



Praktyczne aspekty implementacji narzędzia klasy Business Intelligence

Wojciech Kosiba

Centralny Ośrodek Informatyki Górnictwa S.A.
40-065 Katowice
ul. Mikołowska 100
607-468-348
Wojciech.Kosiba@coig.katowice.pl

Streszczenie: W referacie omówiono możliwości funkcjonalne narzędzia klasy Business Intelligence na przykładzie rozwiązań wykonanych przez COIG S.A. dla klientów z branży wydobywania węgla kamiennego, w oparciu o technologię Business Objects i przy zastosowaniu specjalizowanej dla celów hurtowni danych bazy danych Sybase IQ. Przyjęto krokowy model wdrożenia, wprowadzając ograniczenia zarówno w sferze tematycznej jak i funkcjonalnej. Pierwszym etapem wdrożenia rozwiązania stał się obszar tematyczny Zbytu Węgla w zakresie informacji dostępnych na fakturze. Już na tym etapie stało się możliwe tworzenie takich zestawień, jakich wcześniej nie przewidywali autorzy założeń hurtowni danych i jej warstwy prezentacyjnej, stało się zatem możliwe obserwowanie niedostępnych dotąd przekrojów wielowymiarowej bazy danych. Kolejnymi obszarami dla których zbudowano Tematyczne hurtownie danych (ang. Data Mart, to Logistyka Materiałowa, Koszty oraz Kadry. Referat powstał w oparciu o doświadczenia zdobyte w ciągu wdrożeń systemów eksploatowanych dla celów biznesowych, stąd ściśle praktyczne spojrzenie na proces zaprezentowane przez autora referatu.

1. Wprowadzenie

Efektywne funkcjonowanie każdej firmy wymaga wsparcia decyzji kadry zarządzającej odpowiednio przygotowaną informacją na temat aktualnych parametrów ekonomicznych, lub takich których konsekwencje mogą rzutować na wyniki ekonomiczne. Nowatorskim aspektem rozwiązań przy użyciu narzędzi klasy Business Intelligence jest to, że do rąk użytkownika-analityka merytorycznego, trafia narzędzie umożliwiające eksploatację baz danych, dotychczas dostępną jedynie dla wyspecjalizowanych programistów bazodanowych [1]. Dotychczas osoba zajmująca się zagadnieniami analityczno-merytorycznymi zmuszona była teoretycznie sama wymyślić treść zestawienia, opracowania czy wykresu. Osoba ta nie mając dostępu do bazy danych, mogła obserwować jedynie dostępne jej preselekcjonowane informacje utworzone w formie wydruków, grafiki czy nawet w internetowej lub elektronicznej wersji sztywnych raportów. Obecnie po wdrożeniu rozwiązań klasy Business Intelligence osoba ta dostaje narzędzie o funkcjonalności języka zapytań SQL do bazy danych [2]. Moim zdaniem jest to rewolucyjna zmiana w podejściu do danych. Już nie programista, czy projektant systemu tworzy zestawienia, które mogłyby być użyteczne dla analityka merytorycznego, tylko sam analityk uzyskuje dostęp do wielowymiarowej kostki danych, którą może przecinać dowolnymi dwuwymiarowymi płaszczyznami lub nawet trójwymiarowymi przestrzeniami. Wielowymiarowość niegdyś postrzegana jako obiekt ściśle teoretycznych rozważań, uzyskuje realny, a nawet bardzo prosty kształt. Kolejną konsekwencją wdrożenia hurtowni danych wraz z jej warstwą prezentacyjną jest to, że użytkownik-analityk merytoryczny uzyskuje możliwość porównania danych z różnych, dotychczas rozdzielnych źródeł. W skrócie mówiąc, można powiedzieć, że dotychczas osoba, do której obowiązków należało analizowanie informacji o parametrach firmy miała funkcjonalność typu pakietu Excel na danych takich, jakie udało się jej zdobyć. Dzisiaj osoba ta otrzymuje język zapytań SQL dostępu do baz danych, nie tracąc nic z możliwości modelowania danych do postaci tabeli, macierzy czy wykresów różnych typów.

Dosłownie tłumacząc angielski termin „Business Intelligence” otrzymujemy „wywiad gospodarczy”, czyli jak najobszerniejszą informację o przedsiębiorstwie biznesowym i jego otoczeniu. Istotą hurtowni danych jest możliwość wprowadzenia wszelkich posiadanych informacji i zorganizowania ich w sposób adekwatny do możliwości zastosowanej bazy danych. Z drugiej zaś strony, inteligencja to umiejętność kojarzenia faktów pozornie odległych. Dobrze zorganizowana hurtownia danych daje możliwość kojarzenia wymiarów dotychczas niekojarzonych. Przykładowo, bardzo proste jest stworzenie zestawienia w wymiarze geograficznym sprzedaży

węgla według kaloryczności i sortymentu. Dzisiaj takie zestawienie można stworzyć i natychmiast z niego zrezygnować. Można też wyciągnąć wnioski i zbudować inne zestawienie, w którym wykorzystane zostaną uzyskane wcześniej informacje. Przedtem uzyskanie takiego zestawienia wymagało zatrudnienia programisty znającego narzędzie SQL, co było kosztowne, a jego potrzeba wątpliwa.

Budując hurtownię danych należy brać pod uwagę możliwość wprowadzenia w przyszłości do niej takich informacji, które dzisiaj w ogóle nie są zbierane. Mogą to być np. informacje o trendach sprzedaży u konkurencji, uzyskane zupełnie poza systemami informatycznymi eksploatowanymi w firmie, np. kupione w profesjonalnej wywiadowni gospodarczej.

Wdrożone aplikacje wymagają ciągłej modernizacji pod kątem zarówno informacji, które uzyskiwane są na wyjściu z systemu, jak i architektura hurtowni danych musi uwzględniać zmieniającą się strukturę systemów źródłowych.

Ważnym i obecnie niedocenianym elementem spośród czynności, które należy wykonać wokół wdrożenia hurtowni danych i rozwiązań Business Intelligence jest, jak to już wcześniej podkreślono, wsparcie analityka merytorycznego w zakresie eksploracji danych. Bezpośrednio po wdrożeniu analityk merytoryczny z jednej strony zaczyna zauważać aspekty dotychczas niewidoczne, czyli takie zestawienia, których istnienia nie podejrzewał, z drugiej zaś strony styka się z danymi, do których wcześniej dostęp mieli jedynie programiści bazodanowi. Bardzo istotne jest, aby na tym etapie panować nad wymiarami, które zostaną zaimplementowane w Świecie Obiektów, gdyż może się zdarzyć, że dane miały dotychczas charakter „techniczny”. Udział autorów systemu źródłowego w budowie ETL może skrócić czas implementacji nawet o 75%.

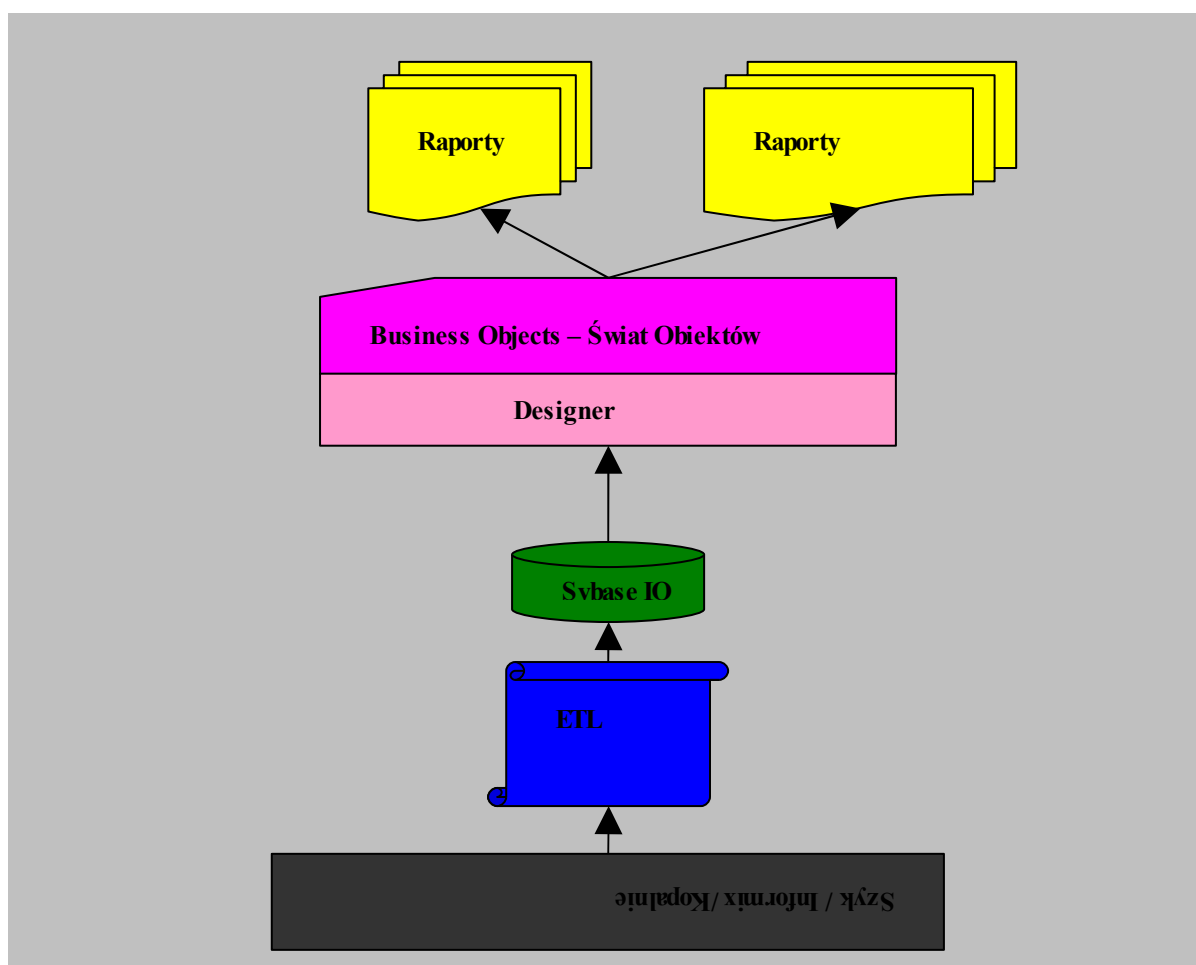
System Business Intelligence może być tak skonfigurowany, aby można było o dowolnej porze uzyskać operacyjną informację na temat najważniejszych, interesujących nas parametrów ekonomicznych. Taki dostęp zwany „kokpitem menadżerskim” umożliwia obserwowanie dynamicznych zmian i relacji pomiędzy nimi za pomocą np. graficznych emulatorów wskaźników analogowych.

2. System źródłowy SZYK/INFORMIX

Istotnym czynnikiem efektywnego wdrożenia hurtowni danych jest kontakt z twórcami systemu źródłowego. W przypadku wdrożenia hurtowni danych u klientów COIG S.A. aspekt ten nie był krytyczny, ponieważ autorem systemu źródłowego Zintegrowanego Systemu Wspomagającego Zarządzanie Przedsiębiorstwem – System SZYK, jest również COIG S. A. [3, 4]. Programowanie hurtowni danych wykonywane było w specjalistycznej pracowni, co powodowało konieczność ścisłych kontaktów również z innymi pracowniami systemu SZYK zajmującymi się rozwiązaniami dla poszczególnych obszarów tematycznych. Pracami objęto: obszar Zbytu Węgla, obszar Logistyki Materiałowej, obszar Rozliczenia Kosztów, obszar Ewidencji Kadrowej oraz obszar Centralnych Rozrachunków. Istota współpracy z autorami systemu źródłowego polega na wyselekcjonowaniu takich danych z Baz Tematycznych Systemu Szyk, które mają wartość merytoryczną. W toku kilkuletniego budowania baz danych dla klientów, niektóre pola mogły zmienić swoją funkcjonalność, inne zaś przestały być wykorzystywane i zawierają nieistotne z merytorycznego punktu widzenia informacje lub w skrajnych przypadkach wartości przypadkowe.

Można sobie wyobrazić wdrożenie narzędzi klasy Business Intelligence prowadzone na nieautoryzowanych danych, wówczas szacuje się wydłużenie procesu wstępnego projektu o 100%. Wykorzystywana przez system SZYK relacyjna baza danych Informix jest bardzo praktycznym narzędziem informatycznym wspierającym aplikacje transakcyjne. Jednak próba wykorzystania narzędzia Informix bezpośrednio jako warstwy bazodanowej hurtowni danych nie powiodła się. Komputer typu IBM F80 z systemem operacyjnym AIX częściowo obciążony pracą na rzecz aplikacji transakcyjnych ulegał permanentnym przeciążeniom. Dlatego też podjęto decyzję o uruchomieniu specjalizowanej bazy danych Sybase IQ, która jest zorganizowana w ten sposób, że dane przechowywane są jako poindeksowane kolumny. W celu przejścia do efektywniejszego środowiska bazodanowego, odrzucenia informacji zbędnych i szkodliwych, zapewnienia właściwej kontroli danych, sformatowania danych do potrzeb nowego środowiska - zastosowano technikę zbierania danych ETL.

3. Technika zbierania danych - warstwa ETL



Rys. 1 Schemat ideowy rozwiązania

Zastosowana w rozwiązaniach wdrożonych u klientów COIG S.A. technika zbierania danych – warstwa ETL (ang. extract, transformation, load) oznacza ekstrakcję z systemu źródłowego oraz transformację danych ze względu na potrzeby, o których wspomniano pod koniec poprzedniego punktu referatu oraz ładowanie danych do systemu docelowego. Każdy z obszarów tematycznych hurtowni danych (ang. Data Mart) charakteryzował się innymi potrzebami. I tak np. w obszarze Zbytu Węgla w rozwiązaniach przygotowanych na bazie Informix wykorzystywano zbyt wiele tablic, co w początkowym okresie nastęcało olbrzymich trudności w wychwyceniu i zaimplementowaniu powiązań pomiędzy danymi w ten sposób, aby na etapie tworzenia raportów, czyli dwuwymiarowych przekrojów bazy danych otrzymywać dane spójne logicznie oraz zgodne formalnie. Oczywiście jest, że narzędzie Business Objects jako generator zapytań w języku SQL może generować bardzo długie zapytania do bazy danych, jednak zapytania te zawsze muszą być zbudowane w sposób logiczny i formalnie zgodny nie tylko ze składnią języka SQL, co jest wbudowane w system, ale zgodny z logiką danych opisanych przez programistę podczas tworzenia Świata Obiektów[5],[6]. W niektórych sytuacjach wygodniej i efektywniej z punktu widzenia architektury narzędzia Business Intelligence jest przygotować dane w taki sposób, aby relacje opisane w bazie danych były maksymalnie uproszczone. W obszarze tematycznym Zbytu Węgla zdecydowano się na utworzenie w warstwie ETL jednej tablicy zawierającej wszystkie informacje z faktury. Ta jedna tablica umożliwia nam tworzenie przekrojów sprzedaży węgla zarówno np. w wymiarze ilości i ceny węgla jak i w wymiarze parametrów jakościowych, czy chemicznych. Pozostałe tablice związane z wymiarem czasu, wymiarem klienta pozostawiliśmy poza tablicą danych podstawowych.

4. Specjalistyczna baza danych – Sybase IQ

Zastosowana do zbudowania hurtowni danych baza danych Sybase IQ charakteryzuje się pionowym składowaniem tablic oraz tym, że tablice są logiczne, a nie fizyczne. Daje to efekt 90% redukcji operacji wejścia wyjścia, przy czym przetwarzane są tylko kolumny niezbędne do zapytania. Drugą bardzo charakterystyczną cechą bazy danych Sybase IQ jest to, że dane tworzą indeks. Jest to opatentowana metoda redukcji danych do bitów „Bit Wise”. Występują indeksy bitmapowe, indeksy do dat, czasu, słów kluczowych oraz indeksy do złączeń. W wyniku zastosowania tych metod nigdy nie występuje skanowanie tablic, a zapytania „ad hoc” działają tak jakby baza była strojona do każdego z nich. Zazwyczaj sumaryczny rozmiar bazy danych jest mniejszy niż czyste dane źródłowe.

Indeks bitmapowy zakładany jest przez użytkownika, gdy liczba unikalnych wartości w kolumnach zawiera się pomiędzy 10 a 500. Indeks „Bit-Wise” zakładany jest przez użytkownika, gdy ilość wystąpień przekracza 1000. Nie należy wówczas używać poleceń typu „distinct”, W przypadku gdy chcielibyśmy używać „distinct” należy zastosować indeks „Bit-Wise”+B-drzewo.

Mówiąc językiem praktyki zastosowanie bazy danych Sybase IQ ułatwiło wdrożenie hurtowni danych oraz przyspieszyło proces rozwoju hurtowni wraz z narzędziem klasy Business Intelligence.

5. Opis danych – Designer

Narzędziem umożliwiającym tworzenie Świata Obiektów dla Business Objects jest oprogramowanie pozwalające opisać dane, nadać im łatwe do zrozumienia dla analityka merytorycznego nazwy oraz opisać relacje pomiędzy danymi – narzędzie Designer. Narzędzie to jest bardzo intuicyjne, a jego używanie dla doświadczonego programisty języka zapytań SQL jest wręcz oczywiste. Graficzny interfejs tego narzędzia pozwala posługiwać się metodą „przeciagnij i upuść”.

6. Warstwa – Świat Obiektów

Utworzony za pomocą narzędzia Designer Świat Obiektów jest najważniejszym elementem projektu hurtowni danych. W skrócie mówiąc warstwa Świat Obiektów w prosty sposób realizuje wizję danych w oczach merytorycznego specjalisty. O ile w bazie danych nazwy niektórych pól są bardzo skomplikowane i nazwane w jakiś mnemotechniczny sposób na użytek programistów dane te otrzymują nazwy spójne i zrozumiałe. Na przykład, jeśli pole „f100” w bazie danych oznacza tak naprawdę numer faktury, to w warstwie Świat Obiektów wskazane jest aby wymiar ten nazywał się po prostu „numer faktury”. Należy podkreślić, że implementacja narzędzia Business Objects nie wymaga żadnych zmian w strukturze wykorzystywanych baz danych, wszystkie transformacje i nowe powiązania tworzone są albo w warstwie ETL albo w warstwie Świata Obiektów. Należy również podkreślić, że nawet jeśli tablice w źródłowej bazie danych duplikują informacje bądź zawierają je w różnych miejscach nie ma to wpływu na opis w warstwie Świata Obiektów. Świat Obiektów może być zorganizowany w sposób spójny i logiczny merytorycznie

7. Dynamiczne Tworzenie Raportów

Absolutnym novum wynikającym ze stosowania narzędzia klasy Business Intelligence jest możliwość dynamicznego zmieniania i tworzenia nowych raportów. W warstwie prezentacyjnej prócz funkcji pozwalających np. na automatyczne wstawienie wykresu, czy dowolnego przekształcenia tabeli z danymi, istnieje możliwość stosowania rozbudowanego aparatu matematycznego. Dostępna jest większość funkcji matematycznych umożliwiająca prowadzenie analiz ekonomicznych. W skrócie mówiąc, w rękach zaawansowanego analityka biznesowego znajduje się cały aparat dostępny programiście i tylko kwestią strojenia wydajnościowego pozostaje, w której warstwie zaimplementować wymagane obliczenia. Należy zaznaczyć, że w przypadku bardzo często wykorzystywanych danych warto stworzyć odpowiedni zapis w bazie danych, na której oparto hurtownię danych, wówczas konieczne może okazać się rozbudowanie procedur warstwy ETL. W przypadku rzadziej wykorzystywanego przekroju danych można stworzyć prekalkulowany obiekt w Świecie Obiektów Business Objects. Ewentualnie jeżeli dana pozycja jest w fazie testowej, jej wyliczanie można pozostawić w warstwie tworzenia raportów. Jeżeli pozycja miałaby być wykorzystywana w dalszej fazie analiz, można ją dodatkowo zapisać w raporcie jako obiekt <wymiar, szczegół lub miarę>. Funkcjonalność narzędzia Business Objects umożliwia pracę zbiorową, ponieważ raport zapisywany jest wraz z tworzącym go zapytaniem. Po przesłaniu innemu analitykowi raport może powrócić do fazy tworzenia zapytania.

8. Organizacja Pracy – Czyli Co dla kogo ?

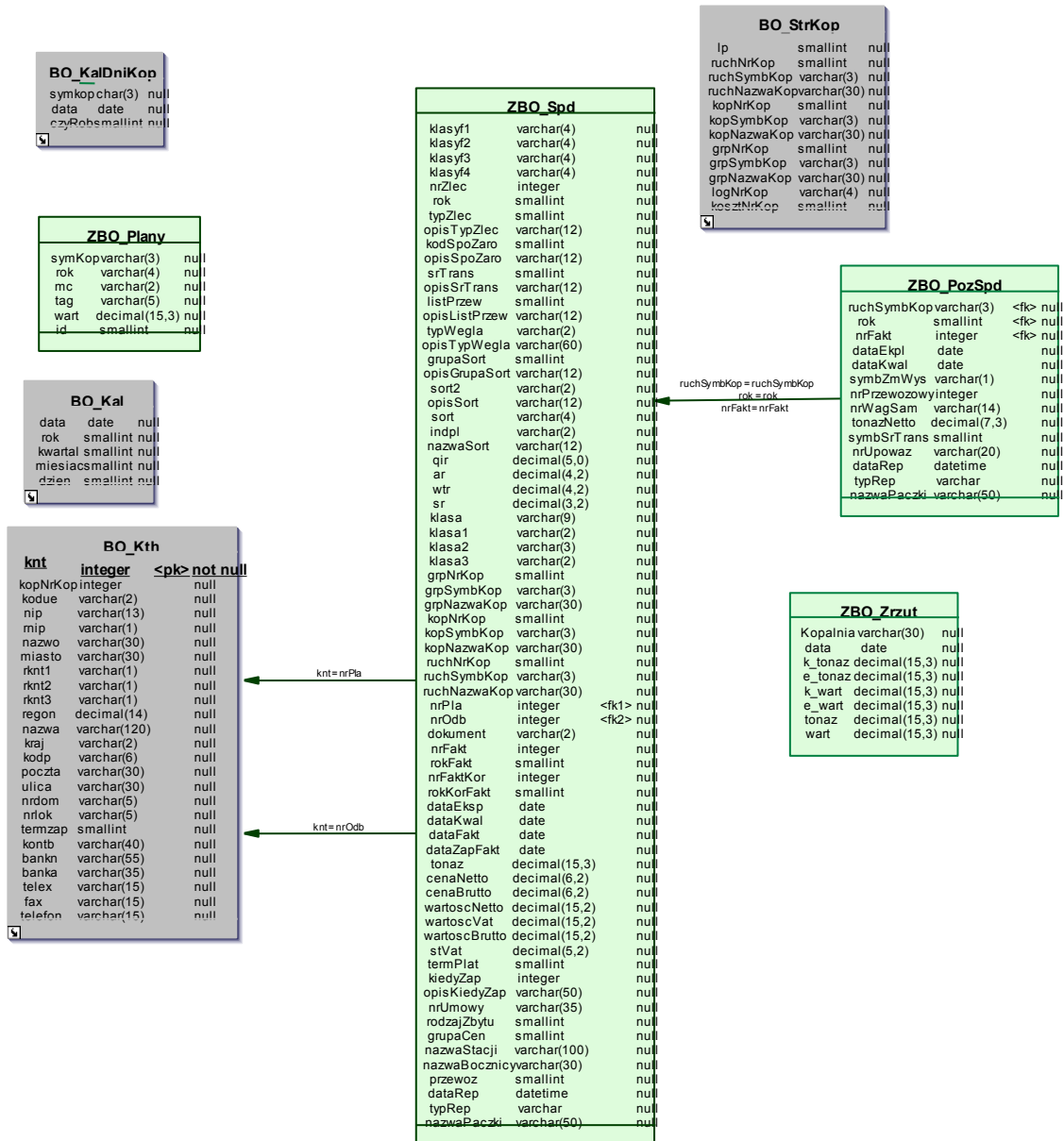
Na początku tego punktu przytoczę cytaty „Pierwsze zapytanie marketingowe, nigdy nie jest właściwym. To jest proces interakcyjny : zapytanie, odpowiedź, zapytanie, aż do wnikięcia w sens biznesu” – Pete Ester[5]. System Business Intelligence jest systemem, który umożliwia tworzenie raportów. Wydawałoby się więc, że największe zastosowanie znajdzie on dla pracowników związanych z szeroko rozumianą sprawozdawczością. Takie wykorzystanie oczywiście jest możliwe, za pomocą Business Objects można generować np. zestawienia sprawozdawczości obowiązkowej. Wydaje się jednak, że bardziej efektywne bardziej naturalne jest udostępnienie Business Objects pracownikom komórek zajmujących się Informacją Zarządzaną tj. szybką informacją, umożliwiającą sprawne podejmowanie operacyjnych decyzji z przedsiębiorstwie. Generalnie w systemie Business Intelligence rozróżnić należy następujące role :

- **Końcowy odbiorca raportów** – może to być zarówno manager wysokiego szczebla jak i na przykład kierownik magazynu. Ten pierwszy otrzymywałby informacje strategiczne, drugi zaś np. szczegółowe zestawienie stanów na określonych grupach materiałowych.
- **Analityk merytoryczny** – posiadający dostęp do Świata Obiektów i do edycji innych raportów, osoba taka odpowiedzialna byłaby za tworzenie raportów dla poprzedniej kategorii pracowników.
- **Designer** – osoba po stronie twórców oprogramowania, odpowiedzialna za stworzenie takiego Świata Obiektów, który będzie łatwy do zrozumienia dla grupy analityków merytorycznych.

9. Data Mart – Zbyt Węgla

Baza danych systemu Zbyt Węgla składa się z jednej tablicy zasadniczej i 7 tablic pomocniczych.

Na podstawie powyższego schematu widać różnicę w organizacji bazy danych przygotowanej dla hurtowni danych w stosunku do standardowej bazy relacyjnej. W celu pełniejszego zobrazowania budowy hurtowni zaprezentuję jeszcze zawartość głównej tablicy faktur.



Rys. 2 Tablice systemu Zbyt Węgl.

1.1.1.1 ZBO_Spd

Name	Code
klasyfikator 1	klasyf1
klasyfikator 2	klasyf2
klasyfikator 3	klasyf3
klasyfikator 4	klasyf4
numer zlecenia	nrZlec
rok zlecenia	rok
typ zlecenia	typZlec
opis typu zlecenia	opisTypZlec
kod sposobu zapłaty	kodSpoZaro
opis kodu sposobu zapłaty	opisSpoZaro
symbol środka transportu	srTrans
opis symbolu środka transportu	opisSrTrans
list przewozowy	listPrzew

opis listu przewozowego	opisListPrzew
typ węgla	typWęgla
opis typu węgla	opisTypWęgla
grupa sortymentu	grupaSort
opis grupy sortymentu	opisGrupaSort
zesłownikowany sortyment	sort2
opis sortymentu	opisSort
sortyment	sort
sposób wzbogacania	indpl
nazwa sortymentu	nazwaSort
kaloryczność	qir
popiół	ar
wilgoć	wtr
siarka	sr
klasa węgla	klasa
klasa1	klasa1
klasa2	klasa2
klasa3	klasa3
numer kopalni (grupa)	grpNrKop
symbol kopalni (grupa)	grpSymbKop
nazwa kopalni (grupa)	grpNazwaKop
numer kopalni (kopalnia)	kopNrKop
symbol kopalni (kopalnia)	kopSymbKop
nazwa kopalni (kopalnia)	kopNazwaKop
numer kopalni (ruch)	ruchNrKop
symbol kopalni (ruch)	ruchSymbKop
nazwa kopalni (ruch)	ruchNazwaKop
numer płatnika (knt)	nrPla
numer odbircy (knt)	nrOdb
dokument	dokument
numer faktury	nrFakt
rok wystawienia faktury	rokFakt
numer korekty do faktury	nrFaktKor
rok korekty faktury	rokKorFakt
data ekspedycji	dataEksp
data kwalifikacji	dataKwal
data wystawienie faktury	dataFakt
data zapłaty za fakturę	dataZapFakt
tonaż	tonaz
cena netto	cenaNetto
cena brutto	cenaBrutto
wartość netto	wartoscNetto
wartość VAT	wartoscVat
wartość brutto	wartoscBrutto
stopa VAT	stVat
termin płatności w dniach	termPlat
kiedy zapłacić (klucz do es17)	kiedyZap
opis kiedy zapłacić	opisKiedyZap
numer umowy (spozaro)	nrUmowy
rodzaj zbytu	rodzajZbytu
grupa cen	grupaCen
nazwa stacji	nazwaStacji
nazwa bocznicy	nazwaBocznicy
kto płaci przewoźne	przewoz
data wstawienia wpisu	dataRep

typ replikacji	typRep
nazwa paczki niosacej dane	nazwaPaczki

10. Wnioski

1. Podsumowując należy stwierdzić, że dzięki rozwiązaniom klasy Business Intelligence zarządzanie przedsiębiorstwem wchodzi w nowy etap. Uciążliwość uzyskiwania sprawozdań w różnym układzie była dotychczas bolączką wielu przedsiębiorstw. W dobie zwiększania efektywności pracy angażowała wykwalifikowanych pracowników do pracy polegającej na przenoszeniu danych z miejsca w miejsce. Dzisiaj istnieje możliwość powiązania dowolnych danych, również z innych systemów, w tym uzyskanych w trybie dostępu przez internet. Business Objects umożliwia dołączanie do Świata Obiektów danych z innych baz danych a także przechowywanych w postaci plików programu Excel czy nawet tekstowych plików płaskich.

2. Dotychczasowe doświadczenia w stosowaniu rozwiązań klasy Business Intelligence potwierdzają ich skuteczność i praktyczność w zaspakajaniu potrzeb wynikających z efektywnego funkcjonowania firmy.

Literatura

1. Praca zbiorowa. *Business Intelligence*, Wydawnictwo Libridis, 2002.
2. Ladanyi H.; *SQL – Księga Eksperta*, Wydawnictwo Heliun. Gliwice, 2000.
3. Koszowski Z., Syrkiewicz J.: *Kierunki przemian w obsłudze informatycznej górnictwa węgla kamiennego*. Wiadomości Górnicze nr 10. Katowice 1994.
4. Koszowski Z., Syrkiewicz J.: *Efektywność wdrażania Zintegrowanego Systemu Wspomagającego Zarządzania Przedsiębiorstwem – System SZYK, w sektorze górnictwa węgla kamiennego*, PTI – Efektywność zastosowań systemów informatycznych – tom III. Wydawnictwo Naukowo-Techniczne, Warszawa – Szczyrk 2002.
5. Liataud B.: *Business Intelligence Od informacji przez wiedzę do zysków*, Warszawa 2003.
6. Syrkiewicz J., Rymaszewski S.: *Zastosowanie systemów klasy Business Inteligence oraz aplikacji korporacyjnych w górnictwie węgla kamiennego*, Szkoła Ekonomiki i Zarządzania w Górnictwie, Bukowina Tatrzańska 2003r. Wydawnictwo AGH. Kraków, 2003.