



POZNAN UNIVERSITY OF TECHNOLOGY

Data Warehouses and Business Intelligence: Technology Overview

Robert Wrembel
Poznan University of Technolog
Institute of Computing Science
Robert.Wrembel@cs.put.poznan.pl
www.cs.put.poznan.pl/rwrembel



Topics

- ⇒ **OLAP/ Business Intelligence/ Data Analytics/ Business Analytics**
- ⇒ **Przetwarzanie OLTP vs. OLAP**



Cele stosowania HD

1. Zapewnienie jednolitego dostępu do wszystkich danych gromadzonych w ramach przedsiębiorstwa
2. Dostarczenie technologii (platformy) przetwarzania analitycznego - **technologii OLAP/BI**



Business Intelligence

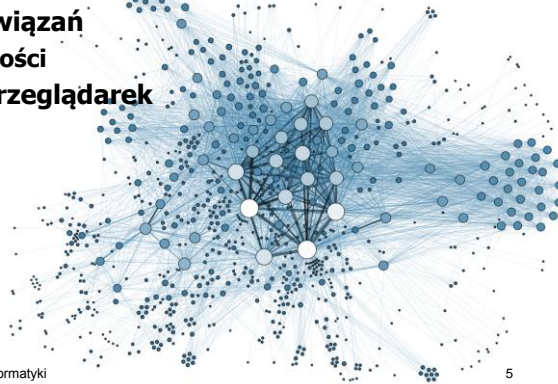
- ⇒ **OLAP - On-Line Analytical Processing**
 - klasyczna analiza danych (dane historyczne, predykcja - what if analysis)
 - analiza trendów sprzedaży
 - analiza nakładów reklamowych i zysków
 - analiza ruchu telefonicznego
 - credit scoring
 - churn analysis
 - customer profiling
 - najczęściej SQL



Business Intelligence

⇒ BI = OLAP+

- eksploracja danych
 - reguły asocjacyjne, profile zachowań
- analiza tekstów (Facebook, Tweeter, ...)
 - hot topics, bezpieczeństwo narodowe
- analiza sieci powiązań
 - liderzy, zależności
- analiza logów przeglądarek



R.Wrembel - Politechnika Poznańska, Instytut Informatyki

5



OLTP a OLAP

	OLTP	OLAP
użytkownik	"zwykły"	analityk
funkcja	bieżące operacje, kluczowe dla działania firmy	wspomaganie decyzji
dane	bieżące, elementarne	elementarne, zagregowane, historyczne
aplikacje	powtarzalność działań	ad hoc
dostęp	odczyt/zapis	odczyt
transakcja	krótka	długa (godziny)
I. przetwarzanych rek.	kilka, kilkadziesiąt	miliony lub więcej
I. użytkowników	kilkudzies., tysiące, setki tys. miliony, dzies. milionów	kilku, kilkunastu
DB size	kilka - setki TB	> setki TB
metric	przepustowość (I. transakcji w jednostce czasu)	czas odpowiedzi

⇒ <http://aukcjostat.pl/>

- Allegro: ponad 45 mln. licytacji dziennie

R.Wrembel - Politechnika Poznańska, Instytut Informatyki

6



Aplikacje BI

- **Zapytania ad-hoc (około 10% aplikacji firmowych)**
 - prosty interfejs prezentacji wyników
 - obliczenia ad-hoc
 - drill-down, drill-across
- **Raporty firmowe (około 90% aplikacji firmowych)**
 - zaawansowany układ graficzny
 - biblioteka predefiniowanych raportów
 - subskrypcja raportów, harmonogram odświeżania raportów i ich dystrybucji
 - uprawnienia użytkowników do raportów



Aplikacje BI

- **Dedykowane aplikacje analityczne**
 - analiza przychodów i promocji
 - przewidywanie trendów, symulacje
 - zawierają specjalizowane algorytmy dla dziedziny zastosowań



KPI

- ⇒ A Key Performance Indicator (KPI) is a metric (measure) for evaluating factors that are crucial to the success of an organization wrt the business the organization is doing
- ⇒ Assesses the most important factors for making progress towards declared goals, e.g.
 - yearly profit from sales > 1M EUR
 - # of faulty products < 1%
 - # of satisfied customers > 90%
 - # of cars sold per month > 20
- ⇒ Represented as visual objects (various types of gauges)



KPI

- ⇒ Relates the current value of a metric with a target value





Leading Indicator

- Measures key drivers of business value
- Allow to predict what will happen soon
- Difficult to figure out which LI would be the most appropriate for a given business
- E.g.: LI(customer satisfaction) → KPI(customer retention)



Aplikacje BI

- Dashboard
 - indicates the status at a specific point in time
- (Performance) Scorecard
 - monitors progress of a goal over time
 - answers question: "how are we doing?"
 - includes: KPIs, goals, alerts





Balanced Scorecards

- ⇒ **The balanced scorecard is a management system aimed at translating an organization's strategic goals into a set of performance objectives that, in turn, are measured, monitored and changed if necessary to ensure that the organization's strategic goals are met**



Balanced Score Card

- ⇒ **Performance measurement approach that considers a financial, customer, internal process and learning and growth perspectives**
 - **financial** perspective includes measures such as: operating income, return on capital, and economic value added
 - **customer** perspective includes measures such as: customer satisfaction, customer retention, and market share in target segments
 - **business** process (procurement, production, and order fulfillment) perspective includes measures such as: cost, throughput, and quality
 - **learning & growth** perspective includes measures such as: employee satisfaction, employee retention, skill sets



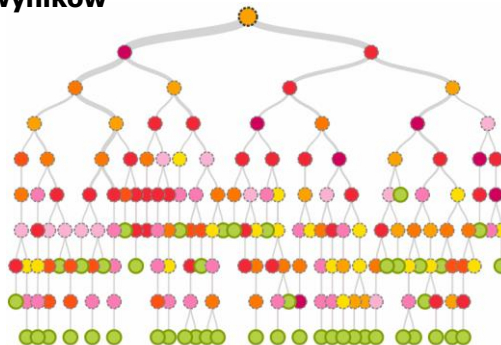
Balanced Score Cards

- Each perspective includes
 - major objectives to be achieved
 - metrics to measure progress toward reaching the objective
 - specific target values for the measures
 - actions to be initiated in order to meet the objectives



Aplikacje BI

- Eksploracja danych
 - złożone obliczeniowo algorytmy
 - dedykowane algorytmy dla dziedziny zastosowań
 - wizualizacja wyników





Użytkownicy

- ⇒ **Aktywni: 10% wszystkich użytkowników systemu BI**
- ⇒ **Równocześnie pracujący: 1% użytkowników systemu BI**



Wielkość systemu

- ⇒ **Mały system HD**
 - HD: kilkaset MB
 - kilkadziesiąt tabel
 - kilka mln rekordów w tabeli faktów
 - 300 użytkowników
 - kilkadziesiąt raportów, kilka kostek
- ⇒ **Duży system HD**
 - HD: kilkaset TB
 - kilkaset tabel
 - kilkaset mln rekordów w tabeli faktów
 - kilka tysięcy użytkowników
 - ponad 1000 raportów, kilkaset kostek



OLAP/BI

⇒ Trends

- **Big Data**
- **mobile BI**
- **in-memory BI**
- **real-time /right-time /active BI**
- **cloud computing**



Systemy komercyjne

⇒ Tradycyjne

- **Oracle11g, Hypersion Essbase - Oracle Corporation**
- **DB2 UDB - IBM**
- **Sybase IQ - Sybase**
- **MS SQL Server - Microsoft**
- **SAP Business Warehouse - SAP**
- **Teradata - Teradata**

⇒ Main memory (in-memory)

- **Netezza - IBM**
- **Exadata - Oracle**
- **SAP Hana - SAP**
- **Teradata DW Appliance - Teradata**



Gartner Report

➔ <http://www.gartner.com/technology/reprints.do?id=1-1DZLPEP&ct=130207&st=sb>

➔ **Assessment criteria**

➔ **Integration**

- BI infrastructure
- Metadata management
- Development tools
- Collaboration

➔ **Information Delivery**

- Reporting
- Dashboards
- Ad hoc query
- Microsoft Office integration
- Mobile BI

➔ **Analysis**

- Online analytical processing (OLAP) - multidimensional analysis, what-if
- Interactive visualization
- Predictive modeling and data mining
- Scorecards- aligning KPIs with a strategic objective
- Prescriptive modeling, simulation and optimization



Gartner Report

