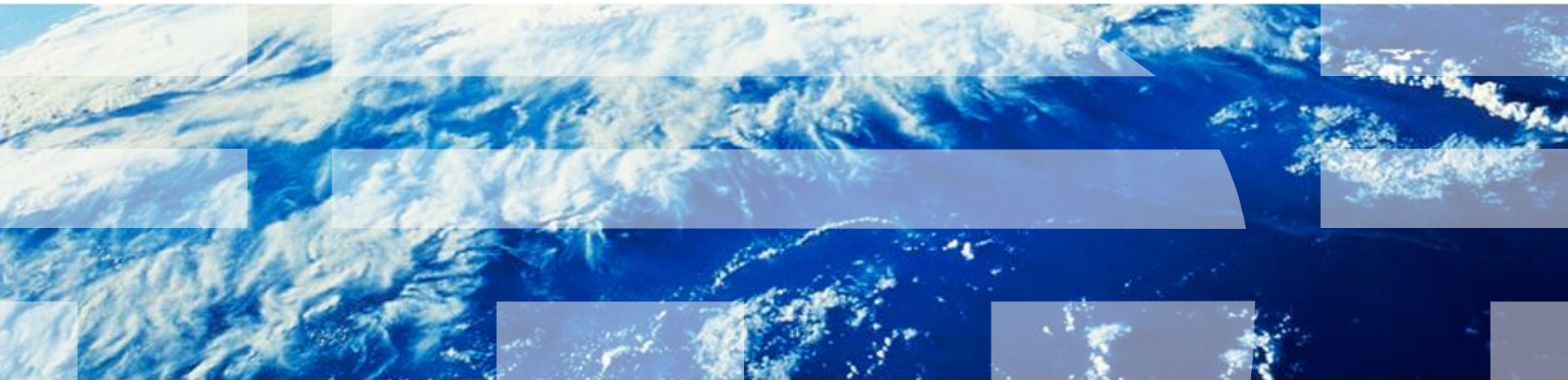


The Evolving Data Warehouse

Budapest DW Forum 2013



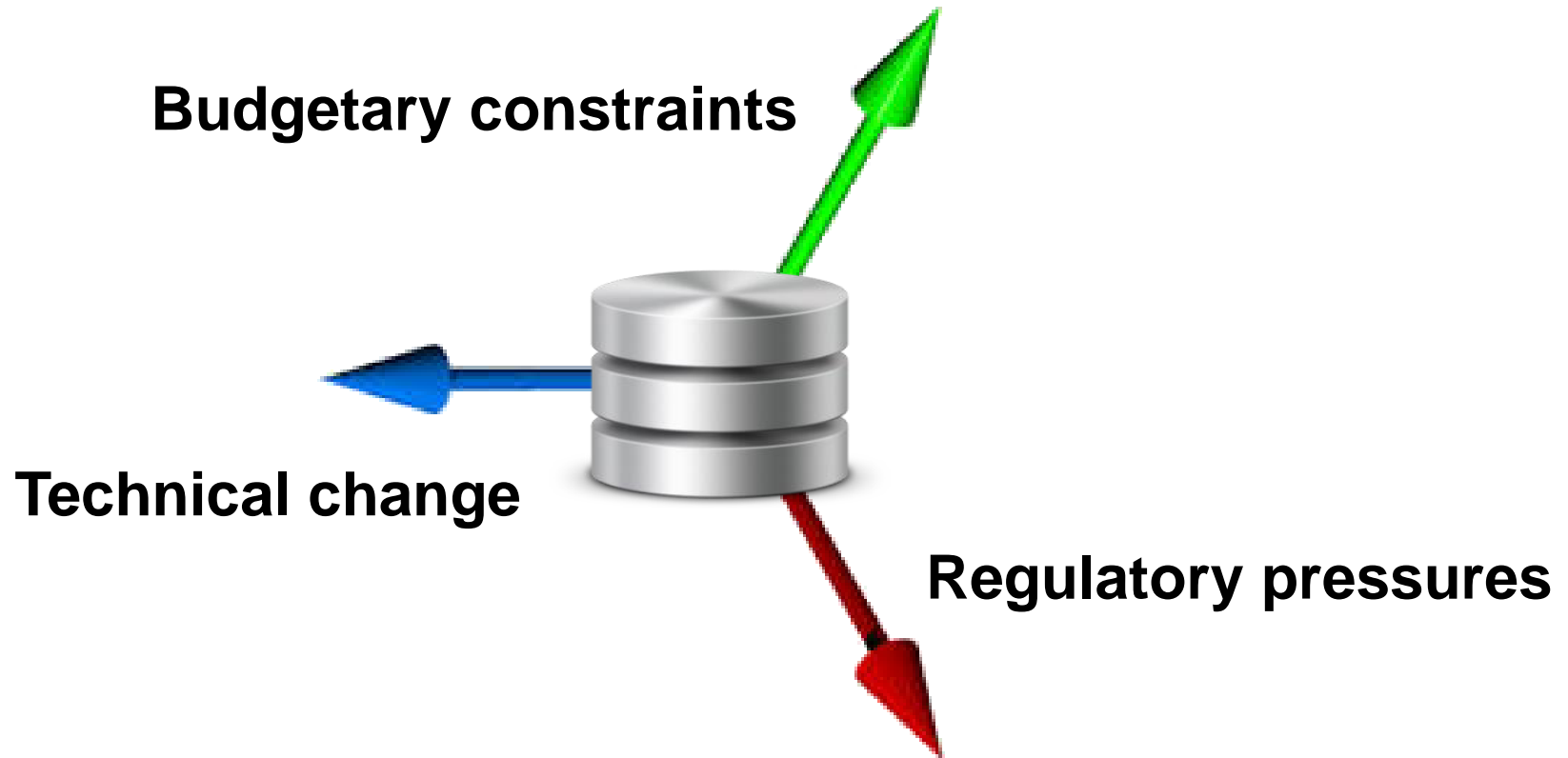
Dirk deRoos

dderoos@ca.ibm.com

 [@Dirk_deRoos](https://twitter.com/Dirk_deRoos)

IBM World-Wide Technical Sales Leader, Big Data

Pressures on the Traditional Warehouse



The New Means of Production

- Land
- Labor
- Capital
- Cloud
- Analytics
- Data

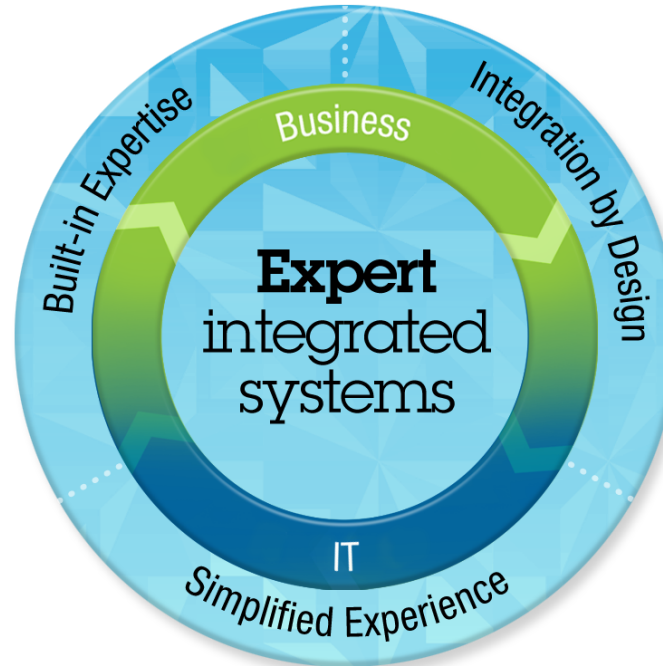


Appliance-Based Warehouse Solutions

PureSystems™

Built-in Expertise

Capturing and automating what experts do – from the infrastructure patterns to the application patterns



Integration by Design

Deeply integrating and tuning hardware and software – in a ready-to-go workload optimized system

Simplified Experience

Making every part of the IT lifecycle easier – with integrated management of the entire system and a broad open ecosystem of optimized solutions

Fit-For-Purpose Architectures

Meeting Big Data Challenges – Fast and Easy!



PureData

System for Transactions

For apps like E-commerce...

Database cluster services optimized for transactional throughput and scalability

PureData

System for Analytics

For apps like Customer Analysis...

Data warehouse services optimized for high-speed, peta-scale analytics and simplicity

Powered by Netezza technology

PureData

System for Operational Analytics

For apps like Real-time Fraud Detection...

Operational data warehouse services optimized to balance high performance analytics and real-time operational throughput

The NoSQL Revolution

- **Different requirements require different tools**
 - Document stores
 - Key/value stores
 - Google BigTable implementations
 - Graph databases
- **Values (there are exceptions)**
 - Huge data volumes – easy scale-out
 - Semi-structured data
 - Extreme performance



Traditional Warehousing vs. NoSQL

ACID vs. BASE

- **Atomicity**
 - **Consistency**
 - **Isolation**
 - **Durability**
- **Basically Available**
 - **Soft state**
 - **Eventually consistent**



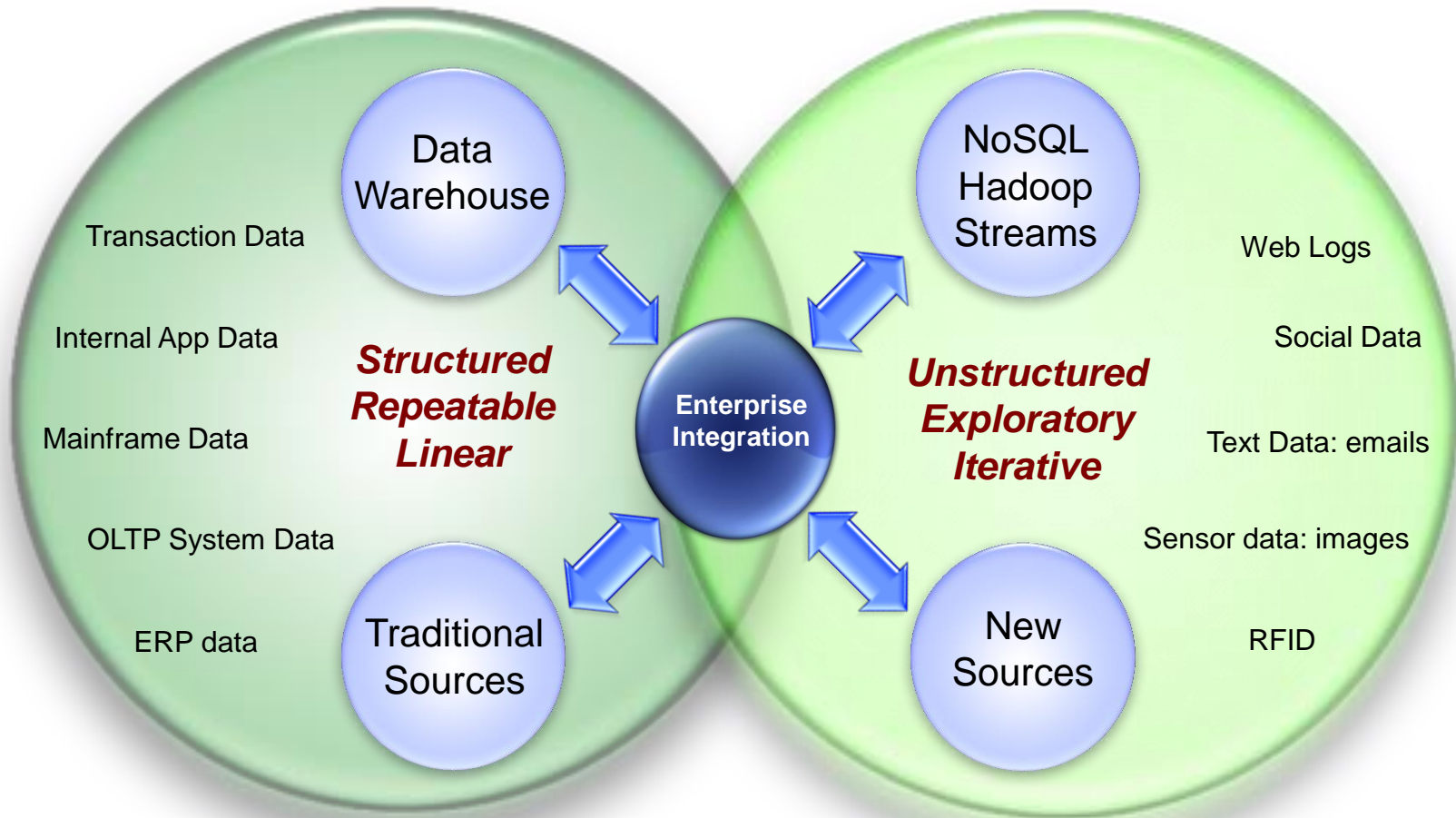
Complementary Analytics

Traditional Approach

Structured, analytical, logical

New Approach

Creative, holistic thought, intuition



Traditional Data Mining and Exploratory Analysis



IBM DB2 as a Graph Store

- **Resource Description Framework support**

- Triple store – ie. subject, predicate, object

- **SPARQL support**

- SPARQL Protocol and RDF Query Language

- **Example:**

"Find all companies that sell a product to a supplier"

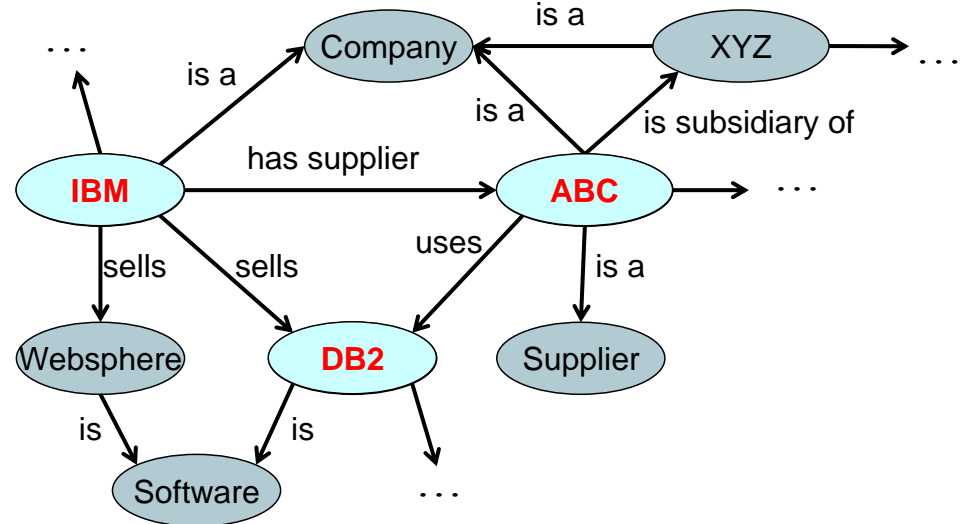
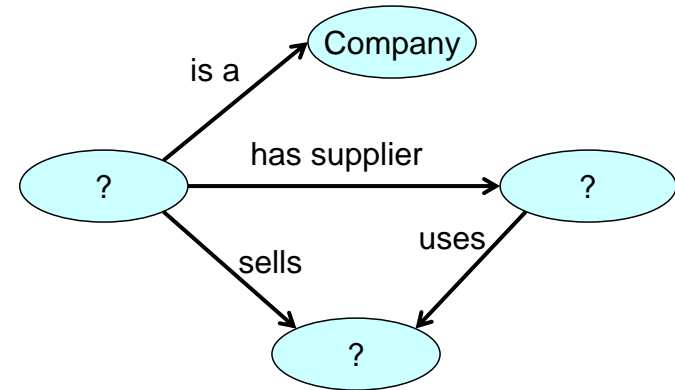
```

SELECT ?comp, ?product, ?supplier
WHERE {
  ?comp <isA> <Company>
  ?comp <sells> ?product
  ?comp <hasSupplier> ?supplier
  ?supplier <uses> ?product
}
    
```

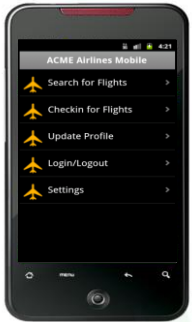
- **Result:**

```

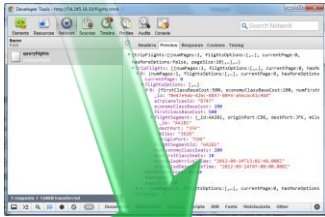
?comp – IBM
?product – DB2
?supplier – ABC
    
```



IBM DB2 JSON Data Store and MongoDB Wire Listener



JSON
JavaScript Object Notation



- **Interoperate seamlessly with modern applications**
 - Flexible schemas allow rapid delivery of applications

- **Preserve traditional DBMS capabilities, leverage existing skills and tools**
 - Multi-statement transactions
 - Management / operations
 - Security
 - Scale, performance and high availability



Data Governance Maturity Disciplines

- **Organizational awareness**
- **Stewardship**
- **Policy**
- **Value creation**
- **Data risk management**
- **Security/Privacy/Compliance**
- **Data architecture**
- **Data quality**
- **Business glossary/metadata**
- **Information lifecycle management**
- **Audit and reporting**



Data Governance Maturity Disciplines

NoSQL Challenges

- Organizational awareness
- Stewardship
- Policy
- Value creation
- Data risk management
- **Security/Privacy/Compliance**
- Data architecture
- **Data quality**
- **Business glossary/metadata**
- **Information lifecycle management**
- **Audit and reporting**



Taming NoSQL: Make it Speak Your Language

- **Native SQL access to NoSQL is the Holy Grail**
 - Every major Hadoop vendor is working on this
 - Focus today is performance
 - Future benefit is some governance capabilities
- **Lens to potentially messy data**
- **With native SQL access, you can apply many governance techniques and tools**



THINK

BIG

BIG