

prof. dr hab. inż. Marian Kopczewski
WSB w Poznaniu

Hurtownie danych narzędziem wspomagającym procesy decyzyjne w zarządzaniu

Data warehousing tool for decision – making in management

Streszczenie: *Każdy decydent powinien pamiętać, że doskonały sprzęt komputerowy to tylko potencjał, który może być wykorzystany dopiero po uruchomieniu odpowiednich systemów informatycznych. Często nie zdajemy sobie sprawy, że komputer jest tak dobry, jak dobre jest jego oprogramowanie, o czym niestety zapominamy wprowadzając systemy informatyczne do zarządzania, w tym podejmowania różnorodnych decyzji, w różnego rodzaju organizacjach. Można zatem stwierdzić, że komputer jest tak dobry, jak dobre jest jego oprogramowanie, pozwalające na zebranie, określenie, zarchiwizowanie, analizę i dystrybucję kluczowych informacji niezawodnych w procesie decyzyjnym. Dynamicznie rozwijające się organizacje (dysponujące już systemami informacyjnymi), aby wkroczyć na wyższy poziom zarządzania, redukować koszty, efektywniej obsługiwać, a szczególnie podejmować racjonalne decyzje, korzystają z hurtowni danych.*

Słowa kluczowe: system informacyjny, relacje systemowe, decyzja, hurtownia danych, przedsiębiorstwo.

Abstract: *Every decision maker should remember that the perfect hardware is only potential, which can be used only after the relevant systems. Often we do not realize that the computer is so good, how good is its software, which unfortunately we forget introducing management information systems, including a variety of decision-making, in all kinds of organizations. It can therefore be concluded that the computer is only as good as its software is good, allowing to gather, define, archiving, analysis and distribution of key information reliable decision-making. Dynamically developing organizations (already possessing information systems) in order to enter the next level of management, reduce costs, operate more efficiently, and especially to make rational decisions benefit from data warehouses.*

Keywords: information system, the relationship system, the decision, data warehouse, company.

Wstęp

Rozwój technologii informatycznych coraz bardziej pomaga przedsiębiorstwom stać się bardziej dochodowymi i konkurencyjnymi. Zdobywanie przewagi nad konkurencją wymaga od przedsiębiorstw przyspieszania procesu podejmowania decyzji, tak aby szybko mogły one reagować na zmiany. Kluczowym w przyspieszaniu procesu podejmowania decyzji jest posiadanie, w odpowiednim czasie, właściwej i łatwo dostępnej informacji. Obecne systemy zarządzania przedsiębiorstwami zawierają dane, które mogą dostarczyć informacji niezbędnej do podjęcia decyzji. Jednak w większości przedsiębiorstw

dane te są rozrzucone po wielu różnych systemach, platformach i miejscach, co sprawia, że zapewnienie integralności danych oraz dostępu do nich w rozsądnie krótkim czasie jest prawie niemożliwe.

W celu dostarczania niezbędnych informacji wypracowujących decyzje, projektuje się hurtownię danych wspomagającą podejmowanie decyzji. Postęp technologiczny sprawia, że stworzenie hurtowni danych dostarczającej na czas wiarygodnej informacji do podejmowania decyzji staje się możliwe praktycznie w każdej organizacji. Celem artykułu jest ogólne przedstawienie hurtowni danych jako elementu wspomagania procesu podejmowania decyzji.

System informacyjny

System informacyjny jest to formalny system komputerowy dokonujący wyboru, udostępniania i integracji pochodzących z różnych źródeł danych po to, by w odpowiednim czasie dostarczyć informacji niezbędnych do podejmowania decyzji. Wyróżniamy następujące typy systemów informacyjnych [2]:

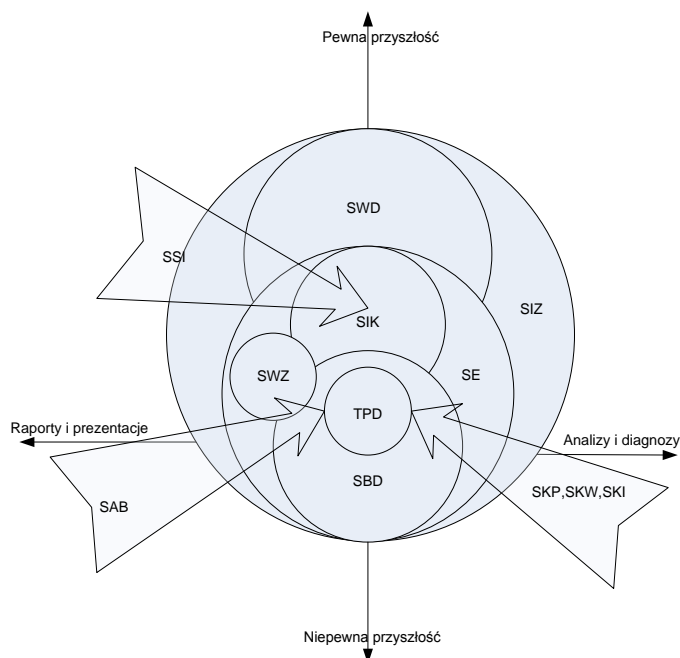
- Systemy Transakcyjne Przetwarzania Danych – TPD – (Transaction Processing Data Systems – TPS) zwane też Systemami Elektronicznego Przetwarzania Danych (Electronic Data Processing – EPD);
- Systemy Informacyjne Zarządzania – SIZ – (Management Information Systems – MIS), na swoim obecnym szczeblu rozwoju utożsamiane z Systemami Zarządzania Relacyjnymi Bazami Danych (Relational Data Base Management Systems (RDBMS));
- Systemy Wspomagania Decyzji – SWD – (Decision Support Systems – DSS);
- Systemy Ekspertkie – SE – (Expert Systems – ES);
- Systemy Informowania Kierownictwa (Executive Information Systems – EIS); utożsamiane niekiedy z Systemami Wspomagania Kierownictwa (Executive Support Systems – ESS), w literaturze polskiej występujące często pod wspólnym akronimem: SIK;
- Systemy Sztucznej Inteligencji – SSI- (Artificial Intelligence Systems) utożsamiane często z Systemami Sieci Neuronowych – SNN – (Artificial Neuron Networks – ANN);
- Systemy Wspomagania Zarządzania – SWZ – (Management Support Systems – MSS) w literaturze polskiej określanej skrótem SWZ;
- Systemy Automatyzacji Biura -SAB – (Office Automation Systems – OAS);
- Systemy Komputerowego Projektowania – SKP – (Computer Aided Designing – CAD);
- Systemy Komputerowego Wspomagania Produkcji – SKW- (Computer Aided Manufacturing – CAM);
- Systemy Komputerowego Wspomagania Inżynierii – SKI – (Computer Aided Engineering – CAE);

W procesach decyzyjnych z wymienionych systemów informacyjnych najważniejsze to [6]:

- Systemy informowania kierowania (SIK), dostarczające informacje o przeszłości, teraźniejszości, przewidywanej przyszłości oraz o istotnych zdarzeniach wewnątrz i na zewnątrz organizacji.
- Systemy wspomagania decyzji (SWD), to interaktywny system przetwarzania i prezentacji danych. Składa się on z następujących elementów:
 - baz modeli;
 - baz danych.
- Systemy eksperckie (SE) to systemy informatyczne, których oprogramowanie wspomaga eksperta – decydenta w poszukiwaniu optymalnych rozwiązań problemu.

Relacje między systemami informacyjnymi

Między systemami informacyjnymi zachodzi szereg wzajemnych relacji:



Rys. 1. Relacje między systemami informacyjnymi

Do najważniejszych relacji należą [7]:

- Istnienie uporządkowanego i prawidłowo funkcjonującego systemu przebiegu informacji w organizacji.
- Kompleksowe skomputeryzowanie obiegu informacji w organizacji.
- Zachowanie kolejności w projektowaniu komputeryzacji obiegu informacji w organizacji – w pierwszej kolejności zaprojektowanie systemu obiegu informacji, a następnie przystosowanie lub adaptacja do jego potrzeb oprogramowania i sprzętu komputerowego.

Analiza systemów informacyjnych

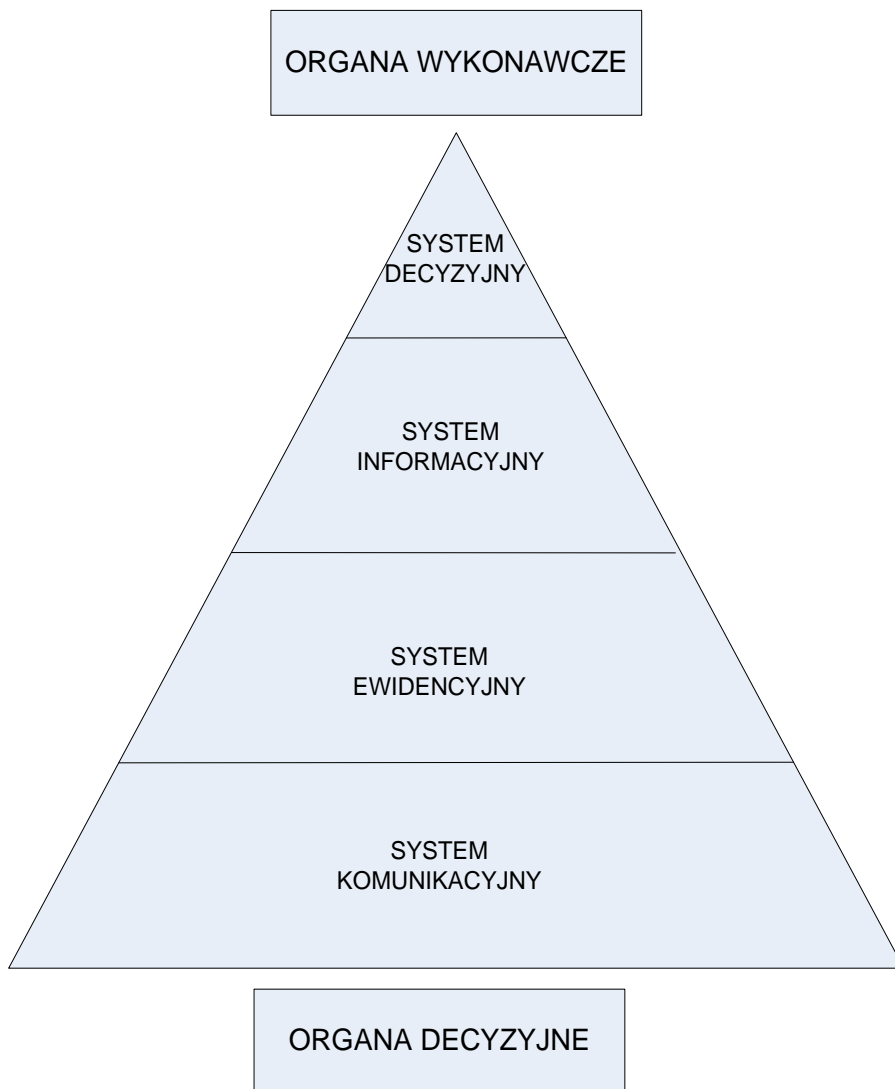
Zakres analizy każdego z systemów informacyjnych można przedstawić jako cel i wynik analiz.

Analiza	
Cel	Wynik
1. Co i przy jakich ograniczeniach system ma robić?	Zestawienie wymagań i założeń, czyli zewnętrzny opis systemu.
2. Jak system ma działać?	Logiczny opis systemu, czyli opis sposobu realizacji przez system postawionych wymagań.

Techniki analiz systemów informacyjnych [1]

FAZY I ETAPY CYKLU ŻYCIA SYSTEMU	TECHNIKA
Analiza potrzeb informacyjnych	<ul style="list-style-type: none"> • Analiza dokumentacji • Wywiad • Ankieta • Analiza dokumentów
Modelowanie funkcji systemu (struktura funkcjonalna systemu)	<p>Diagramy hierarchii funkcji (ang. Function Hierarchy Diagram FHD)</p> <p>Diagramy zależności funkcji (ang. Function Relationship Diagram FRD)</p>
Modelowanie procesów systemu (struktura funkcjonalna systemu)	Diagramy przepływu danych (ang. Data Flow Diagram DFD)
Modelowanie danych (struktura informacyjna systemu)	<p>Diagramy powiązania danych, związków encji (ang. Entity Relationship Diagram ERD)</p> <p>Słowniki danych (ang. Data Dictionary DD)</p>

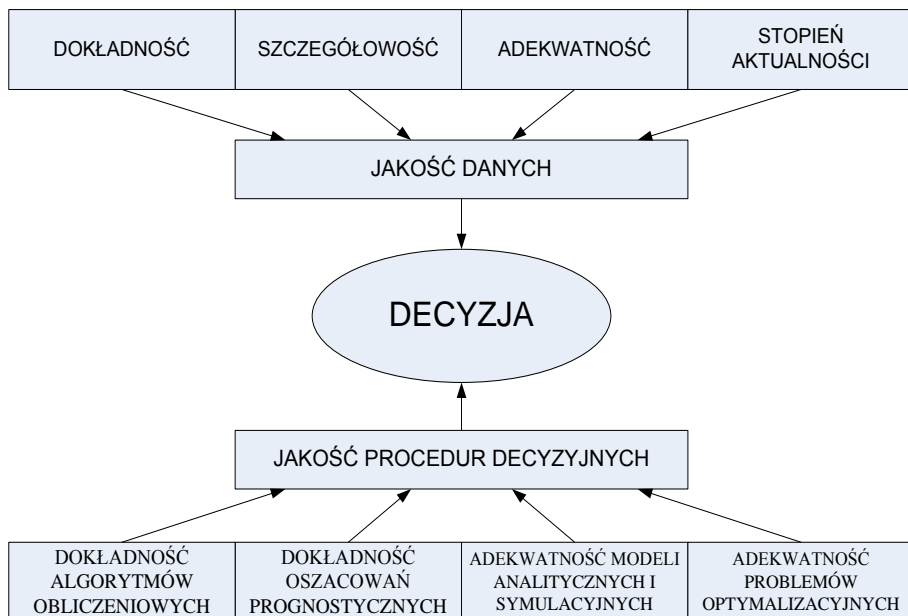
Uwzględniając zakres i techniki analiz systemów informacyjnych, faza cyklu decyzyjnego powinna obejmować następujące procesy:



Rys. 2. Fazy cyklu decyzyjnego

Dlatego uwarunkowania projektowe powinny być rozpatrywane od dolnych faz. Projekty systemów muszą być przygotowywane przez zaplecze naukowo-badawcze organizacji reprezentującej użytkownika wobec organów projektowo-produkcyjnych informatyki.

Reasumując, o jakości systemów informatycznych stosowanych w procesach decyzyjnych, szczególnie kierowania ogniem, świadczy zawsze wartość decyzji [6].



Rys. 3. Czynniki wpływające na decyzje

Hurtownia danych

Jednym z podstawowych, a zarazem najnowocześniejszych w polskim biznesie systemów informatycznych wspomagających proces decyzyjny (spełniający przedstawione wymagania we wcześniejszej części artykułu) jest koncepcja hurtowni danych. Hurtownia danych, według twórcy tej koncepcji W.H. Inmona, to zintegrowany, tematyczny, zmienny w czasie zbiór danych dla wspomaganie procesów podejmowania decyzji zarządczych.

Celem hurtowni danych jest zapewnienie użytkownikom, a przede wszystkim decydentom szybkiego dostępu do danych. Aby ten cel zrealizować, hurtownia danych powinna spełniać kilka warunków [7]:

- przede wszystkim musi gromadzić dane w taki sposób, by istota informacji nie została utracona, chociażby ze względów biznesowych;
- dostęp do informacji powinien odbywać się na przejrzystych zasadach, zaś zasilanie systemu aktualizowanymi danymi musi odbywać się w czasie, który nie podważa sensu istnienia systemu wspomagającego decyzję;
- czas oczekiwania na dane powinien być krótszy od czasu, niezbędnego do podjęcia decyzji;
- użytkownicy hurtowni danych powinni mieć możliwość samodzielnego przeprowadzania złożonych analiz (np. o przeciwniku), bez pomocy informatyków;
- najważniejszą kwestią jest koszt systemu, który nie może być większy od zysków, jakie osiągnie się w wyniku jego wdrożenia.

Hurtownia danych jest dedykowanym systemem baz danych, który w odróżnieniu od systemu transakcyjnego (bazującego na danych operacyjnych) charakteryzuje się następującymi cechami:

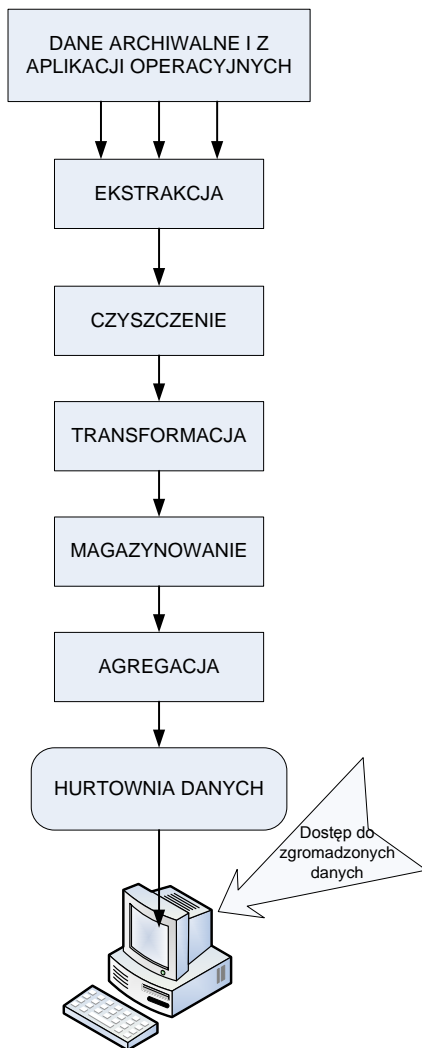
- dane są utrzymywane w dużo dłuższym horyzoncie czasowym;
- utrzymywanie danych jest optymalizowane pod kątem odpowiadania na złożone zapytania pochodzące od analityków i zarządzających;
- dane są pozyskiwane z różnorodnych źródeł przy czym zapewniona jest ich jednolitość.

Cecha	Aplikacje operacyjne	Hurtownie danych
Zasilanie danymi	Równomierny napływ wielu krótkich transakcji	Pojedyncze, bardzo duże transakcje podczas ładowania danych
Wzorzec dostępu do danych	Dobrze określony przez użytkownika	Mało przewidywalny, zależny od bieżących potrzeb użytkownika (ad hoc)
Tryb dostępu do danych przez użytkowników	Zapis, odczyt, modyfikacja	Tylko odczyt, dużych ilości danych w jednym zapytaniu
Horyzont czasowy przechowywania danych	Dane bieżące	Pełna historia

W hurtowni przechowuje się dane różnych rodzajów, jak: elementarne – kopie aktualnych danych źródłowych pozyskanych z baz operacyjnych i przetworzonych zgodnie z potrzebami;

- historyczne – dane elementarne i/lub agregaty dotyczące przeszłości;
- zmaterializowane agregaty – czyli wyliczone wartości obliczeń, w różnych przekrojach i na różnych stopniach agregacji;
- metadane – informacje słownikowe, które opisują strukturę hurtowni danych i źródłowych baz danych, na podstawie których uzyskuje się dane do hurtowni danych, oraz sposób wyliczania danych zagregowanych.

Cykl życia danych w hurtowni zwykle wygląda następująco [1]:



Rys. 4. Cykl życia danych

Hurtownia danych w strukturze organizacji

Sukces organizacji w procesie kierowania w znaczącej mierze zależy od stanu informacji wewnętrznych, jak i zewnętrznych o jego otoczeniu. Wyróżniamy dwa zasadnicze typy systemów wspomagające działalność organizacji: systemy o przetwarzaniu operacyjnym i systemy przetwarzania analitycznego.

Systemy o przetwarzaniu operacyjnym. W codziennej działalności organizacji zmilitaryzowanej, w systemach informatycznych jest wykorzystywane przetwarzanie operacyjne. Są to systemy przetwarzania transakcji OLTP (On-Line Transaction Processing) [6], zaprojektowane do zadań operacyjnych,

związanych z działalnością oraz zoptymalizowane pod kątem przetwarzania niewielkich porcji danych. W bazie danych takiego systemu przechowywane są jedynie aktualne informacje o większości obiektów. Dane w bazie są aktualizowane na bieżąco przez cały czas. Dane historyczne przeważnie przechowuje się na oddzielnych nośnikach i nie udostępnia na bieżąco. Systemy takie nadają się wyłącznie do tworzenia standardowych raportów dokumentów, które obsługują bieżącą działalność. Ten model nie wspomaga procesów analizy danych oraz aplikacji wspomagających podejmowanie decyzji.

Systemy o przetwarzaniu analitycznym [6]. Wspomaganie decyzji stanowi drugi bardzo ważny obszar zastosowania systemów baz danych. Rozwijające się gałęzie przemysłu, nauki i uzbrojenia wymagają składowania i przetwarzania ogromnych ilości danych. Posiadanie danych opisujących działanie organizacji w dłuższym przedziale czasu pozwala na analizę trendów, anomalii, poszukiwania wzorców zachowań. Informacje przechowywane w bazie danych stanowią olbrzymie źródło wiedzy, którą możemy wykorzystać w procesie podejmowania strategicznych i ogniowych decyzji. Dla potrzeb tej analizy opracowano nowy model przetwarzania danych nazywany przetwarzaniem analitycznym OLAP (On-Line Analytical Processing). Zadaniem tych systemów jest wspomaganie zarządzania poprzez dostarczanie właściwych informacji potrzebnych do analizy problemu bądź sytuacji, właściwym ludziom, we właściwym czasie przy niskim koszcie. Korzystając z tej technologii można budować aplikacje, dzięki którym możliwa będzie analiza wszystkich obszarów działalności organizacji, szczególnie ogniowej. Podstawą przetwarzania są porównania oraz analizy wzorców i trendów. Analityczne bazy danych nie przechowują najświeższych informacji, ale informacje odnoszące się do konkretnych punktów w czasie.

Dla potrzeb przetwarzania analitycznego oraz przechowywania ogromnej ilości stałych danych przez długi okres czasu stworzono nowy typ architektury bazy danych nazwany hurtownia danych (ang. Data Warehouse) [1].

Cecha	OLTP	Hurtownia danych
Szczegółowość danych	duża	średnia lub mała
Tryb dostępu do danych	odczyt oraz zapis	tylko odczyt
Konsolidacja danych	mała	duża
Powtarzalność danych	duża	mała
Częstość operacji na danych	bardzo duża	średnia, bądź mała
Aktualność danych	bardzo duża	średnia lub mała
Horyzont czasowy	bardzo krótki lub krótki	długi (dane historyczne)
Rozmiar danych	MB, GB	GB, TB
Liczba przetwarzanych rekordów w jednej operacji	dziesiątki	miliony

Źródło: Vidette Poe, Patricia Klauer, Stephen Brobst, *Tworzenie hurtowni danych*.

Podsumowanie

Gwałtowny rozwój technologii informatycznych spowodował, że systemy komputerowe zaczęły pełnić istotną rolę w procesach decyzyjnych. Ma to miejsce szczególnie tam, gdzie do podjęcia decyzji konieczne jest szybkie przetworzenie ogromnych ilości danych (np. w dużych przedsiębiorstwach), lub gdzie charakterystyka sytuacji decyzyjnej wymaga zastosowania skomplikowanych obliczeniowo modeli. Systemy informatyczne zastąpiły systemy informacyjne. Powstały systemy wspomagania decyzji takie jak: sztuczna inteligencja oraz specjalistyczne systemy decyzyjne. Przedstawione rozważania prowadzą do wniosku, że hurtownie danych są produktem dedykowanym dla konkretnych organizacji. Stworzenie dobrze działającej hurtowni wymaga od dostawcy oprogramowania dużego wysiłku, koniecznego w momencie definiowania zawartości hurtowni i metod przekształcania danych źródłowych w analityce procesu zarządzania, a szczególnie podejmowania decyzji.

Reasumując podstawową wiedzę o hurtowni danych, można stwierdzić, że:

- hurtownia danych automatyzuje zbieranie, przetwarzanie oraz dystrybucję danych użytecznych dla osób podejmujących decyzje, w tym ogniowe;
- hurtownie danych są systemami zdolnymi do wsparcia długoterminowych strategii, przygotowania organizacji do działań w tym ogniowych;
- informacje dostarczane przez komputerowe systemy zarządzania może stanowić podstawę do wielu badań związanych z bieżącą sytuacją i perspektywami rozwoju organizacji, szczególnie w zakresie nowego sprzętu bojowego i systemów dowodzenia. Zagadnienie to nabiera szczególnego znaczenia w przypadku zarządzania rozproszonymi organizacjami, ugrupowaniami bojowymi. Wnikliwa i prawidłowa analiza wymaga ujednoczenia informacji dostarczanych przez różne, często nie zintegrowane ze sobą systemy informacyjne, funkcjonujące w ramach jednej organizacji. Dzięki prawidłowej konfiguracji hurtowni danych i właściwemu wykorzystaniu możliwości związanych z jej funkcjonowaniem poprawia się jakość generowanych informacji. W oparciu o bardziej precyzyjne dane możliwe jest zwiększenie prawdopodobieństwa podjęcia właściwej decyzji;
- interpretacja danych dostarczanych przez hurtownie danych stwarza możliwości badań nad efektywnym zarządzaniem organizacją, a w tym każdym przedsiębiorstwem.

Bibliografia

- Poe V., Klauer P., Brobst S., *Tworzenie hurtowni danych*, WNT Warszawa 2000.
- Jarke M., Lenzerini M., Vassiliou Y., Vassiliadis P., *Hurtownie danych. Podstawy organizacji i funkcjonowania*, Warszawa 2003.
- Pałasz P., http://republika.pl/pawel_palasz/ja/Projektowanie.html
- Golfarelli M., Rizzi S., *Designing the Data Warehouse: Key Steps and Crucial Issues*.
- Armstrong-Smith, Michale i Darlene, *Oracle Discoverer*, Helion 2003.

Kopczewski M., *Hurtownia danych elementem współczesnego systemu Business Objects*, Materiały Konferencji Komputerowo Zintegrowane Zarządzanie, WNT Warszawa 2003.

Kopczewski M., *Hurtownia danych elementem wspomaganie podejmowania decyzji*, Materiały IV KK Metody i Systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim, WNT Kraków 2003.