

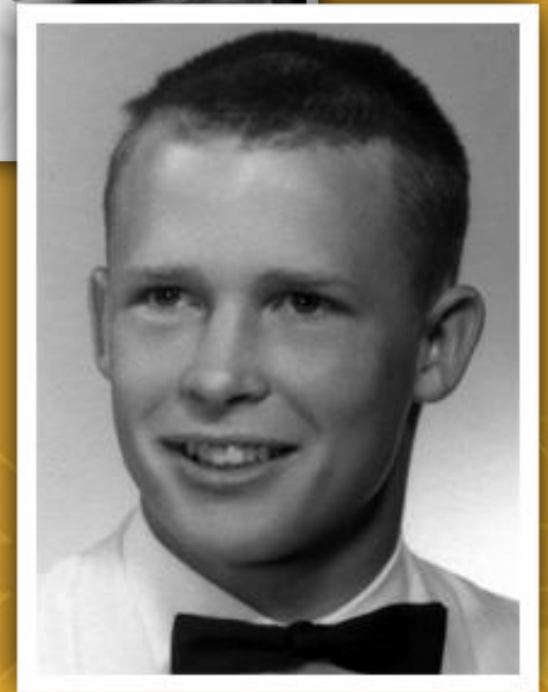
# Database technologies for tomorrow

**Don Haderle**  
**Pat Selinger**

*and Friends*

*Back when  
Relational was  
young...*

*and gas was 23¢  
a gallon BUT disk  
was \$5 a mb.*



# *Our Business: Results, Investments, Strategy*

***Arvind Krishna  
Vice President Data Servers & Development  
IBM Information Management Software***

The IBM logo is displayed in white, bold, sans-serif font on a black rectangular background.

# Our Business is Growing

*IBM Information Management Software*

*#1 Market Share*

*Information*

*Content*

*Data*

*#1 Market Share*

*#2 Market Share*

*Growing Faster Than Market*

**Over 4,500 New Clients  
in the Last 12 Months**

The IBM logo is displayed in white, bold, sans-serif font on a black rectangular background.

# Data Server Innovation

*Milestones in Innovation...Over 3,000 Patents*

**1968**

*First  
Hierarchical  
Data Server*

**1980**

*First IBM  
Relational  
Data Server*

**2006**

*First  
Hybrid  
Data Server*

*Continuous IBM innovation*



# Data Server Innovation

*Defining the Industry.... Setting the Standard*

- System R
- Relational Database
- SQL
- R\* D
- Object
- Para
- DB2
- Rec
- Auto
- Auto
- Mate
- Parallel Database
- Automatic Summary Tables
- Materialized Query Tables
- Index Advisor
- Heterogeneous Database
- Clustering
- er
- er
- on
- er
- XQuery Standard
- pureXML
- Automatic Schema Mapping

**pureXML**  
**2006**  
**Deep Compression**

# Different Needs...Different Servers ...Continuing Investments

*Transactional and analytic applications...*

*Performance and availability characteristics...*

*Mobile devices...Unique features...etc...*

*PHP or Java based web applications...*

# Different Needs...Different Servers ...Continuing Investments





Database  
*Data Servers @ IBM:*  
technologies  
*Cool new Stuff*  
for tomorrow

# IBM

# IDS

# Continued Evolution

2005

**IDS 10**

- **IDS Express**
- **PHP driver**
- **Baan benchmarks**

2006

**OLTP  
Performance**

- **Fastest Informix OLTP engine ever**
- **Enhanced memory management**
- **Enabled for sub-capacity**

2007

**IDS  
"Cheetah"**

- **Location-based data services**
- **Flexible installation footprint**
- **Common Criteria security certification**

**Ease of Development and Operations**



# IBM

# IDS

# Pure OLTP

## Wall Street Equity Company

**All stock quotes,  
every market,  
worldwide**

- **< 1 second lag**
- **Historical and real-time**
- **> 100,000 messages/second**
- **4 CPU, Linux**

streaming updates  
streaming updates



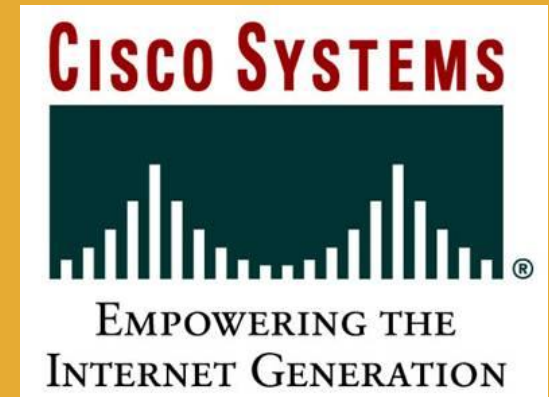
real-time queries  
real-time answers

# IBM

# IDS

# Solution Integration

- **Administration Free, Appliance-style solution**
  - Automated installation, configuration & ongoing management
- **IDS Advantage:**
  - Complete automation
  - 4x faster performance
  - Rich function
  - Ease of development
- **Cisco Solution Expansion**
  - Cisco has 15 products shipping or about to ship with IDS



**IBM**

**IDS**

**What's next?**

*Under Consideration Beyond Cheetah*

- ***Data Compression***
- ***Auto Update Statistics***
- ***Fine Grained Auditing***
- ***Redirected Writes***
- ***Online Table Re-organization***
- ***Range Cluster Index***



# IBM

# IMS

# Continued Evolution

2004

## IMS 9

- **XML database**
- **Web service producer**
- **Online Reorganization**

2006

## SOA Enablement

- **SOAP Gateway**
- **XQuery Preview**

2007



- **More Web Services**
- **XQuery**
- **Greater Parallelism**

**Ease of Development and Operations**

**IBM**

**IMS**

**Version 10**

*Customer Partnership Program: January 2007*

*XQuery for  
Information  
Interchange*

*Enhanced Tooling to  
Ease Development and  
Data Access*

*Web Services for Better  
Business Process  
Integration*



*Dynamic Resource  
Definition  
to Simplify Change*

*Parallel Recon Access  
for High Performance  
and Capacity*

# IBM DB2 System z

## Technical Priorities

- **Unmatched availability**
- **Unmatched scalability**
- **Unmatched reliability**
- **Rock solid security**
- **Support for all applications**
  - **Legacy and new**
- **Efficient data warehousing**
- **Low TCO**

**TD Bank Financial Group**

**10 YEARS**

**Continuous Availability**

**Land Registry**



**23TB**



**World's Largest Known  
Database**

**OLTP**



# IBM DB2 System z

*DB2 9 for System z Available: NOW*

<i>Availability, Scale &amp; Resiliency</i>	More Online Schema Changes Volume Level Backup & Recovery		Business Flexibility Faster, Cheaper, Granular Recovery
<i>Compliance &amp; Security</i>	Database Roles & Trusted Context		Efficient Auditing & Compliance
<i>Easier Application Development</i>	PureXML		Streamlined Data Integration
<i>OLTP &amp; Warehousing</i>	Query Optimization		Highly Available, Secure Data
<i>Reduce Cost of Ownership &amp; Skills</i>	Index Compression		50% Disk Savings

# Bank of China

IBM DB2

SYSTEM Z

## *FNS BANCS Application on IBM Mainframe Redefines Transaction Processing Performance in Banking Industry*

**SYDNEY, AUSTRALIA and ARMONK, NY--(MARKET WIRE)--Feb 8, 2007**



中國銀行

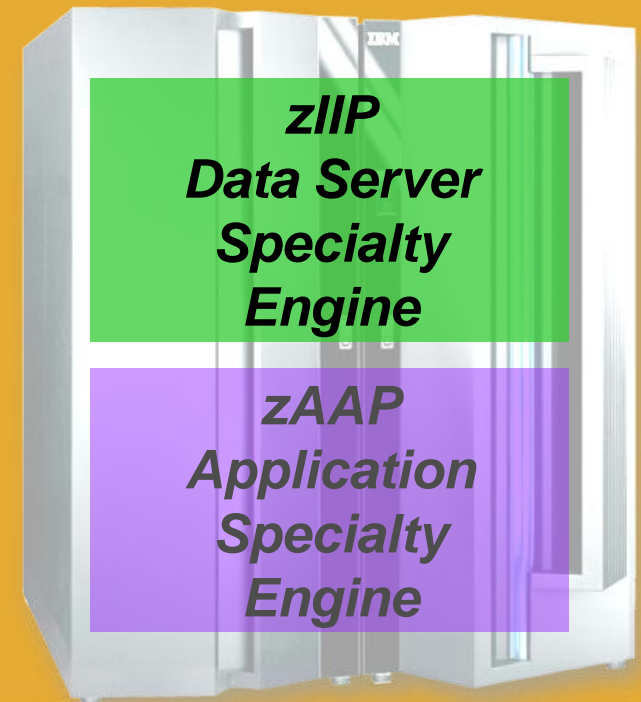
BANK OF CHINA

- **380 million** client accounts
- **34 million** transactions per hour
- **9,445** transactions per second
  - Original goal: **4,100 TPS**

# IBM DB2 System z

## *Specialty Engines Optimize Cost*

- ***Significant value, available now***
  - Improved response times
  - Improved TCO
- ***Multiple workloads benefit***
  - SAP & other DRDA workloads
  - SAP BW & other warehouses
  - DB2 Utilities
  - DB2 9 remote native SQL stored procedures
- ***Strong customer adoption***



**IBM**

**DB2**

**Linux, Unix, Windows**

*Directions*



- ***High Performance & Scalability***
- ***Reduced TCO & Accelerated Time-to-value***
- ***Resiliency through High Availability & Security***

# IBM

# DB2

# Linux, Unix, Windows

## *What's Coming in Performance Leadership?*

- **Go Faster...**
  - Faster engine
  - Faster utilities
    - Relational and XML
  - Consistent query performance
- **Scale Further...**
  - Faster redistribute
  - Manage workloads
  - Expanded options for IBM Balanced Warehouse



*“With DB2 9, we've seen a five-to-10 times performance improvement over a non-database environment.”*

CEO, Skytide

# IBM

# DB2

# Linux, Unix, Windows

## *What's Coming in Hands-off Administration?*

- ***Prior Releases: Management-by-exception***
  - Notifies on unhealthy conditions
  - Advisors
- ***V9: Up-and-running***
  - Install, Provide e-mail / pager, “db2start”
  - Small to mid-size businesses
- ***Future: Continue to automate...***
  - Simplify upgrades
  - More automated storage
  - Integration with flash copy
  - Backup automation
  - “Just in Time” statistics
  - Auto-create compression dictionary



# IBM **DB2** Express-C

**New**

*12 Months License & Subscription*

## Do It Yourself

- **Free** to build, deploy, distribute
  - pureXML included
  - Community assistance
- **On conference DVD**

## ...or Let IBM Help

- 24\*7 support, FixPaks
- SQL Replication
- High Availability
- ... at a price that is competitive with open source data servers

Download or learn more at [www.ibm.com/db2/express](http://www.ibm.com/db2/express)

Database  
*great news for*  
technologies  
*app developers*  
for tomorrow

*: - )*



# IBM

# Data Server App. Dev.

## Leveraging Technologies and Communities

- **Exploiting New Converged SQL**

- **Key Database Technologies**

- SQL, SQL Procedures
- XML
- SOA, Web Services

- **Developer Communities**

- COBOL, PL/I, REXX, C, C++
- Java (JDBC / SQLJ)
- .NET (C#, VB .NET)
- Open Source
  - PHP
  - Perl
  - Python
  - Ruby on Rails



**Database**  
*Dynamic Warehousing:*  
**technologies**  
*Extending Warehouses*  
**for tomorrow**  
*to Create*  
**New Business Value**

# Dynamic Warehousing

*Every Person, Every Transaction, Every Asset*

## Information On Demand

Optimize Each Transaction  
as It Happens



*Dynamic  
Warehousing*

## OLAP & Data Mining

Recommend  
Future Action



*Traditional  
Data  
Warehousing*

## Query & Reporting

Understand What  
Happened



# Dynamic Warehousing

*Every Person, Every Transaction, Every Asset*

## Information On Demand

Optimize Each Transaction  
as It Happens



*Dynamic  
Warehousing*

**OLAP**

### *Dynamic Warehousing Attributes:*

- 1. Real-time access – in context*
- 2. Analytics – part of a business process*
- 3. Unstructured information – extracted knowledge*
- 4. Extended database capabilities*

**Query**

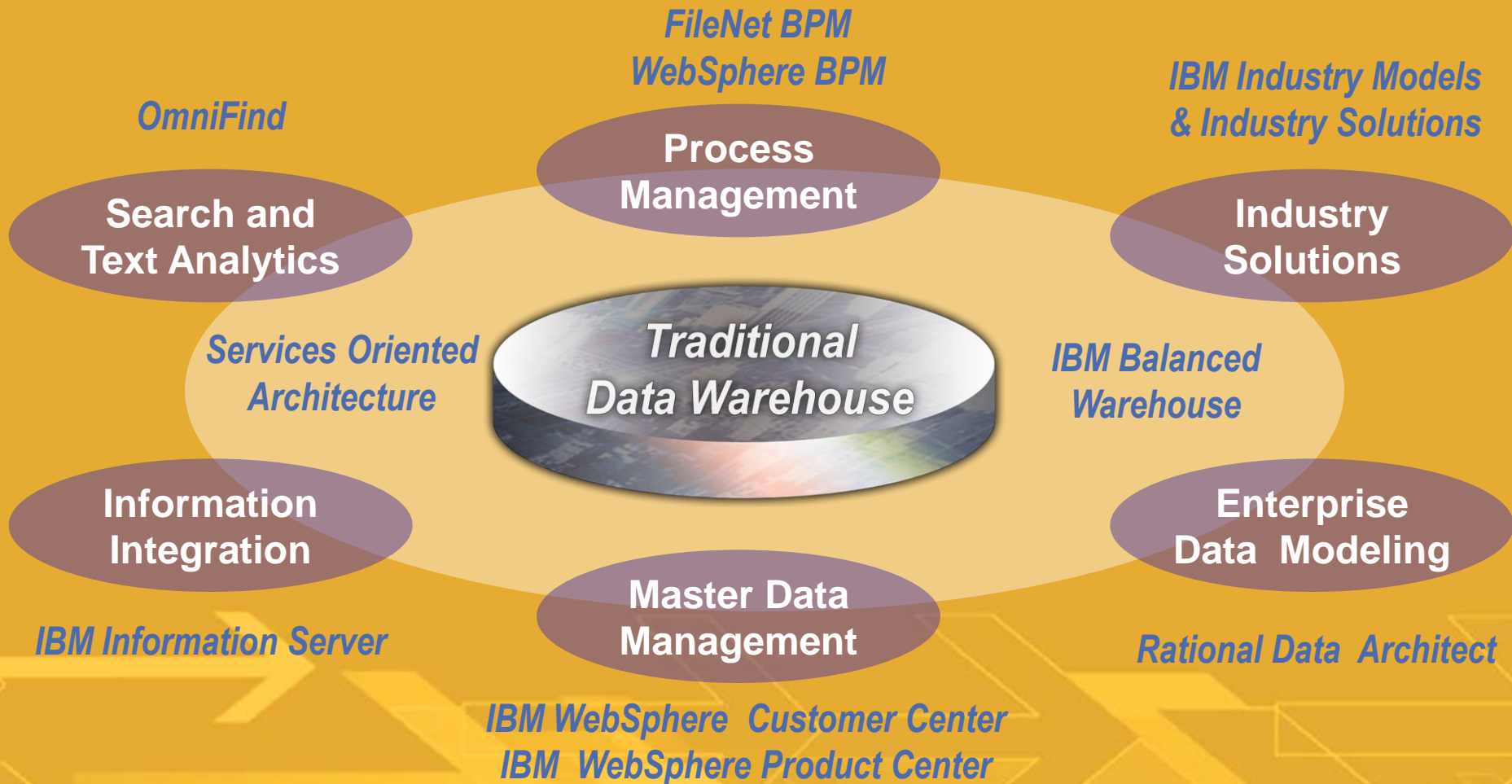
Under

Happened



# Dynamic Warehousing

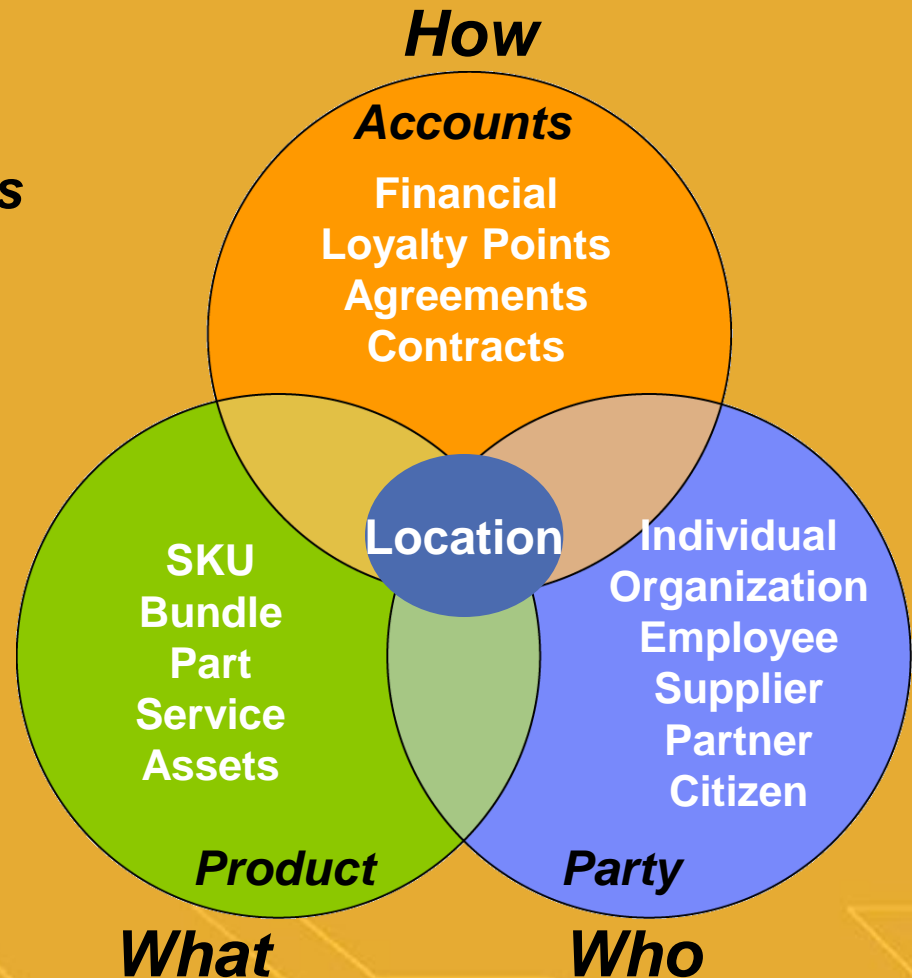
## *Related IBM Offerings*



# Database technologies *Information* for tomorrow

# Master Data

- **Describes Core Business Entities**
  - Customers, suppliers, partners,
  - Products, materials,
  - Accounts, location and employees
- **High Value Information**
  - Used repeatedly
  - Across many business processes
  - Agreed common values and views
  - Decisive Data
    - Currency, quality



# Master Data Management

- **Ensure Consistent Master Data**

- Across transactional & analytical systems

- **Address Key Issues Proactively**

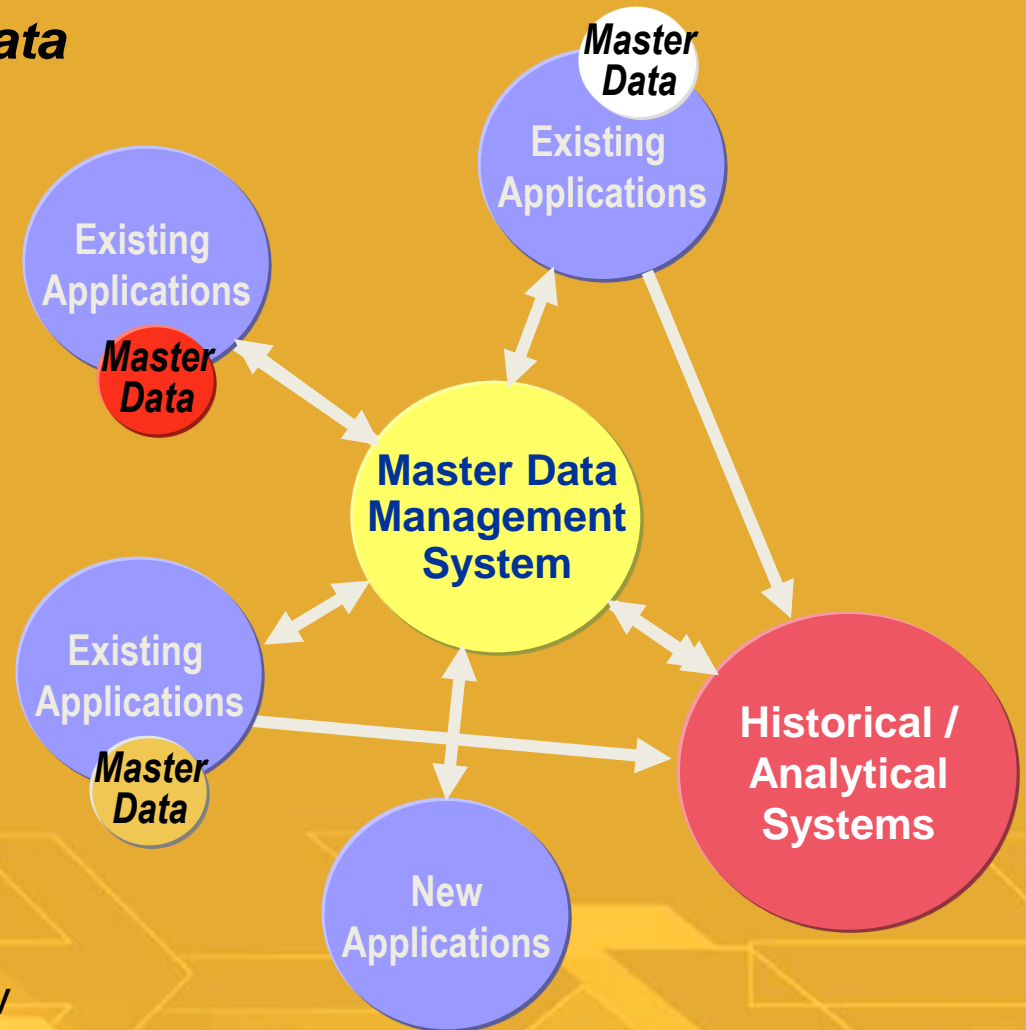
- Data quality
- Data consistency

- **Decouple Master Data from Individual Applications**

- Application independent resource

- **Simplify Ongoing Tasks**

- Integration
- New application development





# IBM

# MDM Strategy

## 1. Provide Multi-form MDM Platform

- Collaborative, Operational, Analytical
- Multiple Domains
  - Customer, Product, Supplier, Location, Account...

## 2. Manage the Full Data Lifecycle

- Proactively detects and responds to events within the data lifecycle process
- Leverages existing investments

## 3. Leverage IBM's Technology Portfolio, Industry Models and Industry Expertise



*IBM is the most experienced MDM provider*

# Multiple Forms of MDM

- Time to market
- Diverse product mix
- Collaboration

- System of record
- Fast transactions
- Threat & fraud detection
- Conflict management

- Quality of master data feeding analysis
- Compliance



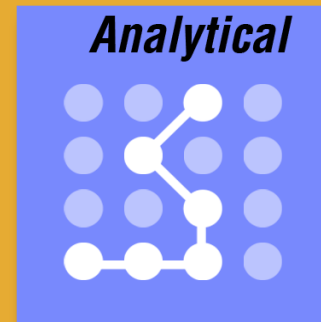
## ***Define, Collaborate, Publish***

- Define, enrich, and publish definitional data
- Workflow and advanced security capabilities
- Trading partner communication



## ***Create, Transact, Track***

- Single source
- Flexible Architecture
- Historical data
- “In-line” Analytics
- Performance



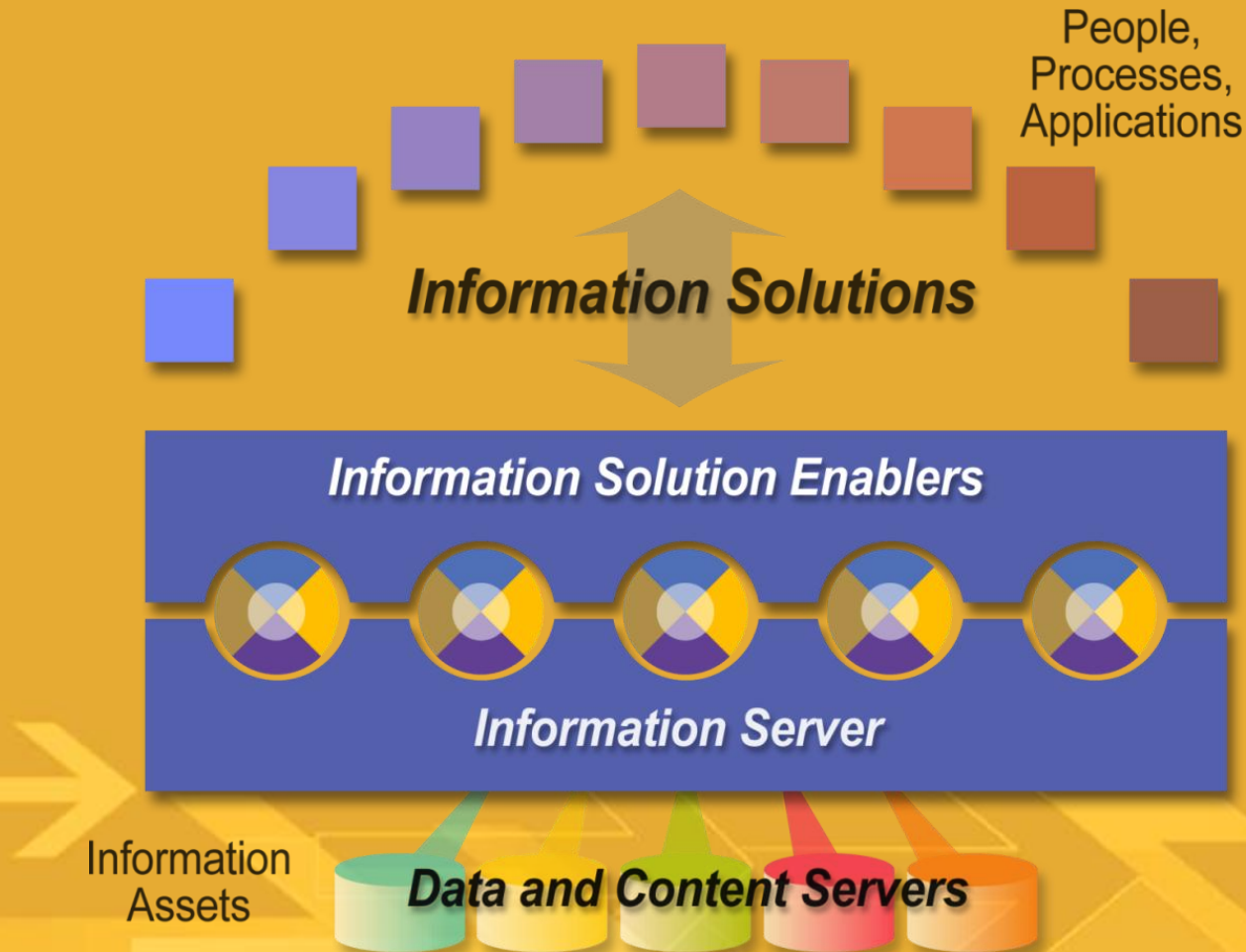
## ***Aggregate, Analyze, Act***

- Persist results from analytical environments
- Feed master data to warehouse
- Simple domain analytics



# Master Data Management

*Open Standards, Flexible Architecture, Metadata Driven*





# Information Server

**Business Users**



**Subject Matter Experts**



**Architects**



**Data Analysts**



**Developers**



**Database Administrators**



## *Role Based Modules that Support the Data Lifecycle*

Unified Deployment

*Understand*



*Cleanse*



*Transform*

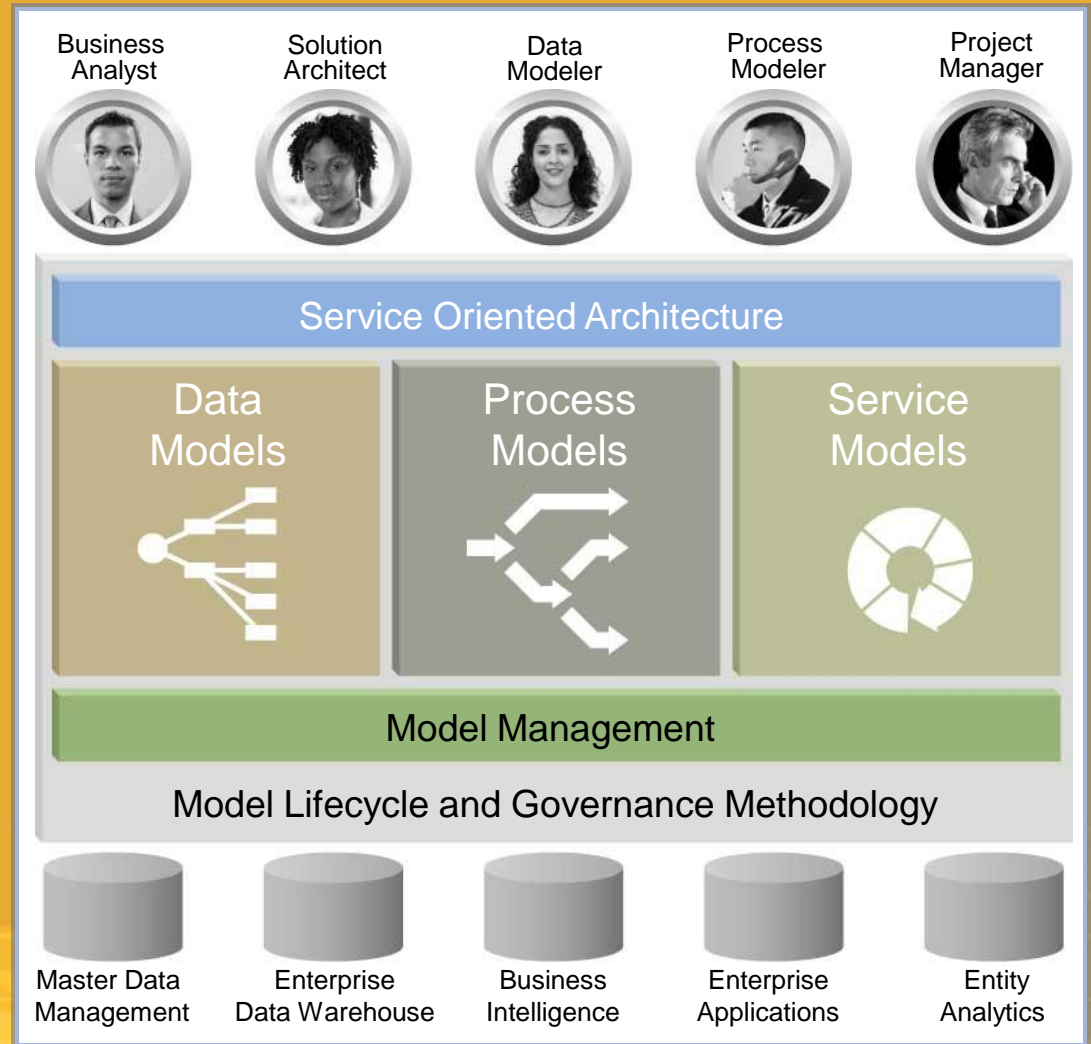


*Collaborate*



Unified Metadata Management

- **Enables business users to easily scope and customize their own requirements**
- **Facilitates step-by-step business focused development and roll-out**
- **Delivers regularly updated business, technical and regulatory content**
- **Creates open technology platform for any application or integration solution**
- **Manages definitions and standards in complex IT environments**



## Usable Content for Business-driven Development

### Banking



- Profitability
- Relationship Marketing
- Risk Management
- Asset and Liability Management
- Compliance

### Financial Markets



- Risk Management
- Asset and Liability Management
- Compliance

### Health Plans



- Claims
- Medical Management
- Provider and Network
- Sales, Marketing and Membership
- Financials

### Insurance



- Customer centricity
- Claims
- Intermediary Performance
- Compliance
- Risk Management

### Retail



- Customer centricity
- Merchandising Management
- Store Operations & Product Mgmt.
- Supply Chain Management
- Compliance

### Telecommunications

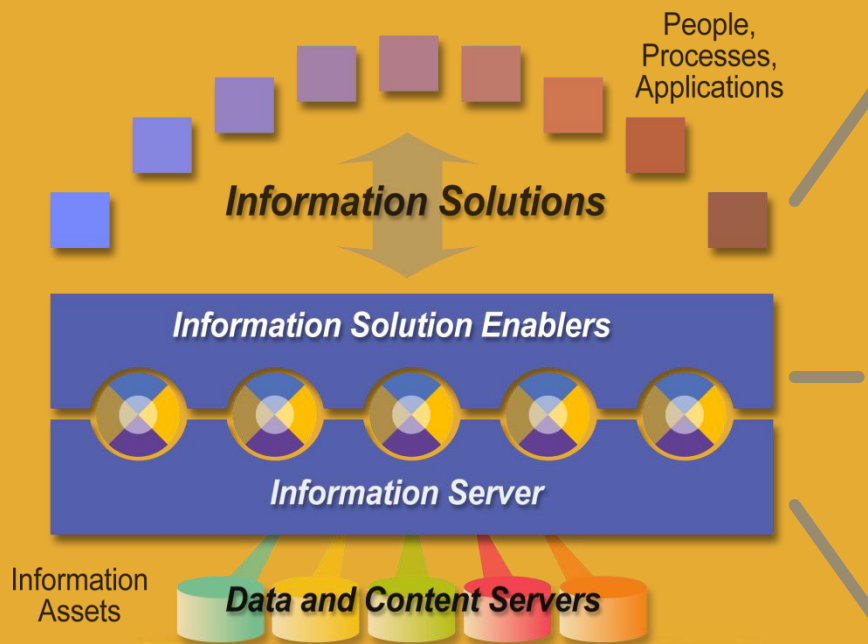


- Churn Management
- Relationship Mgmt & Segmentation
- Sales and Marketing
- Service Quality & Product Lifecycle
- Usage Profile

# IBM

# Information On Demand

*Tying it all together*



**Information Solutions**  
Create new business value and innovation from information

**Master Data Management & Industry Models**  
Define and synchronize critical information for operations and business insight

**Information Server**  
Understand, Cleanse, Transform, Deliver; Metadata Driven

Database  
*hot topics:*  
innovative THINKing  
at IBM Research  
for tomorrow





Database  
*software in action:*  
technologies  
*search & discovery*  
for tomorrow

# Integration of Unstructured Data Will Move BI to the Next Level

*Detailed analysis of data often leads to non-obvious  
conclusions that go against conventional wisdom...*

*“Freakonomics” by Levitt & Dubner*

# Common Real-Estate Terms from Ads

*Fantastic*

*Granite*

*Spacious*

*State-of-the-art*



*Corian*

*Charming*

*Maple*

*Great Neighborhood*

*Gourmet*

***Which ones correlate with higher final sales price?***

# Common Real-Estate Terms from Ads

~~Fantastic~~

Granite

~~Spacious~~

State-of-the-art



Corian

~~Charming~~

Maple

~~Great Neighborhood~~

Gourmet

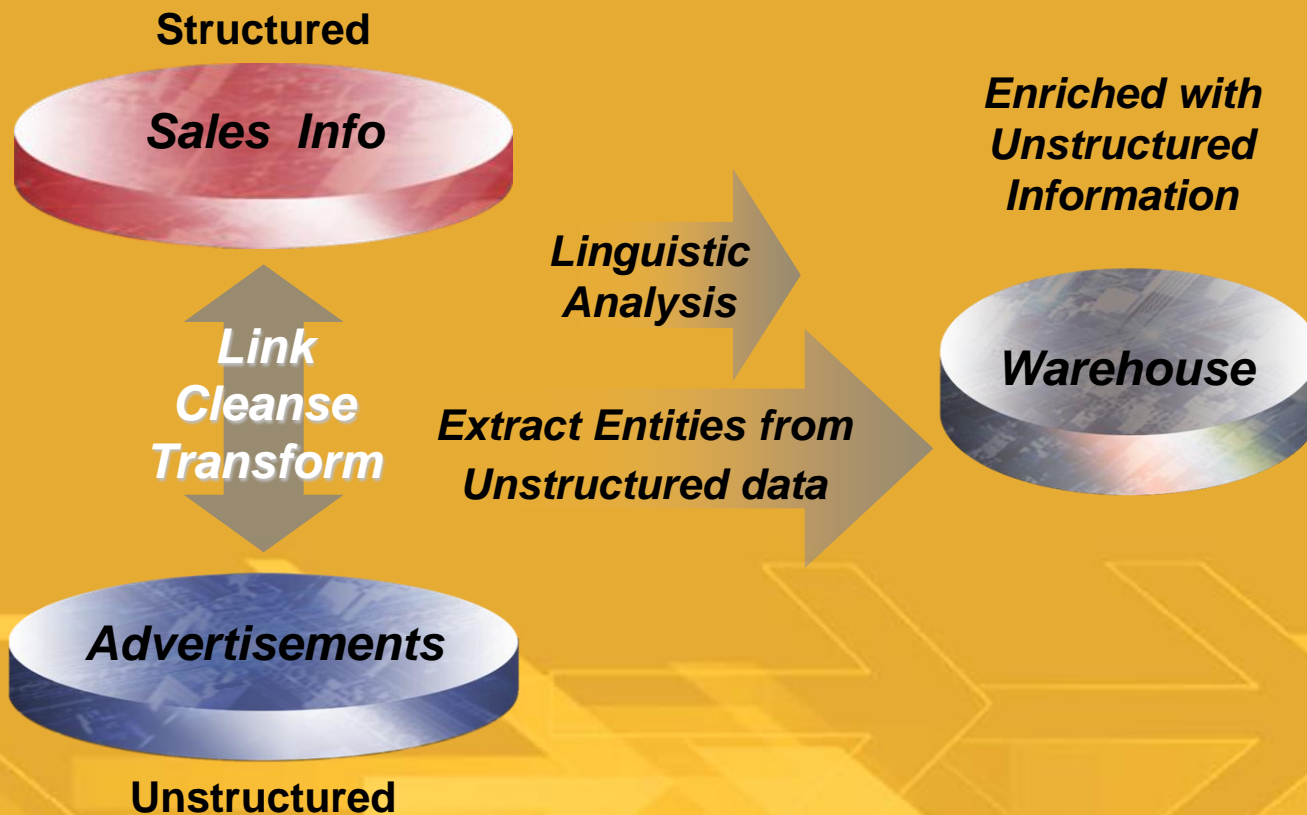
***Some words do correlate with higher price.***

***Some are best avoided...***

***and are when real estate agents sell their own homes...***

# Omnifind Analytics Edition

*Augments Dynamic Warehouse with the ability to correlate Unstructured Data (e.g. Ads) with Structured Data (e.g. Sales data)*



**Visualization & Interactive Mining**



- ***Internal IBM IT Helpdesk for Employees***

- ***Each ticket***

- Unstructured text            problem description, resolution...
    - Structured attributes        date, id, type, duration....

- ***3 Million Tickets Loaded into OmniFind***

- ➔ ***What are the most common problems?***
  - ➔ ***Which issues take a long time to solve?***
  - ➔ ***What are the issues that are specific to DB2?***

Vertical category: Noun Sequence

List:  Subcategories  Keywords

Sort:  Frequency  Correlation  Alphabet

Max lines to display:  20  50  100  200  500

1) Mail related

2) Password reset

3) Voice mail

Search	Keywords	Frequency		
<input type="checkbox"/>	tec server	56035		1.0
<input type="checkbox"/>	Users Affected	48492		1.0
<input type="checkbox"/>	Workaround Available	48255		1.0
<input type="checkbox"/>	lotus Notes	44123		1.0
<input type="checkbox"/>	password reset	42796		1.0
<input type="checkbox"/>	error message	36648		1.0
<input type="checkbox"/>	user id	24062		1.0
<input type="checkbox"/>	Applications What version	23479		1.0
<input type="checkbox"/>	server copy	21719		1.0
<input type="checkbox"/>	Windows XP	21317		1.0
<input type="checkbox"/>	mail file	21292		1.0
<input type="checkbox"/>	da URA D	18689		1.0
<input type="checkbox"/>	system ref	17636		1.0
<input type="checkbox"/>	voice mail	16843		1.0

# 2D View Allows Structured and Unstructured Dimensions to be Correlated

SEARCH CONDITION Searched documents: 13659 documents

AND 13659 docs EXCEPT KEYWORD: automatically 2914786 docs

KEYWORD: DB2 restrict analysis to only cases that mention DB2

VIEWS 2D Map

Vertical category: Noun Sequence  
List: Keywords  
Sort: Frequency  
Max lines to display: 10

Horizontal category: DURATION  
List: Keywords  
Sort: Alphabet  
Max lines to display: 10

Structured

Unstructured

If DBA appears in text case takes longer

Subcategories/keywords		00:00:00 1023	00:00:00:10 717	00:00:00:15 590	00:00:00:20 455	00:00:00:25 1721	00:00:00:30 1930	00:00:00:35 596	00:00:00:40 281	00:01:00:00 380
tec server 723	1 0.0	137 2.1	73 1.4	52 1.2	31 0.8	89 0.7	136 1.1	18 0.3	3 0.0	17 0.5
shift T 512	1 0.0	10 0.1	19 0.4	15 0.4	14 0.4	106 1.3	198 2.3	10 0.2	6 0.2	3 0.0
error message 508	50 0.5	78 1.5	17 0.3	19 0.5	19 0.6	66 0.8	73 0.8	15 0.4	9 0.4	24 1.0
CLI driver 469	41 0.4	81 1.7	17 0.4	10 0.2	17 0.6	22 0.2	53 0.6	32 1.0	16 0.9	35 1.8
system ref 408	0 0.0	30 0.6	52 1.7	75 3.2	31 1.5	26 0.3	59 0.7	40 1.5	16 1.0	1 0.0
DBA DB2 STF 308	0 0.0	1 0.0	16 0.5	15 0.6	28 1.7	105 2.1	93 1.7	6 0.1	1 0.0	0 0.0
work station 283	2 0.0	17 0.4	22 0.9	8 0.2	8 0.3	17 0.3	17 0.2	6 0.1	10 0.8	4 0.1



# Deriving Business Value

Agent Can't Handle Problem... Needs a DBA...Delay  
*Could Better Training Help?*

Title	X1@@ / PMIDB2 / LOADJV / DATLD / RC=U0044
Text	<p><b>.DESCRIPTION:</b> ** TIMELINE ** 23.54.12 JOB16523 -LOADJV DATLD U0044 23:54 - Job LOADJV from application PMIDB2 abended on step DATLD with RC=U0044. 00:07 - Called Susan Puckett. She said that, for these DBA jobs we should call for <b>DBA DB2</b> instead of her. She said though for Ops to restart this one job from top once, and if it fails, then call for DBA. 00:16 - Job restarted. 00:21 - Job ended. BET Ops 3rd shift</p> <p><b>.PROBLEM_RESULT:</b> Job was restarted,</p>
Detail	<a href="#">DETAIL</a>
Standard Date	20060202

Title	EME/R1EIMD01/JFBSC0/RC=08/ICCMVS4/IGF
Text	<p><b>.DESCRIPTION:</b> 23:03 - Job R1EIMD01 from application EMEAR1EDLY abended on step JFBSC0 with return code 08 on MVS4. 23:05 - OPS contacted <b>DBA DB2</b> on-call Vishnu D Pavor. 23:25 - Vishnu refreshed from EMG and restarted. It abended. 23:29 - Vishnu refreshed from EMG and restarted the job again. It ran fine. Patrick Wiens 3rd Shift OPS</p> <p><b>.PROBLEM_RESULT:</b> Cleaned up conflicting item, restarted job. Job ran to completion.</p>

# **IBM** Research

## *More Goodies...*

- ***Visualization & Text Analytics***
  - *Semantic search*
- ***Information Mashups***
- ***Metadata discovery, analysis, integration & understanding***
- ***Business Integrity; Risk & Compliance***
- ***Extreme Compression***



Database

*Thank You!*

for tomorrow