą być ulepszone poprzez administrację bazą danych bądź sieci. W wielu przypadkach kwestia osiągnięć tkwi jednak w konstrukcji i wywodzi się z ustalenia procesów decyzyjnych, które wymagają żmudnych analiz. Należy zrozumieć gruntownie jak są wykorzystywane zasoby hurtowni. Często zespół hurtowni danych zarządza danymi a inna grupa wspiera narzędzia analityczne. Nie raz decyzje o wykorzystywaniu danych są podejmowane wyłącznie wewnątrz grup biznesowych. Wymagane jest ogólne spojrzenie na zbiór danych, aby zoptymalizować projekt i procesy. Jest wiele rozwiązań tego problemu jednym z możliwych jest coś, co nazywamy „warehouse inventory management process”. Jest to aktywne podejście bazujące na kreatywnym monitorowaniu, analizie wzorów zachowań i adaptacji reguł poprawiających operacje. Jest to forma zarządzania jakością stosowana przy hurtowniach danych.

Rozwój istniejącego środowiska

Hurtownia danych może być jedynie wtedy dobrze prosperująca, kiedy się rozrasta. Wzrost mierzony jest w trzech kierunkach: zawartości, kategorii obsługiwanych użytkowników oraz możliwych analiz.

Wzrost może być uzyskiwany poprzez:

- Ponowne użycie rozwiązań, które są zbliżone do innych zapotrzebowań. Jeśli obsługujemy np. księgowość jednego działu to księgowość drugiego działu jest potencjalnym kandydatem.
- Pionowy rozwój jest drugą możliwością, jeśli analizy były skierowane np. na badanie zadowolenia klientów wtedy rozwinięcie analiz o planowanie sprzedaży może być kolejnym krokiem.
- Dodanie danych zewnętrznych. Niejednokrotnie potrzebne są dane nie tylko z obszaru firmy np. marketing potrzebuje danych z produkcji i sprzedaży, ale również danych pochodzących z innych źródeł, aby móc prognozować wyniki.

Dokonywanie cyklicznych przeglądów architektury

Zmiany technologiczne są wyznacznikiem przeglądów a cykl przyrostowy wprowadzi stabilność w złożonym środowisku. Często po wdrożeniu hurtowni pojawiają się szybko zapotrzebowania na kolejne Data Mart'y w innych działach. Pamiętać należy przy tym o elastyczności hurtowni. Często się zdarza, że między kolejnymi Data Mart'ami zachodzi redundancja danych oraz ich niespójność, co prowadzi do uniemożliwienia odpowiedzi na postawione pytania biznesowe przedsiębiorstwa, nie mówiąc już o ogromnym koszcie utrzymywania tych oddzielnych systemów.

Integracja pomiędzy hurtownią danych a CRM

Dobrze prosperujący CRM powinien bazować na analizach ze zintegrowanego zbioru danych bieżących i historycznych pochodzących z całego przedsiębiorstwa. W związku z tym można powiedzieć, że jeżeli przedsiębiorstwo nie posiada hurtowni danych nie posiada również systemu CRM.

Czynne składowanie danych

Obecnie nie jest wystarczające by hurtownia danych odpowiadała na pytania „Co się wydarzyło?”. By móc skutecznie konkurować z innymi przedsiębiorstwami użytkownicy muszą mieć możliwość skutecznego przewidywania tego jak ich działania wpłyną na rozwój biznesu. W związku z tym oczekują od hurtowni, że na podstawie zachowań z przeszłości będą mogli przewidzieć przyszłość. Obszarami tego typu analiz mogą być:

- a) Skuteczne metody pozyskiwania klientów,
- b) Przewidywanie odchodzenia klientów,
- c) Przewidywanie zjawisk windykacyjnych,
- d) Dobór oferty „lojalnościowej” dla klientów,
- e) Optymalizacja kampanii marketingowych.

Hurtownie danych czasu rzeczywistego

Tradycyjne hurtownie danych nie zawierają danych ładowanych na bieżąco. Są najczęściej ładowane danymi z systemów operacyjnych, co tydzień lub niekiedy, co noc, ale są w każdym przypadku spojrzeniem na przeszłość. Dzisiejszy biznes zaczyna coraz bardziej potrzebować aktualnych danych. Tak samo jak decyzje w biznesie stają się coraz bardziej decyzjami czasu rzeczywistego tak i systemy, które wspierają te decyzje muszą spełniać ich wymagania. Jest rzeczą oczywistą, że DW, BI, DS oraz narzędzia OLAP szybko zaczną zawierać dane czasu rzeczywistego. Wyzwania stojące przed takim podejściem to:

f) Uzyskanie procesów ETL czasu rzeczywistego.

Jedną z najtrudniejszych rzeczy przy budowaniu hurtowni jest proces ETL. Wykonywanie procesu ETL w czasie rzeczywistym przysparza dodatkowych wyzwań. Większość tych procesów pracuje aktualnie w trybie wsadowym. Przy ładowaniu na bieżąco nie może być żadnych przerw w udostępnianiu hurtowni danych użytkownikom.

g) Modelowanie tabeli faktów czasu rzeczywistego.

Zaimplementowanie podejścia czasu rzeczywistego w istniejącej hurtowni lub zaprojektowanie nowej opartej na tych zasadach porusza wiele ciekawych kwestii. Hurtownia danych posiadająca zagregowane dane na wielu poziomach opartych na wymiarach czasu musi umieć rozpoznać, że zagregowana informacja może być niesynchronizowana z danymi czasu rzeczywistego.

h) Zapytania OLAP a zmieniające się dane

OLAP i narzędzia raportujące zostały zaprojektowane do pracy na statycznych danych historycznych.

Nowa generacja hurtowni danych

Zagadnienie hurtowni danych znane jest już od początku lat 90-tych. Stanowiło ono sztandarowe hasło postępu technologicznego tamtego okresu. Często bywało, że celem wdrożenia hurtowni danych było nie tyle usprawnianie biznesu, ile poprawianie marketingowego wizerunku firmy wobec klientów i firm konkurencyjnych. Niezaprzeczalne jest jednak, że przez ten czas cały czas rozwijała się wiedza o hurtowniach poprzez praktykę kolejnych wdrożeń, zarówno tych zakończonych sukcesem jak i porażką. Użytkownicy stali się bardziej świadomi swoich oczekiwań i bardziej wymagający. Obecnie można powiedzieć, że wkraczamy w kolejne pokolenie hurtowni danych. Powinny się one stać krytycznym komponentem dla przedsiębiorstwa.

Chociaż można przeczytać dużo definicji hurtowni danych, które mówią, że te systemy są zaprojektowane na potrzeby osób podejmujących strategiczne decyzje mało jest publikacji o faktycznym wykorzystaniu hurtowni przy podejmowaniu takich decyzji. Zdefiniujmy pojęcie podejmowanie decyzji strategicznych. Prawdopodobnie są tysiące opublikowanych definicji. Dla celu tego referatu można powiedzieć, że strategiczne decyzje są jednymi z tych, które wymagają dużych nakładów finansowych niejednokrotnie wraz z przydzieleniem dużych zasobów ludzkich i bywa to przyczyną wielkich problemów lub sukcesów do czasu, kiedy efekty podjętej decyzji są przesłaniane przez kolejną.

Większość przypadków wykorzystywania hurtowni nie jest związana z podejmowaniem decyzji strategicznych. Prawdopodobnie najważniejszym powodem jest to, że strategiczne decyzje nie są podejmowane tak często. Większość hurtowni jest wykorzystywana przede wszystkim do monitorowania efektów podjętej decyzji. Nie mniej jednak niektóre hurtownie są wykorzystywane przy podejmowaniu decyzji i przynoszą wiele korzyści.

Jak wobec tego można wykorzystać hurtownię przy podejmowaniu strategicznych decyzji?

Przygotowywanie specjalnych danych, modelowanie i raportowanie są najbardziej czasochłonnymi zajęciami podczas wykorzystywania hurtowni przy podejmowaniu decyzji.

Rozwiązania na potrzeby podjęcia decyzji mają stosunkowo krótki cykl życia. Niekiedy czas wykorzystywania tych systemów liczony jest w dniach. Te kilka dni wykorzystywania systemów mogą jednak przynieść większy zysk niż niektóre systemy raportowe przyniosły przez całe lata.

Takie rozwiązania muszą być wykonane zazwyczaj w terminie od jednego popołudnia do kilku tygodni. Wnioski muszą być wyciągnięte przez analityka biznesowego, który niejednokrotnie nie ma czasu na formalne wywiady czy testy. Wymagania są otrzymywane na spotkaniach biznesowych, które mogą być trochę problematyczne lub relacjonowane z drugiej ręki przez uczestników tych spotkań. Te wymagania są niejednokrotnie niejasne i trzeba wtedy wywnioskować, jaka jest faktyczna potrzeba biznesowa.

Trzeba będzie prawdopodobnie dokonać specjalnej agregacji danych, wykorzystać inne kalkulacje lub połączyć dane, które nie były nigdy łączone. Wykonana praca pozwoli w inny sposób spojrzeć na problem niż przez ogół biznesu. Innymi słowy częścią wielu efektywnych decyzji stra-

tegicznych jest przedstawienie problemu biznesowego w innej perspektywie. Potrzeba wykonania tej pracy jest spowodowana tym, że tworzona hurtownia danych była projektowana na podstawie spojrzenia na ówczesny biznes.

Często trzeba uruchomić wielokrotne zapytania do utworzenia nowej części hurtowni danych. Taka nowa część może być utworzona raz poprzez pojedyncze zapytanie ze złożonymi warunkami lub może być potrzeba wielokrotnego wykorzystania nowych agregatów i kalkulacji z jednoczesnym wykorzystaniem danych spoza produkcyjnej bazy danych lub z hurtowni jednak niełatwych do połączenia. Ze względu na prostotę i efektywność najlepszym wyjściem jest utworzenie specjalnej bazy danych. Niejednokrotnie można myśleć, że nie po to się zbudowało hurtownie żeby teraz tworzyć takie wyjątki jednak takie wyjątki są nie do uniknięcia.

Wiele prac przy podejmowaniu decyzji za pomocą hurtowni kończy się dostarczeniem danych do arkuszy kalkulacyjnych użytkowników. Może być to realizowane poprzez linki do hurtowni danych lub załadowanie danych bezpośrednio do arkuszy. Arkusze są wykorzystywane, ponieważ użytkownicy potrzebują zmieniać złożone kalkulacje, zazwyczaj, dlatego że stale jest wątpliwość jak pewne są kalkulacje a użytkownicy są najbardziej zorientowani do dokonywania zmian w własnie w arkuszach. Większość z tych kalkulacji są wielowymiarowymi analizami.

Nieraz trzeba wykonać specjalne raporty. Informacja z hurtowni danych musi być dostarczona ludziom, którzy nie potrzebują bezpośredniego dostępu do hurtowni. Niejednokrotnie na potrzeby podjęcia strategicznej decyzji niektórzy użytkownicy będą potrzebować wydrukowanych raportów, aby kogoś przekonać. Raporty te będą wymagać dopracowanego wyglądu, aby były wiarygodne, niejednokrotnie z wykorzystaniem różnych wykresów.

Dzisiaj największym zagrożeniem jest brak informacji lub nawet gorzej niekompletna informacja. Decyzje podejmowane są bez uwzględnienia wszystkich faktów, natomiast ryzyko oraz okazje oceniane są na podstawie anegdotycznej, niekompletnej lub przestarzałej informacji. Zgadywanie nie jest z pewnością najlepszą praktyką w biznesie. W związku z tym narasta szum wokół terminu Business Intelligence (BI). Jednak długo przed tym jak termin business intelligence został sprecyzowany, było wiele odmiennych i niezależnych procesów biznesowych zaprojektowanych do zbierania, monitorowania, analizy i raportowania informacji w przedsiębiorstwie. Różnica dzisiaj polega na dojrzałości, integracji, automatyzacji oraz skalowalności technologii rozwiniętej aby wesprzeć business intelligence umożliwiając przepływ w krótkim czasie wiernych, bieżących oraz pewnych informacji z całego przedsiębiorstwa. W artykule przedstawione zostaną dzisiejsze korzyści biznesowe, jakie przynosi najnowsza technologia.

Inżynieria informacyjna

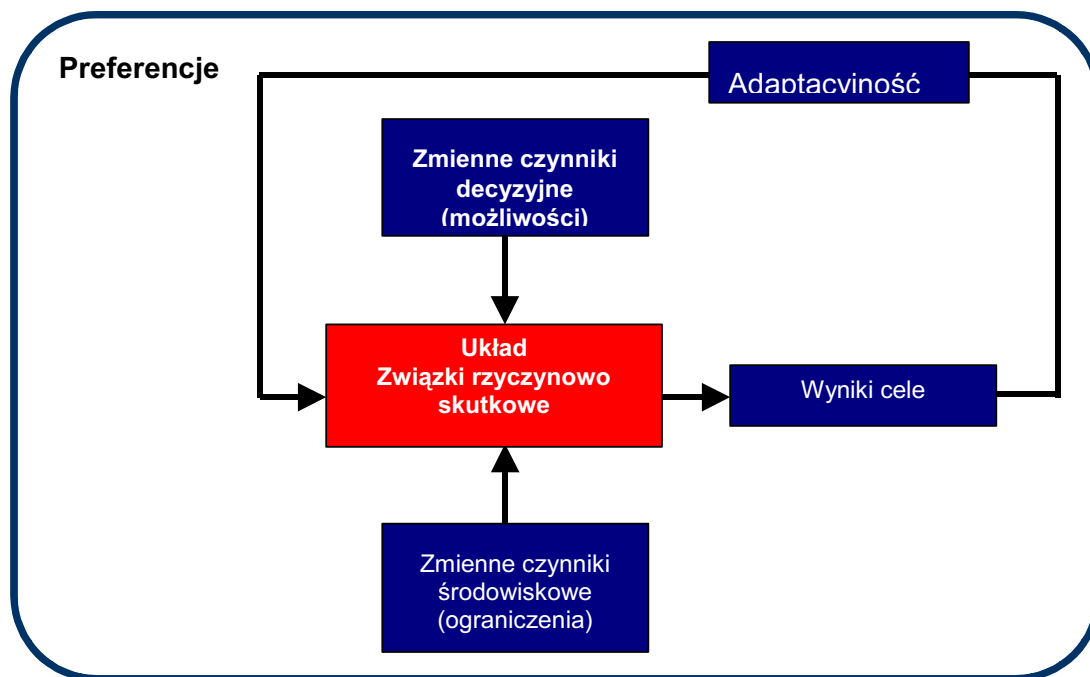
Inżynieria informacyjna jest nauką mającą za zadanie identyfikację potrzeb informacyjnych oraz rozwój systemu informatycznego, który udostępnia informacje odbiorcom. Inżynieria informacyjna jest również odpowiedzialna za proces filtrowania który redukuje masy zbędnych danych do komunikatu który zawiera wymaganą informację. Prostem celem inżynierii informacyjnej jest dostarczenie właściwych danych do potrzebujących je ludzi w każde wymagane miejsce na czas w czytelnej formie i po ustalonym koszcie, tak żeby mogły te osoby podjąć właściwe decyzje oraz wykonać optymalne akcje.

Podejmowanie decyzji

Chociaż można przeczytać dużo definicji hurtowni danych, które mówią że są to systemy zaprojektowane na potrzeby osób podejmujących strategiczne decyzje, mało jest publikacji o faktycznym wykorzystaniu hurtowni przy podejmowaniu takich decyzji. Zdefiniujmy pojęcie podejmowanie decyzji. Prawdopodobnie są tysiące opublikowanych definicji. Dla celu tego referatu można powiedzieć, że podejmowanie decyzji związane jest z procesem wyboru odpowiedniego kursu działania w danej sytuacji. Wymaga to dokonania wyboru najlepszej opcji z wielu możliwych rozwiązań. Istnieje wiele rodzajów problemów decyzyjnych jednostki w przedsiębiorstwie, przed którą stoi zadanie dokonania pojedynczego wyboru.

Kluczową kwestią przy podejmowaniu decyzji jest stopień pewności możliwy do przyjęcia co do systemu, w którym odbywa się działanie. Podejmowanie decyzji w warunkach niepewności oznacza, że prawdopodobieństwo zaistnienia różnych opcji oraz wyników nie jest znane. Jeśli prawdopodobieństwo jest określone na podstawie doświadczeń z przeszłości lub obserwacji, nazywamy tę sytuację podejmowaniem decyzji w warunkach ryzyka.

W procesie podejmowania decyzji w warunkach ryzyka główny problem stanowi określenie kompromisu pomiędzy sprzecznymi celami przy założonym poziomie ryzyka. Podejmowanie decyzji w warunkach niepewności wymaga natomiast określenia możliwych kompromisów.



Podczas podejmowania decyzji, przy założeniu określonych preferencji w stosunku do wyników, mamy do czynienia z układem w którym występują określone:

- zmienne czynniki środowiskowe (ograniczenia) - wszystkie płaszczyzny przestrzeni, które znajdują się poza wpływem działających lub podejmujących decyzje,
- zmienne czynniki decyzyjne (możliwości) - wszelkie płaszczyzny przestrzeni znajdujące się pod ich kontrolą, w zasięgu ich wpływu,
- przekonania dotyczące związków przyczynowo – skutkowych powstałe w procesie adaptacyjnego uczenia się.

Wyniki zależą zarówno od wpływu zmiennych czynników środowiskowych, jak i czynników decyzyjnych. Aby osiągnąć dany cel lub pożądany stan, należy określić ograniczenia i w jaki sposób się nimi zająć. Dlatego też, poza określeniem stanu teraźniejszego i stanu pożądanego, skuteczne podejmowanie decyzji wymaga określenia ograniczeń, w ramach których jednostka działa lub będzie działać. Decyzje podejmowane przez ludzi oraz środki, jakie wybiorą to często wypadkowa zakładanych ograniczeń. Ograniczenia te występują w procesie planowania jako czynniki, które mogą, ale nie muszą się zmieniać, jednak nie podlegają arbitralnej kontroli jednostki lub organizacji.

Różne strategie podejmowania decyzji są odpowiednie po wzięciu pod uwagę stanu tych zmiennych. Preferencje co do wyników (celów) np. mogą być wyraźne lub niewyraźne. Podobnie

działania przyczynowo – skutkowe wpływające na wynik mogą być pewne lub niepewne. Różne typy sytuacji, w których podejmuje się decyzje wymagają zastosowania odmiennych strategii:

- w przypadku, gdy istnieje pewność zarówno co do przyczyn i preferencji wyników, należy zastosować strategię obliczeniową,
- w przypadku, gdy preferencje wyników są wyraźne, ale przyczyny są niepewne, należy posłużyć się strategią rozsądkową uwzględniającą nieprzewidziane wypadki,
- w przypadku, gdy istnieje pewność co do przyczyn ale brak jasności w odniesieniu do preferencji wyników, sytuacja wymaga strategii kompromisowej,
- w przypadku, gdy istnieje niepewność zarówno co do przyczyn jak i preferencji wyników, sytuacja wymaga skupienia się na wyższym poziomie celów.

Rodzaje użytkowników końcowych

Tylko kilka usług hurtowni danych jest dostarczanych na poziom użytkownika końcowego ale są one prawdopodobnie najbardziej ważnymi usługami w hurtowni danych. Usługi te są udostępniane poprzez narzędzia, które umożliwiają użytkownikom dostęp do danych zgromadzonych w hurtowni. Większość ocen doświadczeń użytkowników z hurtownią danych będzie determinowana, poprzez to jak dobrze te narzędzia spełnią ich potrzeby. Dla użytkownika końcowego reszta hurtowni jest tylko instalacją. Użytkownicy hurtowni chcą tylko pracować i jest niedobrze jeśli tak nie jest. Ludzie z IT często zapominają o tym, że użytkownicy są znacząco zróżnicowani pod względem ich zdolności technologicznych. Często są zaskoczeni jak trudne dla wielu ludzi w sferze biznesowej jest zrozumienie tego co uważają za prostą technologię. Hurtownia danych powinna wspierać pełen zakres technicznych umiejętności i złożoności analitycznych użytkowników. Tabela przedstawia, gdzie są usytuowane poszczególne grupy zdolności technicznych i jakie narzędzia są stosowane aby wesprzeć ich potrzeby. Profil ten zmienia się cały czas jak tylko komputery stają się bardziej powszechne oraz dostępne są coraz lepsze publikacje. Jednak są to powolne zmiany.

Typy użytkowników				
Przestrzeń Wykorzystania	Paper User	Push-Button	Ad Hoc	Power User
Generalnie Komputer	żadne	e-mail, edytory tekstu	edytory tekstu, arkusze kalkulacyjne, prezentacje	makra, funkcje, publikacje WWW
Hurtownia Danych	polegają na innych przy nawigacji	standardowe raporty, parametry domyślne, EIS	pisanie prostych zapytań, modyfikacja istniejących zapytań, przeglądanie/zmiana parametrów, nawigacja w hierarchiach	budowanie skomplikowanych zapytań, bezpośredni dostęp do bazy danych

Zapotrzebowanie na informacje

Kolejna tabela przedstawia cztery główne kategorie potrzeb informacyjnych i kilka korepondujących atrybutów dla każdego zapotrzebowania. Jak do tej pory poznaliśmy role

użytkowników w każdej kategorii, częstym przypadkiem jest że ta sama osoba ma potrzeby w kilku kategoriach. Chociaż są przypadki że dana osoba gra jedną rolę przez większość czasu, role zmieniają się szczególnie w dzisiejszym środowisku biznesowym.

Kategoria zapotrzebowania na informacje	Rola użytkownika	Kategoria dostępu do danych	Narzędzia	Rozmiar grupy odbiorców
Monitorowanie na wysokim poziomie – główne metryki, alarmowanie	Zarządzanie wysokiego szczebla	Push-Button – „dash board”	Interfejsy w stylu EIS, niektóre narzędzia tworzenia zapytań	Mała
Kontrolowanie biznesu – rynek, produkty, klienci itd.; drill down	Zarządzanie średniego szczebla z zakresu: sprzedaży, marketingu, obsługi klienta itd.	Standardowe raporty - parametryzowane	Narzędzia raportowe, OLAP, zapytania	Duża
Poszukiwanie – wyjątki, nowe problemy lub okazje, projekty do okazjonalnych potrzeb biznesowych	To samo co powyżej razem z analitykami biznesowymi	Analizy ad hoc	Narzędzia OLAP, zapytania, narzędzia do zaawansowanej analizy	Średnia
Złożone analizy – zapytania zagłębione, analiza statystyczna, modelowanie biznesowe	Analitycy biznesowi oraz eksperci od analiz	Zaawansowane analizy	Narzędzia do zaawansowanej analizy, narzędzia statystyczne, Data Mining, algorytmy specjalizowane	Mała

Warto zwrócić uwagę że liczba odbiorców nie jest bezpośrednio skorelowana z ich potencjalnym wpływem na biznes lub ich wagą dla hurtowni. Można pominąć wiele frustracji jeśli w procesie projektowania zostanie wzięte pod uwagę kto jest głównym klientem hurtowni danych, jakie są jego potrzeby informacyjne, które hurtownia powinna spełnić jednocześnie znając możliwości spełnienia tych potrzeb.

Korzyści Hurtowni Danych

Odpowiedzią na dzisiejsze potrzeby zarządzania jest technologia wielowymiarowej struktury danych, której celem jest wspomaganie zarządzania poprzez dostarczanie informacji zgodnie z zasadami inżynierii informacyjnej. Dzięki tej technologii informacje potrzebne do analiz, które mogą pochodzić z wielu różnych systemów informatycznych wykorzystywanych w przedsiębiorstwie są gromadzone w jednym miejscu, dzięki czemu łatwe jest ich utrzymywanie oraz wykorzystanie.

Zysk z wdrożenia tego typu systemu polega z jednej strony na redukcji kosztu związanych dotychczasowym raportowaniem oraz oszczędności i eliminacji zagrożeń związanych z pełniejszym i szybszym dostępem do informacji. W grupie kosztów możemy wyróżnić: koszty związane z wdrożeniem i koszty eksploatacji. Te pierwsze zwykle są duże, ale jednorazowe i można je znacznie ograniczać tworząc system adekwatny do potrzeb przedsiębiorstwa. Jeżeli chodzi o koszty eksploatacji to są one wyraźnie niższe ze względu na pełną automatyzację procesu pobierania, czyszczenia i ładowania danych.

Następna kwestia to zarządzanie danymi historycznymi. Systemy transakcyjne zazwyczaj są zaprojektowane do wydajnego przetwarzania niewielkich porcji danych takich jak: wystawienie faktury czy obsługa zamówienia a nie do generowania zbiorczych analiz, poza tym przechowują tylko aktualne informacje o większości obiektów. Dane historyczne przeważnie przechowuje się na oddzielnych nośnikach i nie udostępniane na bieżąco, co utrudnia wykonywanie raportów porównujących aktualne wskaźniki z wartościami z lat ubiegłych. Zdarza się również, że nastąpiła zmiana systemu w przedsiębiorstwie i odczytanie archiwalnych danych jest wręcz niemożliwe.

Korzystając z tej technologii można budować aplikacje, dzięki którym możliwa będzie analiza we wspólnym systemie miar biznesowych wszystkich obszarów działalności przedsiębiorstwa: analiza finansów, sprzedaży, klientów, marketingu, dystrybucji itd. Pozwoli to nie tylko skutecznie mierzyć biznes, ale również skutecznie go planować. Hurtownia danych dzięki wspólnej architekturze udostępnia wspólny kontekst, w którym wszystkie dane mogą być zdefiniowane pod względem ich znaczenia. Metadane są projektowane, zarządzane, integrowane, składowane stając się ogólnie dostępnymi, dla każdego kto chce wykorzystać zintegrowane informacje.

Korzyści standardowych raportów

W każdym wykorzystywanym w przedsiębiorstwie systemie informatycznym mamy możliwość tworzenia różnorodnych raportów. Każdy z tych raportów pozwala wybrać pewne dane transakcyjne według wybranego kryterium i posegregować je według pewnego klucza. Bez względu na to, jak dużo takich raportów istnieje, często zdarza się, że żaden z nich nie zawiera dokładnie tej informacji, którą potrzebujemy. Czasem by uzyskać potrzebne dane należy stworzyć kilka, bądź kilkanaście raportów, często opartych o dane pochodzące z wielu różnych systemów. Zdarza się także, że pewnych informacji w ogóle nie da się uzyskać. Nawet, jeżeli dysponujemy narzędziami wspomagającymi tworzenie raportów, sposób ułożenia danych w bazach systemu czy też systemów, może skutecznie uniemożliwić uzyskanie pożądanego zestawienia.

Innym istotnym problemem jest taka budowa wzorca, w której dane są zależne od dużego zbioru elementów ich opisujących np. czas, lokalizacja, typy. Bazy relacyjne są bardzo trudne do raportowania w tym kontekście. Jeszcze trudniejsza jest zmiana perspektywy spojrzenia na dane, ponieważ wiąże się to z utworzeniem odrębnych raportów dla każdej z nich. Jeżeli np. chcemy zbadać wartość sprzedaży w zależności od czasu lub lokalizacji, zazwyczaj wiąże się to z utworzeniem dwóch odrębnych raportów. Kolejną trudnością jest analizowanie informacji z raportów. Po wydrukowaniu stu stron zestawienia trudno odnaleźć pozycje, które rzeczywiście nas interesują, np. ze względu na ich wysokie bądź niskie wartości. Jeśli natomiast wydrukujemy zbiorcze zestawienie, musimy generować kolejne raporty, które na większym poziomie szczegółowości ukażą nam składniki pewnej wybranej z raportu syntetycznej pozycji. Generowanie raportów pochłania wiele czasu i obciąża system komputerowy. Problem potęguje się, gdy do uzyskania jednej wartości, konieczne jest utworzenie kilku różnych zestawień. Taka metoda uzyskiwania informacji ma ponadto, inne istotne wady: czas tworzenia raportów jest długi, a używanie danych transakcyjnych jako źródła informacji nie jest efektywne, gdyż nie są one usystematyzowane "oczyszczone", czyli pozbawione błędów.

SYSTEMY RAPORTOWE OPARTE O BUSINESS INTELLIGENCE SPRAWNIE ROZWIĄZUJĄ WYŻEJ WYMIENIONE PROBLEMY DOSTARCZAJĄC WYMAGANYCH RAPORTÓW POPRZEZ PORTALE KORPORACYJNE, PRZEGLĄDARKI INTERNE-

TOWE LUB W FORMIE WYDRUKOWANEJ ORAZ PRZESYŁAJĄC NA KONTA POCZTOWE LUB DO SERWERA PLIKÓW.

Korzyści analiz oraz zapytań Ad Hoc

Obecnie dynamika rozwoju firm jest tak duża, że klasyczne metody pozyskiwania informacji stają się zbyt wolne i nieoptymalne. Pomimo ogromu danych przechowywanych w różnych systemach informatycznych poziom zadowolenia z dostarczanych informacji nie jest zbyt wysoki. Systemy te nie posiadają skutecznych mechanizmów pozwalających na dokonywanie wielowymiarowych analiz prowadzonego biznesu, a tym samym nie są w stanie efektywnie dostarczyć informacji istotnie wspomagających zarządzanie. Natomiast dzięki wielowymiarowej strukturze mamy możliwość wykonywania dowolnych analiz i raportów będących źródłem wyczerpujących informacji, które pozwalają na ocenę obserwowanych zjawisk i stanowią podstawę do podejmowania odpowiednich i szybkich decyzji biznesowych. Dane zostają zamienione w wiedzę, która pozwala na lepsze zrozumienie otoczenia, w którym funkcjonuje przedsiębiorstwo, obniżyć koszty i ryzyko działalności przy jednoczesnym zwiększeniu zysku a przede wszystkim zbudować i utrzymać długotrwałą przewagę konkurencyjną.

Koncepcja wielowymiarowej struktury danych pozwala na osiągnięcie niżej wymienionych korzyści:

- otrzymywanie na czas aktualnej informacji,
- możliwość tworzenia przekrojowych analiz danych,
- szybsze dostrzeżenie pojawiających się problemów i możliwości ich rozwiązania,
- możliwość „zagłębiania się” w dane, drażnienia ich od ogólnego spojrzenia na firmę jako całość do szczegółowej analizy wybranych, interesujących składników,
- możliwość prowadzenia prognoz – analiz, co by było gdyby, co będzie za miesiąc, rok,
- użytkownicy mają okazję wykazania się w twórczej pracy nad analizą danych a nie tylko nad ich ewidencją,
- dane zebrane w systemach transakcyjnych stają się dostępne do sporządzania dowolnych raportów,
- zwiększenie czytelności danych, gdyż o ich umiejscowieniu decyduje tematyka a nie np. pochodzenie czy przeznaczenie,
- integracja danych, dbanie o ich spójność i zgodność z przyjętymi standardami (napływające z różnych źródeł dane są w razie potrzeby przekształcane poprzez ujednoczenie formatów, jednostek miar, itp.),
- zasilanie danymi odbywa się automatycznie, co zmniejsza ryzyko błędów w danych,
- tworzenie raportów porównujących wartości bieżące ze wskaźnikami z poprzednich okresów dzięki przechowywaniu danych historycznych.

Najlepszym wydaje się rozwiązanie polegające na utworzeniu warstwy pośredniej w postaci struktury wielowymiarowej pomiędzy systemami transakcyjnymi a warstwą raportową. Nie spowoduje to konieczności rezygnacji z dotychczasowych raportów a jedynie poszerzy obecną funkcjonalność systemu przy jednoczesnej poprawie jakości systemu raportowego. Nowy system pozwoli na eksplorację informacji i generowanie analiz na dowolnym poziomie agregacji danych. Możliwość dogłębnej analizy i wielowymiarowej interpretacji danych daje korzyści w podejmowaniu szybszych decyzji, dostrzeganiu korelacji i czynników niewidocznych przy tradycyjnych metodach analiz takich jak statyczne raporty i arkusze kalkulacyjne.

Korzyści analiz statystycznych i specjalizowanych algorytmów

Strategiczne decyzje są jednymi z tych, które wymagają dużych nakładów finansowych niejednokrotnie wraz z przydzieleniem dużych zasobów ludzkich i bywa to przyczyną wielkich problemów lub sukcesów do czasu, kiedy efekty podjętej decyzji są przesłaniane przez kolejną strategiczną decyzję. Większość przypadków wykorzystywania hurtowni nie jest związana z podejmowaniem decyzji strategicznych. Prawdopodobnie najważniejszym powodem jest to, że strategiczne decyzje nie są podejmowane tak często. Większość hurtowni jest wykorzystywana przede wszystkim do monitorowania efektów podjętej decyzji. Nie mniej jednak niektóre hurtownie są wykorzystywane przy podejmowaniu decyzji i przynoszą wiele korzyści. Jak wobec tego można wykorzystać hurtownię przy podejmowaniu strategicznych decyzji? Przygotowywanie specjalnych danych, modelowanie i raportowanie są najbardziej czasochłonnymi zajęciami podczas takiego procesu. Rozwiązania na takie potrzeby mają stosunkowo krótki cykl życia. Niekiedy czas wykorzystywania tych systemów liczony jest w dniach. Te kilka dni wykorzystywania systemów może jednak przynieść większy zysk niż niektóre systemy raportowe przyniosły przez całe lata.

Takie rozwiązania muszą być wykonane zazwyczaj w terminie od jednego popołudnia do kilku tygodni. Wnioski muszą być wyciągnięte przez analityka biznesowego, który niejednokrotnie nie ma czasu na formalne wywiady czy testy. Wymagania są otrzymywane na spotkaniach biznesowych, które mogą być trochę problematyczne lub relacjonowane z drugiej ręki przez uczestników tych spotkań. Te wymagania są niejednokrotnie niejasne i trzeba wtedy wywnioskować, jaka jest faktyczna potrzeba biznesowa.

Potrzeba będzie prawdopodobnie dokonać specjalnej agregacji danych, wykorzystać inne kalkulacje lub połączyć dane, które nie były nigdy łączone. Wykonana praca pozwoli w inny sposób spojrzeć na problem niż przez ogół biznesu. Innymi słowy częścią wielu efektywnych decyzji strategicznych jest przedstawienie problemu biznesowego w innej perspektywie. Potrzeba wykonania tej pracy jest spowodowana tym, że tworzona hurtownia danych była projektowana na podstawie spojrzenia na ówczesny biznes. Często trzeba uruchomić wielokrotne zapytania do utworzenia nowej części hurtowni danych. Taka nowa część może być utworzona raz poprzez pojedyncze zapytanie ze złożonymi warunkami lub może być potrzeba wielokrotnego wykorzystania nowych agregatów i kalkulacji z jednoczesnym wykorzystaniem danych spoza produkcyjnej bazy danych lub z hurtowni najczęściej niełatwych do połączenia. Ze względu na prostotę i efektywność najlepszym wyjściem jest utworzenie specjalnej bazy danych. Niejednokrotnie można myśleć, że nie po to się zbudowało hurtownie żeby teraz tworzyć takie wyjątki jednak takie wyjątki są nie do uniknięcia.

Wiele prac przy podejmowaniu decyzji za pomocą hurtowni kończy się dostarczeniem danych do arkuszy kalkulacyjnych użytkowników. Może być to realizowane poprzez linki do hurtowni danych lub załadowanie danych bezpośrednio do arkuszy. Arkusze są wykorzystywane, ponieważ użytkownicy potrzebują zmieniać złożone kalkulacje, zazwyczaj, dlatego że stale jest wątpliwość jak pewne są kalkulacje a użytkownicy są najbardziej zorientowani do dokonywania zmian we własnych arkuszach. Większość z tych kalkulacji jest wielowymiarowymi analizami.

Nieraz trzeba wykonać specjalne raporty. Informacja z hurtowni danych musi być dostarczona ludziom, którzy nie potrzebują bezpośredniego dostępu do hurtowni. Niejednokrotnie na potrzeby podjęcia strategicznej decyzji niektórzy użytkownicy będą potrzebować wydrukowanych raportów aby kogoś przekonać. Raporty te będą wymagać dopracowanego wyglądu aby były wiarygodne, niejednokrotnie z wykorzystaniem różnych wykresów.

Potrzebne są algorytmy, które wesprą klientów w doskonaleniu efektywności i jakości procesów biznesowych. Bardzo ważnym elementem jest adaptacyjność, dzięki niej system jest w stanie określić użyteczność decyzji i w czasie rzeczywistym wybierać decyzje optymalne w danych warunkach. Dzięki temu przedsiębiorstwo może szybko przystosować się do zaobserwowanych trendów zmian i podjąć właściwe decyzje na podstawie niepewnych i niekompletnych informacji. Większość dotychczasowych metod jest niewystarczająca wobec złożoności współczesnych procesów biznesowych. Ograniczają się one do rozważania prostych zależności przyczynowo-

skutkowych i prostych modeli probabilistycznych, nieadekwatnych do natury modelowanych zjawisk. Przedsiębiorstwa nie są dłużej zainteresowane "tym co było" - interesuje je "to co będzie". Wiedza dotycząca przeszłości jest niewystarczająca do określania celów marketingowych czy też polityki cenowej. Niezbędne jest wsparcie przedsiębiorstwa poprzez wyposażenie w narzędzia wspomagające prognozowanie zmian. Do tej pory używa się do analizy takich zjawisk metod matematycznych oraz statystycznych, rzadko natomiast narzędzi obliczeniowych. W rezultacie opracowane dotąd techniki obliczeniowe były oparte na założeniach, które unikały nie analizowalnych matematycznie, złożonych zjawisk zastępując je nadmiernie uproszczonym modelem zmiennych niezależnych. Taki model był często nieadekwatny, lecz poddawał się rozwiązaniom analitycznym.

Podsumowanie – wnioski

Nie można oczekiwać, że użytkownicy mają świadomość siły technologii. Do czasu, kiedy nie będzie użytkowników z dobrym rozeznanieniem w technologii do zadań analityka biznesowego będzie należeć inspirowanie wyobraźni użytkowników.

Użytkownicy będą zarówno niedoceniać jak i przeceniać możliwości hurtowni danych w procesie podejmowania decyzji. Znaczy to, że można nieraz stracić okazje lub być postawionym przed niemożliwymi zadaniami. Pomimo ogólnie stosowanej polityki uświadamiania użytkowników, wcześniejsza budowa zaufania podejmujących decyzje dużo pomaga.

Podczas projektowania hurtowni nie należy projektować jej pod względem każdej ewentualnej potrzeby decyzyjnej. Nie jesteśmy zdolni przewidzieć wszystkich potrzeb przy zadaniach na potrzeby podjęcia decyzji. Nie należy łądować do hurtowni danych wszystkiego, o czym tylko pomyslimy. Trzeba jak najdokładniej określić główne wymiary wykorzystywane w analizie biznesu.

Nie można pozwolić, żeby wiedza na temat systemu pozostawała jedynie w umysłach zewnętrznych konsultantów. Niejednokrotnie konsultanci są niedostępni, kiedy nadarza się sposobność wykorzystania ich wiedzy. Należy poznać arkusze kalkulacyjne użytkowników a także jak hurtownia może z nimi współpracować. Często zapomina się przy hurtowniach, że arkusze kalkulacyjne są jednym z najczęściej wykorzystywanych narzędzi przy podejmowaniu decyzji. Osoby obsługujące hurtownie powinny poznać języki skryptowe wykorzystywane w arkuszach, żeby mieć możliwość zaoferowania na czas rozwiązań podczas procesu decyzyjnego.

Nie należy zakładać, że hurtownia sama z siebie poprawi wyniki decyzji strategicznych. Jeśli osoba jest mierna w podejmowaniu decyzji sama technologia nie uczyni go lepszym w tych działaniach.

Ciężko jest skalkulować ROI z projektu hurtowni danych. W większości przypadków biznes musi uwierzyć, że wysiłek będzie tego wart. Z tego powodu sukces podczas podjęcia decyzji strategicznej z pomocą hurtowni danych może znacznie podnieść wrazenie, że hurtownia danych była warta włożonej w nią pracy.

Bibliografia

1. Devlin, B.A. and Murphy, P.T. An Architecture for a Business and Information System. IBM Systems Journal. Volume 27, No. 1, 1988.
2. W.H. Inmon, Building the Data Warehouse, QED/Wiley, 1991.
3. Ralph Kimball, The Data Warehouse Toolkit : Practical Techniques for Building Dimensional Data Warehouses, John Wiley & Sons, 1996.
4. Michael Haisten, cykl artykułów opublikowanych na łamach magazynu DM Review.
5. W.H. Inmon, „Building the Data Warehouse”, QED/Wiley, 1991
6. Ralph Kimball, „The Data Warehouse Toolkit :Practical Techniques for Building Dimensional Data Warehouses”, John Wiley & Sons, 1996
7. M.H. Bracket, "The Data Warehouse Challenge"

8. W.H. Inmon, R.D. Hackharthon, "Using the Data Warehouse"