

Justyna STASIEŃKO

Państwowa Wyższa Szkoła Techniczno-Ekonomiczna im. ks Bronisława Markiewicza
w Jarosławiu, Instytut Inżynierii Technicznej

BI IN-MEMORY – NOWA GENERACJA NARZĘDZI ANALITYCZNYCH

Streszczenie. Rynek systemów Business Intelligence (BI) jest jednym z najbardziej dynamicznych segmentów rynku oprogramowania. Dotychczasowe rozwiązania analityczne, oparte na kostkach OLAP, okazują się jednak niejedynym i nienajszybszym rozwiązaniem. Celem niniejszego artykułu jest prezentacja nowej technologii, tzw. BI in-memory. Technologia ta umożliwia zarządzanie informacją w taki sposób, jak to dzieje się w ludzkim mózgu.

Słowa kluczowe: Systemy Business Intelligence, analiza, raport, podejmowanie decyzji, OLAP

BI IN-MEMORY – THE NEW GENERATION OF ANALYTICAL TOOLS

Summary. The market of Business Intelligence Systems is one of the most dynamic sections of the software industry. The previous analytical solutions based on OLAP turn out not to be the only available and quickest ones. The aim of this article is to look through the traditional methods of analyzing and presenting new technology called BI in-memory. This technology makes it possible to manage the information in such a way as it takes place in human brain.

Keywords: Business Intelligence Systems, analysis, report, decision making, OLAP

1. Narzędzia przetwarzania analitycznego

Narzędzia analityczne są to narzędzia, które wspomagają decydenta, dostarczając mu wiedzę potrzebną do podejmowania decyzji w postaci raportów opartych na historycznych i bieżących danych. Narzędzia analityczne pozwalają na standardowe raporty oraz zaawanso-

wane analizy z wykorzystaniem statystyki, prognozowania, wyszukiwania zależności pomiędzy danymi, badanie trendów [4].

Do podstawowych narzędzi analitycznych zalicza się narzędzia generowania zapytań i raportowania (Query&Report – Q&R), arkusze kalkulacyjne, eksplorację OLAP oraz narzędzia wizualizacji danych [1].

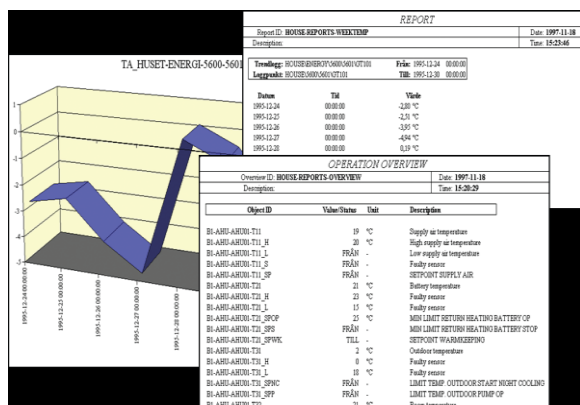
Narzędzia zapytań i raportowania są najbardziej podstawowym narzędziem analizy danych, w szczególności zawartych w hurtowniach danych. Rozróżnia się dwa rodzaje raportowania: standardowe i tzw. ad hoc. Najczęściej odpowiadają one napytania typu „co się stało?”, „na jakim poziomie kształtował się sprzedaż w poprzednim roku?” itp.

W drugiej połowie lat 90. XX wieku zauważono, że dane przechowywane w bazach danych transakcyjnych systemów wykorzystywanych w przedsiębiorstwach, sprawiają sporo kłopotów analitykom, którzy dokonywali oceny stanu biznesowego przedsiębiorstwa. Problem ten rozwiązano, wprowadzając techniki analityczne, realizowane między innymi w arkuszach kalkulacyjnych [1], które umożliwiają tworzenie modeli generujących cykliczne raporty w sposób automatyczny [10] (rys.1c). Arkusz kalkulacyjny oferuje elastyczność, jeśli chodzi o definicję warunków analizy i możliwości łatwej obsługi. Trudność pojawia się wtedy, kiedy podstawą analizy są duże wolumeny danych lub duża złożoność modelu. Aby osiągnąć pożądaną elastyczność analityczną, musi być zbudowany dosyć szeroki schemat przetwarzania, często oparty na dużej liczbie makropoleczeń i arkuszy. Takie rozwiązanie rodzi trudności związane z zarządzaniem przetwarzania danych, a przy tym podatne jest na błędy użytkownika. Arkusze mają też ograniczoną pojemność roboczą dla danych, co wręcz eliminuje ich przydatność dla analiz dużych porcji danych, idących w setki tysięcy transakcji. Na przeciw nowym wyzwaniom wyszły koncepcja baz wielowymiarowych oraz technologia OLAP, która pozwoliła na dynamiczną i wielowymiarową analizę danych biznesowych.

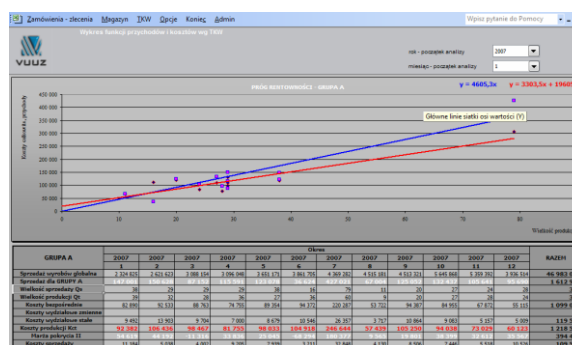
Architektura OLAP spotykana jest w hurtowniach danych i narzędziach do analizy, takich jak języki zapytań, eksploracji danych, sztucznej inteligencji, a także generatorach raportów. Poprzez odpowiedni sposób prezentacji, wizualizacji i agregacji, pozwala ona na wyświetlanie i oglądanie danych z różnych punktów widzenia, umożliwiając użytkownikowi szybką ich analizę. Dodatkowo cechuje ją możliwość interaktywnego raportowania bez znajomości języków programowania oraz uzyskiwania odpowiedzi na skomplikowane i często niestandardowe zapytania w trybie bieżącym. Dlatego też narzędzia OLAP często są używane do wykonywania analiz trendów sprzedaży, analiz finansowych (hurtownia danych) [6] czy też do wstępnego przeglądania zbioru danych przez analityka we wstępnej fazie analiz statystycznych [10].

Narzędzia wizualizacji danych służą zwiększeniu przejrzystości i czytelności prezentowanych informacji. Większość narzędzi analitycznych oferuje proste ilustracje zależności pomiędzy danymi.

Poza wyżej przedstawionymi podstawowymi narzędziami analitycznymi należy wymienić również zaawansowane narzędzia analityczne, takie jak pakiety statystyczne (np. STATISTICA), Data Mining oraz Text Mining [1].



a)



b)

Rys. 1. Wizualizacja analizy danych w: a) generatorze raportów, b) arkuszu kalkulacyjnym MS Excel

Fig. 1. Data visualization of analysis in: a) report generator b) spreadsheet MS Excel

2. Istota Business Intelligence

Business Intelligence to dziś najważniejszy i nieunikniony punkt styku informatyki oraz biznesu. Narzędzia Business Intelligence pozwalają na łatwy dostęp do informacji, jej analizę oraz współdzielenie w ramach całej organizacji i jej otoczenia biznesowego [2]. Dają możliwość integracji danych z różnych źródeł i kompleksowej analizy pod kątem potrzeb biznesowych. BI dają podgląd całości biznesu organizacji. Ich celem jest wspomaganie skutecznego zarządzania przedsiębiorstwem oraz planowania biznesu przez dostarczenie właściwych informacji [11]. Wspomagają prace menedżerów zarządzających kluczowymi obszarami działalności przedsiębiorstwa. Najbardziej ogólnie można przedstawić je jako proces przekształcania danych w informacje, a informacji w wiedzę [4], która może być wykorzystana do zwiększenia konkurencyjności przedsiębiorstwa. W grupie tych narzędzi znajdują się zarówno systemy zarządzania zasobami informacyjnymi, narzędzia raportowania i analiz, jak również rozwiązania pozwalające zarządzać wydajnością.

J. Surma [11] definiuje BI jako „zorientowany na użytkownika proces zbierania, eksploracji, interpretacji i analizy danych, który prowadzi do usprawnienia i racjonalizowania procesu podejmowania decyzji”. Według [2], oprogramowanie tworzące system BI składa się z: hurtowni danych, narzędzi do raportowania i zapytań *ad-hoc*, narzędzi OLAP, eksploracji danych.

Początkowo systemy BI zarezerwowane były dla wspomaganie decyzji w sferach strategicznej i taktycznej [5] – przy tworzeniu i rozwijaniu produktów, zarządzaniu finansami, za-

rządaniu wydajnością procesów. Aplikacje BI nowej generacji mają za zadanie udostępniać informacje niezbędne do podejmowania decyzji także na szczeblu operacyjnym.

Ogólnodostępnym systemem BI jest serwis Google. Jego zadaniem jest penetrowanie zasobów sieci internetowej i udostępnianie informacji o tych zasobach w wynikach wyszukiwania. Inteligencja Google określa typy zawartości serwisów internetowych, klasyfikuje je, określa ich atrakcyjność dla osób poszukujących informacji oraz nadaje priorytet. Potrafi też kojarzyć fakty, tematy, podobne znaczenia pojedynczych słów lub dłuższych fraz. Potrafi wyszukać informacje w innych językach niż zostało zgłoszone zapytanie. Potrafi interpretować zapytania użytkownika. Stara się zinterpretować zapytanie i udzielić odpowiedzi nawet wtedy, kiedy zapytanie nie jest precyzyjnie sformułowane, co jest istotą inteligencji.

3. BI in-memory

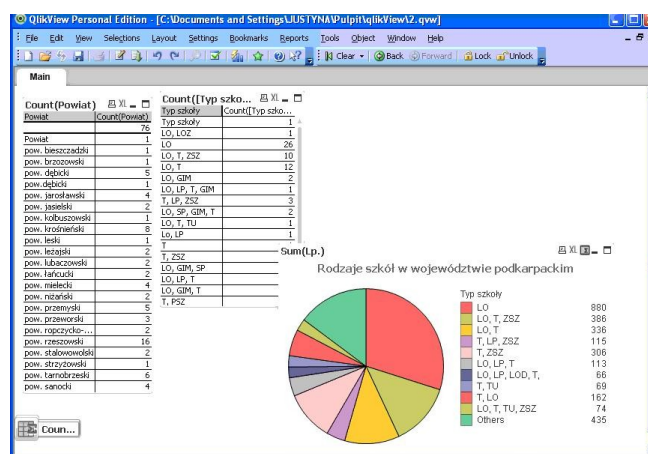
Rosnące wymagania użytkowników i klientów wpłynęły na ewolucję systemów wspomagających procesy decyzyjne. Rok 2008 przyniósł nową technologię – in-memory. Umożliwia ona ładowanie i przetwarzanie podstawowych bloków danych w pamięci operacyjnej. Oznacza to, że opracowane w niej rozwiązania analityczne można użytkować szybciej niż tradycyjne aplikacje. Wykorzystywanie pamięci operacyjnej do analizy i tworzenia raportów ma wpływ na architekturę BI. Ogranicza potrzebę budowania relacyjnych baz danych oraz używania kostek OLAP. Często znika też konieczność budowania hurtowni danych i następuje eliminacja większości ograniczeń starej technologii OLAP. Brak jest ograniczeń co do liczby wymiarów, które podlegają jednoczesnej analizie. Proces wdrażania tego typu narzędzia trwa zaledwie kilka tygodni, a nakłady na niego są proporcjonalnie niższe. Wymagania sprzętowe są dużo niższe niż wymagały tego tradycyjne systemy BI, dzięki temu czas realizacji wdrażania liczony jest już nie w miesiącach, lecz tygodniach, zaś dane po załadowaniu i silnym skompresowaniu poddawane są analizom w pamięci operacyjnej. Kalkulacje, nawet te najbardziej skomplikowane, wykonywane są z ogromną prędkością – czasy oczekiwania na odpowiedzi systemu są rekordowo niskie (nierzadko mierzone w ułamkach sekund), nawet przy olbrzymich wolumenach danych, sięgających miliardów rekordów. Jest to przełom wśród narzędzi analitycznych i systemów BI, który zdominuje rynek.

Dodatkowo nowa technologia może przyczynić się do popularyzacji narzędzi analitycznych wśród małych i średnich przedsiębiorstw, które do tej pory nie mogły sobie pozwolić na zakup aplikacji BI. Jest ona tańsza, gdyż nie wymaga zmian sprzętu oraz jest pewniejsza, jeśli chodzi o liczbę udanych wdrożeń. Istotną zaletą dla użytkowników małych i średnich przedsiębiorstw, gdzie rzadko są specjaliści z dziedziny informatyki, są mniejsze wymagania w zakresie znajomości SQL, bazy danych, hurtowni danych czy też języków programowania.

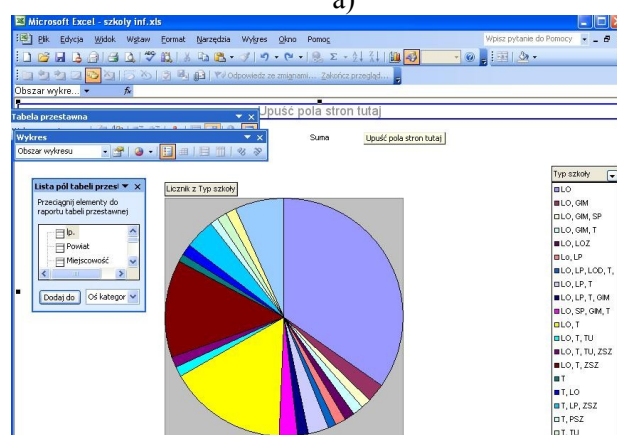
Jedynym istotnym ograniczeniem może być droga pamięć, a tym samym koszt ładowania do niej dużych ilości danych.

3.1. QlikView – nowa generacja BI

Liderem wśród nowej generacji narzędzi BI in-memory jest rozwiązanie QlikView firmy QlikTech [www.qlikview.com]. Oprogramowanie to służy do wykonywania wizualizacji analiz biznesowych, naśladując ludzki sposób myślenia. Obecnie nazywa się to asocjacyjną logiką QlikView, która lepiej odzwierciedla pracę umysłu człowieka – ludzie myślą w sposób nieliniarny, budując asocjacje. QlikView łączy wszystkie dane do pamięci i dokonuje kalkulacji metryk w czasie rzeczywistym, w momencie kliknięcia przez użytkownika. Analizy asocjacyjne wymagają intuicyjnego, wizualnie interaktywnego interfejsu, a także szybkości i elastyczności. Wszystko to zapewnia technologia operowania na danych w pamięci.



a)

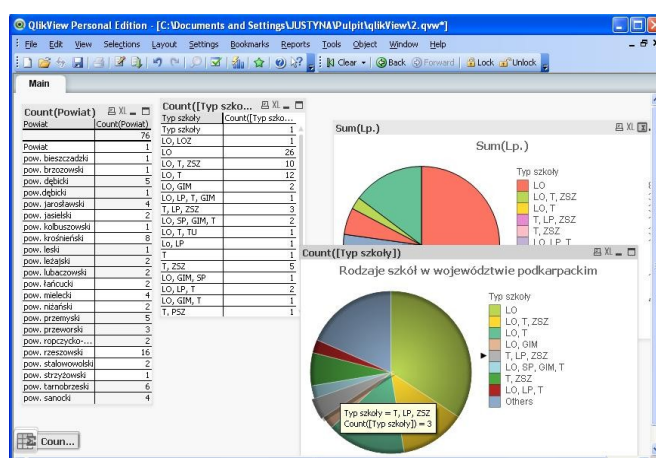


b)

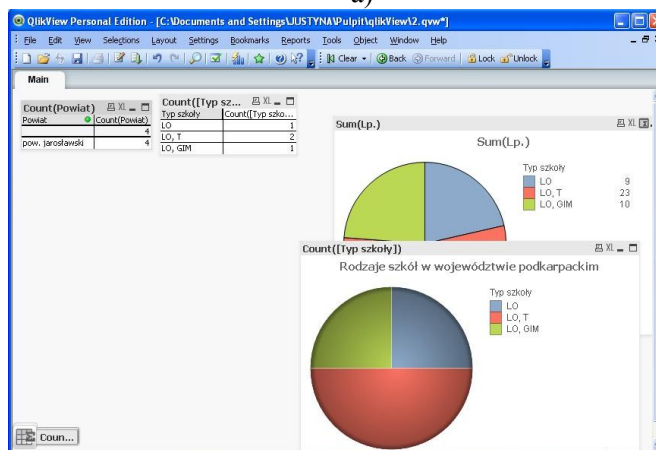
Rys. 2. Analiza danych: a) w aplikacji QlikView, b) w arkuszu kalkulacyjnym MS Excel
Fig. 2. The data analysis: a) in Qlik View application, b) in MS Excel

W porównaniu do innych narzędzi analitycznych czy też BI opartych na „kostkach” OLAP, aplikacja QlikView w szybszy i prostszy dla użytkownika sposób przedstawia anali-

zowane dane (rys. 2). QlikView daje możliwość nakładania kryteriów na dane. W przystępny i szybki sposób QlikView daje możliwość powrotu do wcześniejszych danych czy też na istniejących kryteriach dokładać kolejne (rys. 3a i rys. 3b). Ze względu na łatwość i prostotę obsługi oraz przyjazny i efektowny interfejs użytkownika, QlikView jest nowoczesną i niezwykle wydajną aplikacją. Posiada również, nieosiągalne dla innych rozwiązań, czas dostarczania nowych analiz, moc obliczeniową oraz elastyczność. Elastyczność oprogramowania polega na braku ograniczeń co do liczby wymiarów i miar, zaś jego moc na praktycznie natychmiastowej reakcji systemu na zapytania, nawet przy bazach danych liczących do pół miliarda rekordów. Daje także możliwość natychmiastowego przejścia do poziomu pojedynczej transakcji.



a)



b)

Rys. 3. Analiza rodzajów szkół na terenie woj. podkarpackiego: a) we wszystkich powiatach, b) w powiecie jarosławskim

Fig. 3. The analysis of all types of schools in Podkarpackie voivodeship: a) in all poviat, b) in jaroslawski poviat

Oprogramowanie QlikView umożliwia przenoszenie wyników analiz, które pozostają wciąż tak samo funkcjonalne, na inny sprzęt komputerowy. Aplikacja umożliwia drukowanie wyników w postaci raportów, wyeksportowanie ich do MS Excela lub zapisanie w formacie

PDF. Jest praktycznie w stanie integrować wszystkie formaty danych – od standardowych danych relacyjnych do raportów tekstowych, danych z Excela i strumieni XML.

Tabela 1 prezentuje zalety aplikacji QlikView w porównaniu z tradycyjnymi systemami BI.

Tabela 1

Różnice pomiędzy tradycyjnymi systemami BI a BI in-memory, np. QlikView

Tradycyjne BI	QlikView
Duża liczba narzędzi: hurtownie danych, składnice danych, OLAP, zapytania i narzędzia raportowania, eksploracja danych	Prosta architektura – wszystkie dane trzymane w pamięci
Techniki prezentacji analizy (jednostronnicowe raporty, raporty)	Techniki prezentacji analizy (wykresy)
Wielu użytkowników	Jeden użytkownik
Długi czas wdrażania (ok. 18 miesięcy)	Krótki czas wdrażania (kilka tygodni)
Wysokie koszty wdrażania	Niskie koszty wdrażania
Mało wygodny i elastyczny dla użytkownika	Łatwy w użyciu, elastyczny
Czasochłonny i złożony proces przetwarzania informacji	Szybki czas odpowiedzi, nieograniczona liczba wymiarów i miar możliwych do zmiany w ciągu kilku minut

Pod koniec 2010 roku pojawiła się kolejna wersja programu – QlikView 10. Nowością jest możliwość wdrażania oprogramowania na dowolnej platformie: lokalnie, cloud computing oraz na urządzeniach mobilnych. Udostępnienie na platformie „w chmurze” odbywa się za pośrednictwem serwisu Amazon Elastic Compute Cloud (EC2). Dysponuje również gigantyczną mocą analityczną, pozwalającą elastycznie przetwarzać wielkie zbiory danych przy zachowaniu dostępu do szczegółów. Nie jest ograniczony żadną liczbą wymiarów. Zmiany w zaprojektowanych aplikacjach mogą być wykonywane szybko i bez posiadania zaawansowanej wiedzy programistycznej. Aplikacja oferuje nowe możliwości prezentowania wyników analiz, korzystając m.in. z technologii AJAX. Nowe sposoby wizualizacji danych ułatwiają zrozumienie prezentowanych danych. Usprawnione zostało też wyszukiwanie informacji poprzez funkcje wyszukiwania asocjacyjnego.

Przykładem oprogramowania, które oparte jest na nowej technologii, jest darmowa aplikacja **Kick IT & Qlik IT** firmy QlikTech. Aplikacja jest przykładem mobilności BI in-memory oraz możliwości operowania na dużych zbiorach danych. Aplikacja dostarcza odpowiedzi na pytania związane z piłką nożną oraz pozwala na dokonywanie porównań za pomocą kilku ruchów. Pozwala fanom na nieograniczony dostęp do statystyk dotyczących Mistrzostw Świata w Piłce Nożnej. Dzięki tej aplikacji, dostępnej online z przeglądarki WWW lub za pomocą dowolnego urządzenia mobilnego, entuzjaści piłki nożnej mogą przeglądać ponad 80 lat historii mistrzostw. Mogą analizować dane historyczne i poznawać związane

z nim ciekawostki. Fani podczas Mistrzostw Świata w Piłce Nożnej w 2010 roku mogli na bieżąco korzystać z danych i analizować wydarzenia w RPA.

Innymi tego typu aplikacjami firmy QlikView są **The QlikView Cycling App** and **GlobalWaterViews**. Pierwsza, podobnie jak wyżej wymieniona Kick It & Qlik It, dostarcza fanom kolarstwa informacji na tematy związane z wyścigami Tour de France, które mogą być analizowane na wiele sposobów, np. rok, kraj, kolarz, drużyna itp. Natomiast GlobalWater Views dostarcza analiz związanych ze światowym kryzysem zanieczyszczenia wody.



Rys. 4. Strona główna aplikacji Kick IT & Qlik IT
Fig. 4. The home page of KickIt&QlikIt application

4. Zakończenie

Tradycyjne systemy BI integrują środowisko analityczne. Zapewniają samodzielność użytkownikom przez wykorzystanie wcześniej zdefiniowanych raportów i możliwość tworzenia analiz „na żądanie”. Udostępniają informacje w wymaganych przekrojach. Natomiast BI in-memory jest szybsza (jeśli chodzi o dostęp do danych, analiz), łatwiejsza w użyciu, bardziej elastyczna i skalowalna.

W systemach BI in-memory podejmowanie decyzji biznesowych odbywa się na podstawie aktualnych i rzetelnych informacji. Dostęp do zebranych informacji jest natychmiastowy, a informacje prezentowane są w czytelnej i przejrzystej formie analiz i raportów. Dzięki temu następuje poprawa organizacji informacji w przedsiębiorstwie. Zapewniona jest także ochrona oraz pełna kontrola dostępu do strategicznych zasobów informacyjnych firmy. Przyczynia się to do identyfikacji braków danych i braków w informacji zarządczej. Nowa technologia umożliwi szybką i elastyczną reakcję na potrzeby analityczne różnych odbiorców.

Kluczowe dla analityki biznesowej będą w najbliższym czasie cztery aspekty: prostota, spójność i integracja, podejście społecznościowe oraz skuteczna realizacja strategii.

BIBLIOGRAFIA

1. Dudycz H.: Przetwarzanie analityczne podstawą rozwiązań informatycznych klasy Business Intelligence. Materiały konferencyjne SWO, Katowice 2004, http://www.swo.ae.katowice.pl/_pdf/138.pdf.
2. Januszewski A.: Funkcjonalność Informatycznych Systemów Zarządzania. T. 2. Systemy Business Intelligence. PWN, Warszawa 2008.
3. Matouk K., Owoc M.: Rola integracyjno-kooperacyjna business intelligence w systemach informatycznych zarządzania. Materiały konferencyjne SWO, Katowice 2004, http://www.swo.ae.katowice.pl/_pdf/142.pdf.
4. Nycz M., Smok B.: Business Intelligence w zarządzaniu. Materiały konferencyjne SWO, Katowice 2008, http://www.swo.ae.katowice.pl/_pdf/421.pdf.
5. Olszak C. M, Ziemia E.: Approach to Building and Implementing Business Intelligence Systems, [in:] Interdisciplinary Journal of Information, Knowledge, and Management, Vol. 2, Informing Science Institute, Santa Rosa USA 2007, <http://ijikm.org/Volume2-/IJKMv2p135-148Olszak184.pdf>.
6. Olszak C., Ziemia E.: Systemy Business Intelligence w rozwoju holistycznej infrastruktury wspomagającej podejmowanie decyzji w organizacji. Roczniki, Kolegium Analiz Ekonomicznych, nr 16, SGH, Warszawa 2006, <http://www.ae.katowice.pl/images/user/File-/katedra%20informatyki%20ekonomicznej/NTIE2006-Olszak-Ziemia.pdf>.
7. Olszak C.: Systemy Business Intelligence we wspomaganiu decyzji zarządczych organizacji, [w:] Knosali R. (red.): Komputerowo zintegrowane zarządzanie. WNT, Warszawa 2004, s. 241-248.
8. Stasińko J.: Technologie Business Intelligence w procesie podejmowania decyzji. Roczniki KAE SGH, Warszawa 2009, s. 149-162.
9. Stefanów P., Zając A., Kajrunajtyś D., Kuraś M.: Narzędzie zarządzania. Raport Specjalny Business Intelligence. ComputerWorld, nr 09/848, 2009.
10. Sierocki R.: OLAP to efektywna technologia przetwarzania danych analitycznych. Controlling i Rachunkowość Zarządcza, nr 1/2007, <http://www.infosynergia.eu/pliki/publikacje%20-%20Robert%20Sierocki%20%20OLAP%20to%20efektywna%20technologia.pdf>.
11. Surma J.: Business Intelligence – Systemy Wspomagania Decyzji Biznesowych. PWN, Warszawa 2009.
12. <http://www.qlikview.com> [20.07.2010].
13. <http://www.hogart.pl> [20.01.2011].
14. <http://www.businessintelligence.pl/index.php/aktualnosci/89-narzedzie-qlikview-9-wyznacza-nowe-standardy-bi09> [20.07.2010].

15. http://searchdatamanagement.techtarget.com/sDefinition/0,,sid91_gci213571,00.html [10.08.2010].
16. <http://accre.pl/index.php?solution/businessintelligence/narzedzia> [10.08.2009].
17. <http://www.businessintelligence.com/> [09.07.2009].
18. <http://www.informationbuilders.com/new/magazine/v11-2/pdf/DRESNERI.PDF> [05.05.2010].

Recenzenci: Dr inż. Henryk Josiński

Dr hab. inż. Mirosław Zaborowski

Wpłynęło do Redakcji 31 stycznia 2011 r.

Abstract

Existing analytical data acquisition and processing tools, such as report and queries or spreadsheets were not able to fully satisfy the needs of managers, who were supposed to make a more thorough analysis relatively fast. To meet these challenges there came a concept of multidimensional databases and OLAP technology, which enabled a dynamic and multi-dimensional business data analysis. You can come across OLAP architecture in data warehouses and analysis tools such as query languages, data exploration, artificial intelligence, as well as report generators. Another device used for processing data into information, information into knowledge, and multidimensional analysis of these relationships are Business Intelligence systems.

“Business intelligence (BI) is a broad category of applications used for gathering, storing, analysing, and providing access to data to help managers make better business decisions. BI applications encompass decision support systems, query and reporting, online analytical processing (OLAP), statistical analysis systems, forecasting, and data mining” [portal searchdatamanagement.bitpipe.com]. Increasing user expectations call for improvements of existing tools and methods. Attempt to limit the building of relational databases, OLAP cubes and even strenuous data warehouse technology has led to the development of “in-memory” technology. This technology enables loading and processing the basic sets of data in the operational memory which contributes to the quicker use of analyses. The use of memory for analysis and reporting has an impact on BI architecture. The latest analytical tools pose much lower hardware requirements; they do not need OLAP 'cubes' since they directly integrate with external systems. There is no need to build the relational databases and relating on the

labor-intensive OLAP cubes. The new technology does not restrict the number of dimensions that are analysed.

QlikView is a leader in the new generation of BI 'in-memory' tools made by QlikTech Company. The analysis made in QlikView can be transferred to another computer still remaining totally functional. The results can be printed in the form of reports, exported to MS Excel or saved in PDF. Qlik View enables integrating all data formats – from standard relational data to texts reports, Excel and XML data .

Adres

Justyna STASIEŃKO, Państwowa Wyższa Szkoła Techniczno-Ekonomiczna im. ks Bronisława Markiewicza, Instytut Inżynierii Technicznej, ul. Czarnieckiego 16, 37 -500 Jarosław, Polska, justyna.stasienko@pwste.edu.pl.