

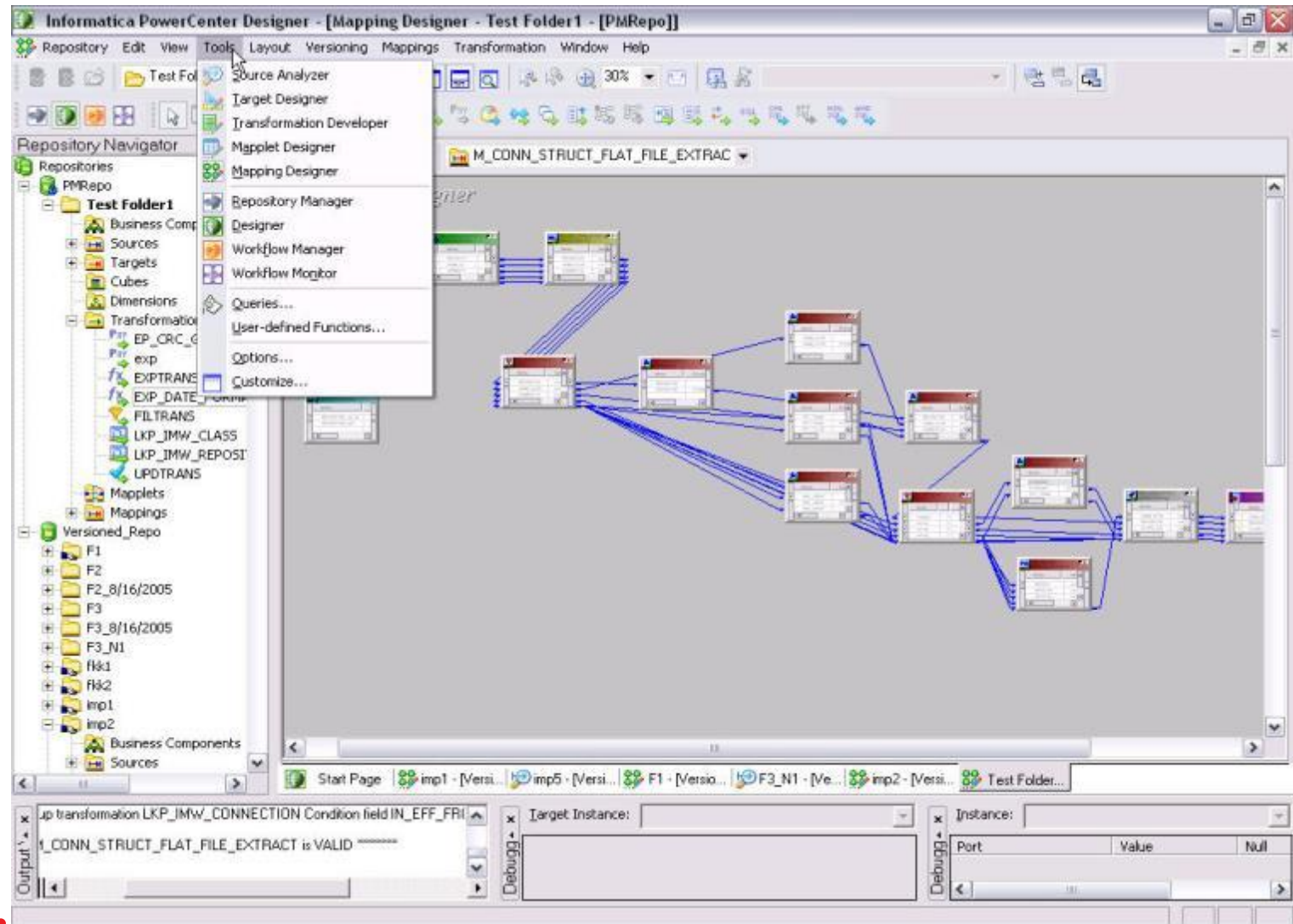


企业级数据交换中心分享 **Data Integration Hub (DIH)**

Eddie Hui

Principal Sales Consultant

传统ETL工具



Data Integration Hub

Gartner建议的数据交换模式

Gartner

Research

Publication Date: 6 June 2011

ID Number: G00213541

Data Integration Hubs: Drivers, Benefits and Challenges of an Increasingly Popular Implementation Approach

Ted Friedman

Point-to-point
costs, crea
managem
integration
integration

Data Integration Hubs Grow in Popularity

As an alternative to point-to-point and project-specific data integration development and deployment, an increasing number of organizations are moving to a delivery approach in which data integration processes are managed in a cross-project/application fashion (see Figure 1). Although starting from a small base, Gartner noted an increase of more than 200% in client inquiries on this topic from 2009 to 2010. Sometimes referred to as a "data integration hub," in reference to a "hub and spoke" architecture, this approach is being adopted to address the challenge including integration (hub) and

利用一个中央监控平台和数据模型

实现多个数据源对多个目标系统的数据整合

Data Integration Hub

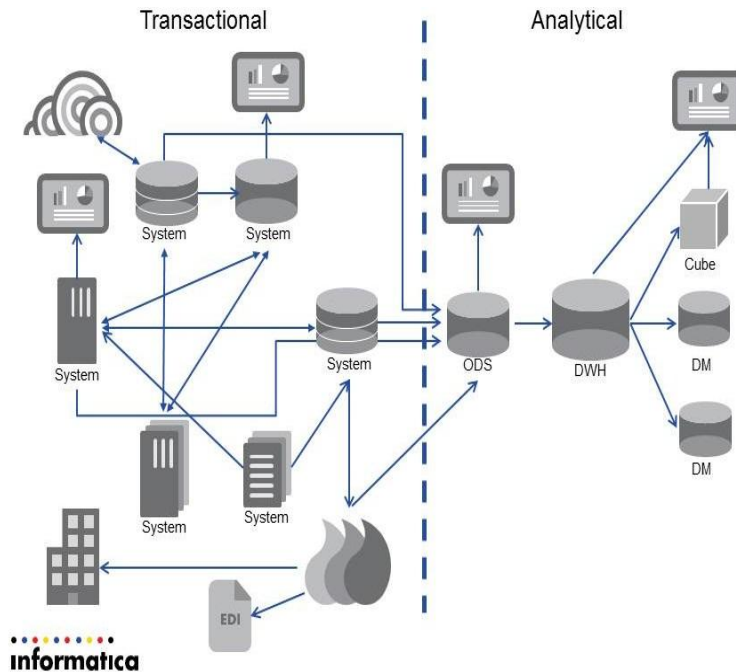
Application to Application Integration

Turning

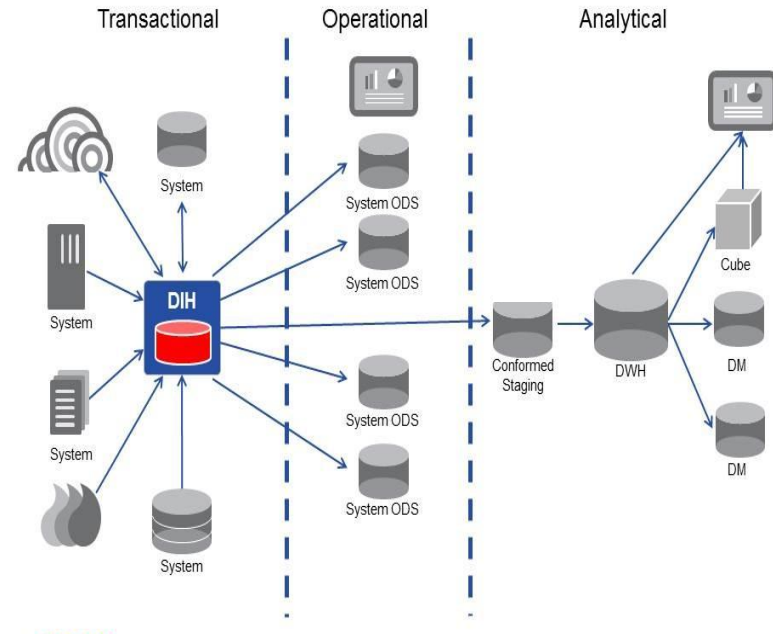
Chaos...

Into

Order...

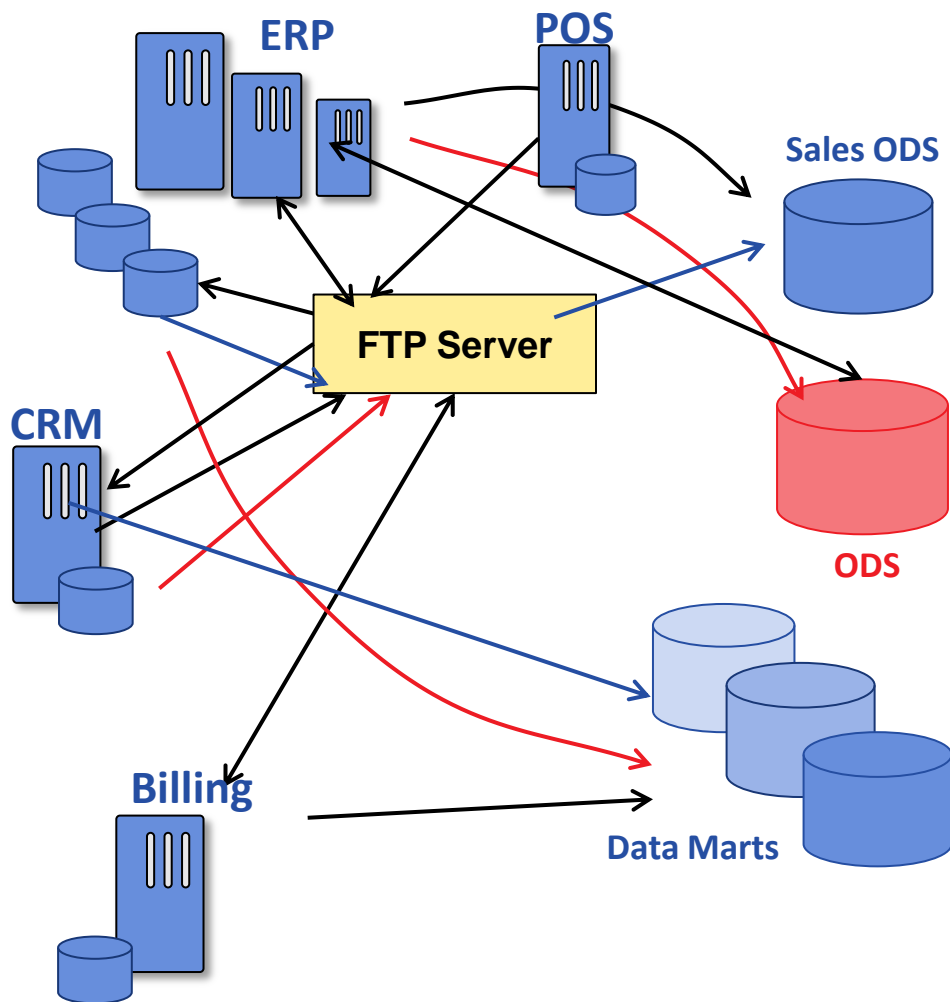


ETL - 点对点模式



DIH - 发布/订阅模式

传统方案



Complex Transformation



XML / JSON



Flat Files



Logs



Industry Standards



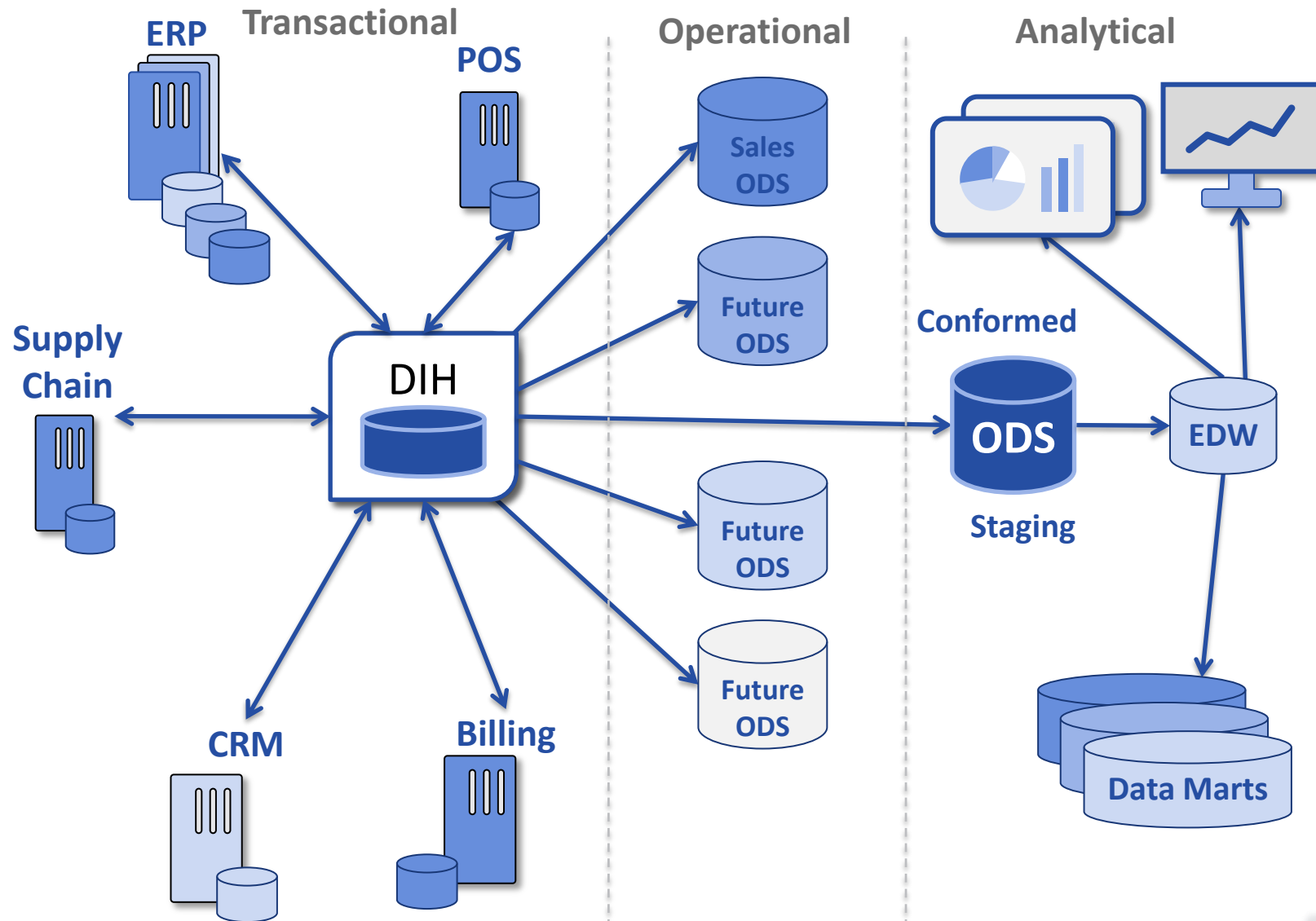
费时

昂贵



DIH数据传输架构

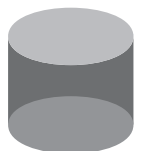
Desired End-state



End-to-End Metadata Visibility

Data Integration Hub

发布

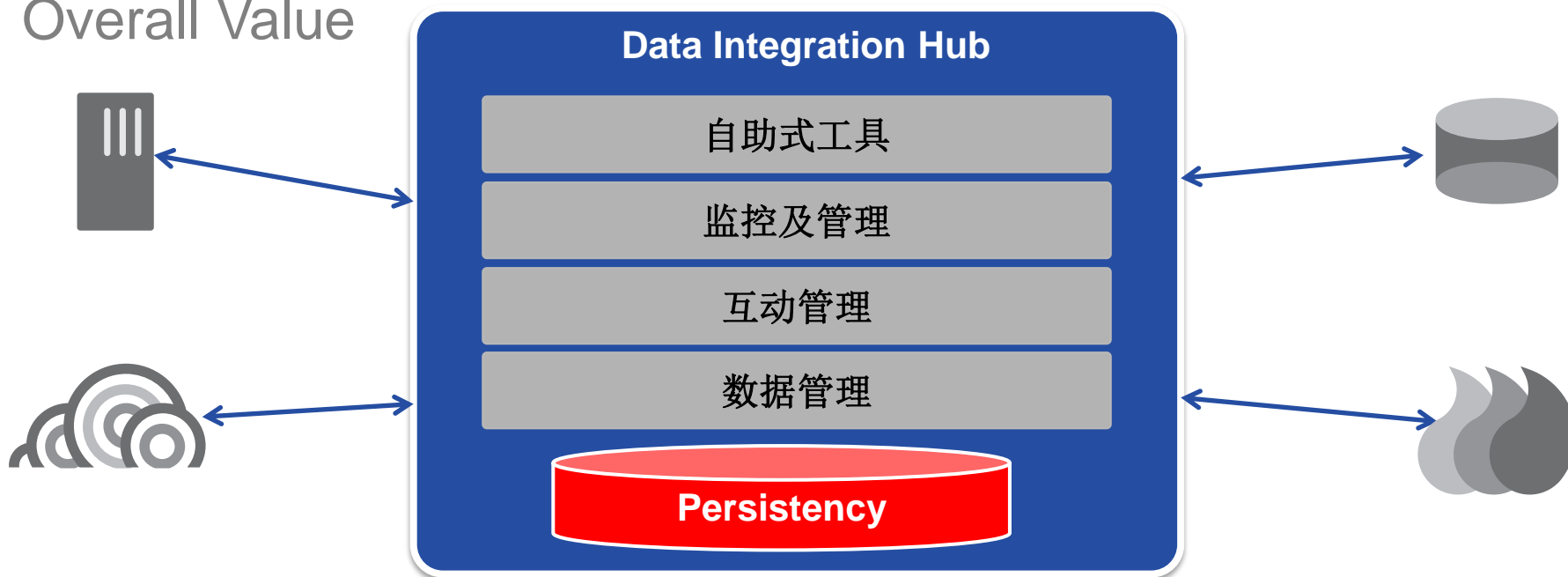


订阅



Data Integration Hub Capabilities

Overall Value



Capability	Value
发布/订阅模式数据交换	减少系统之间的依赖性
自助式数据持久性管理	简易储存及归档
自动生成参数选项接口	缩短上线时间
保持数据一致性	确保数据可靠
数据沿袭和影响分析	提供可见性

Data Integration Hub Value Proposition



- 减少点对点数据交换
- 免除重复处理同类数据
- 引入数据管理



- “适时”数据集成
- 自助式数据交换
- 中央管理

Universal Data Access

Messaging, and Web Services



WebSphere MQ
JMS
MSMQ
SAP NetWeaver XI

Web Services
TIBCO
webMethods

JD Edwards
Lotus Notes
SAS
PeopleSoft
Qlikview

SAP
MS Dynamics CRM
Oracle E-Business
Siebel

Packaged Applications



Relational and Flat Files



Oracle
DB2 UDB
DB2/400
SQL Server

Informix
Sybase
ODBC
JDBC

SalesForce CRM
Force.com
RightNow
NetSuite

ADP
Hewitt
SAP By Design
Oracle OnDemand

SaaS/BPO



Mainframe and Midrange



ADABAS
Datacom
DB2
IDMS

IMS
VSAM
C-ISAM
Binary Flat Files

EDI-X12
EDI-Fact
RosettaNet
HL7
HIPAA

AST
FIX
Cargo IMP
MVR

Industry Standards



Unstructured Data and Files



Word, Excel
PDF
StarOffice
WordPerfect
Email (POP, IMPA)
HTTP

Flat files
ASCII reports
HTML
RPG
ANSI
LDAP

XML
LegalXML
IFX
cXML

ebXML
HL7 v3.0
ACORD (AL3, XML)

XML Standards



MPP Appliances



EMC/Greenplum
Vertica
Teradata
Hadoop: HDFS, Hive

Netezza
AsterData
Exadata
SAP HANA

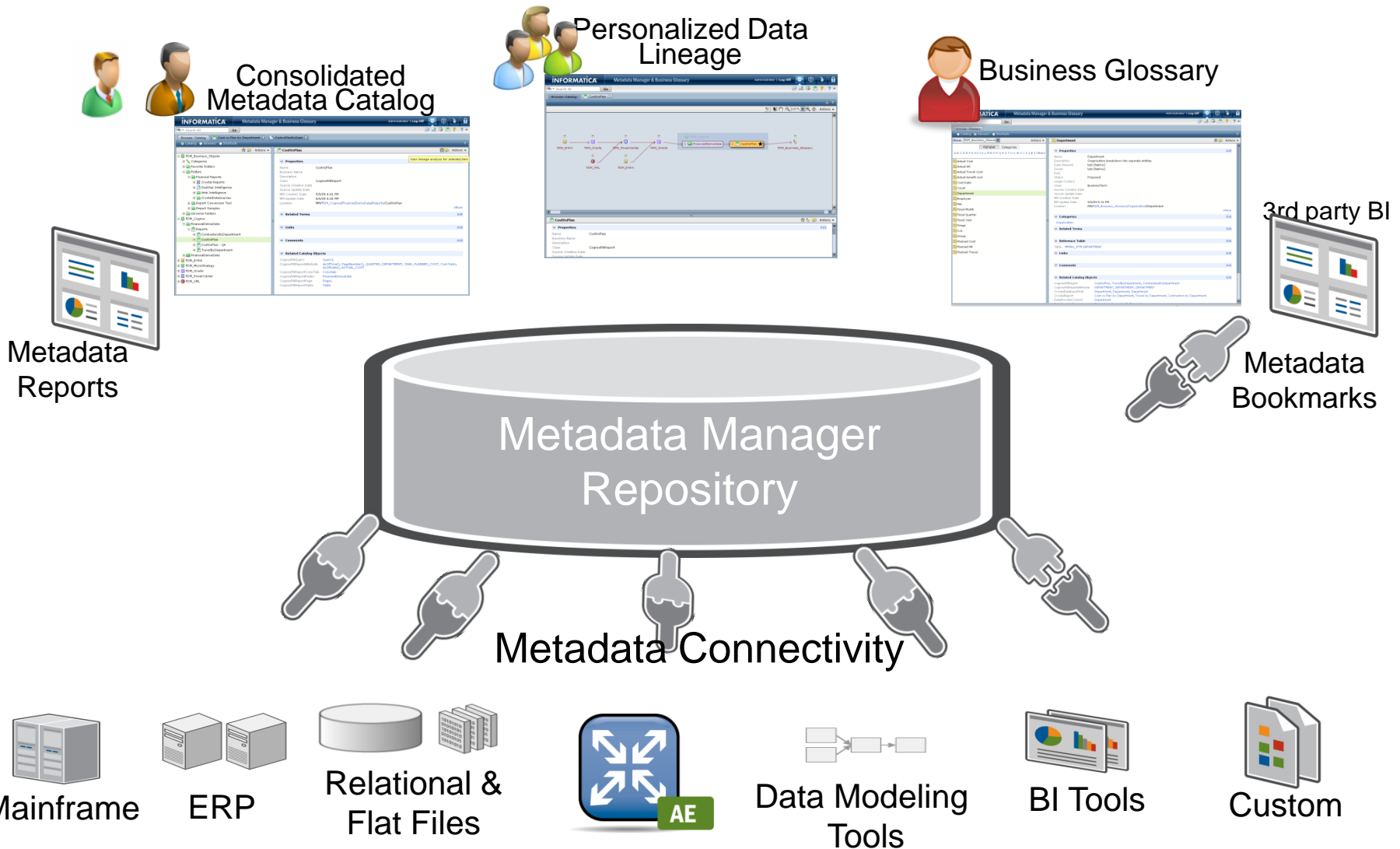
Facebook
Twitter
Bloomberg

LinkedIn
Reuters



Social Media

元数据管理



血缘分析

Informatica Metadata Manager & Business Glossary

Administrator Logout | Manage | Help

Browse: Glossary Account Name

Search All Go Actions

DIH Workflow

DIH Repository

- CRM
- Customer
- Finance
- Sales Datamart

DIH Workflow

Account Name

Properties

Name: Account Name

Description: Business Meaning: This data is common for the investor and the counterparty of the derivative trade. It holds the customer's name for the investor's account. It holds the counterparty's name. Name of the BEA customer who is investor of the derivative contract. For the counterparty of the derivative contract, it is the name of the counterparty.

Source Data #1

System: BG3 (Alnova)

Interface Name: BG3P.BGFD0015.BG3TOZPB.AIOINF.D&RDATE

Field Name: -FHP02-ACC-LONG-FNAME

Protecting Sensitive Data

Restrict Access, Mask Private Data



- Development and Testing
 - Training
 - Support
 - Data Analysis
- Outsourcing and Offshoring

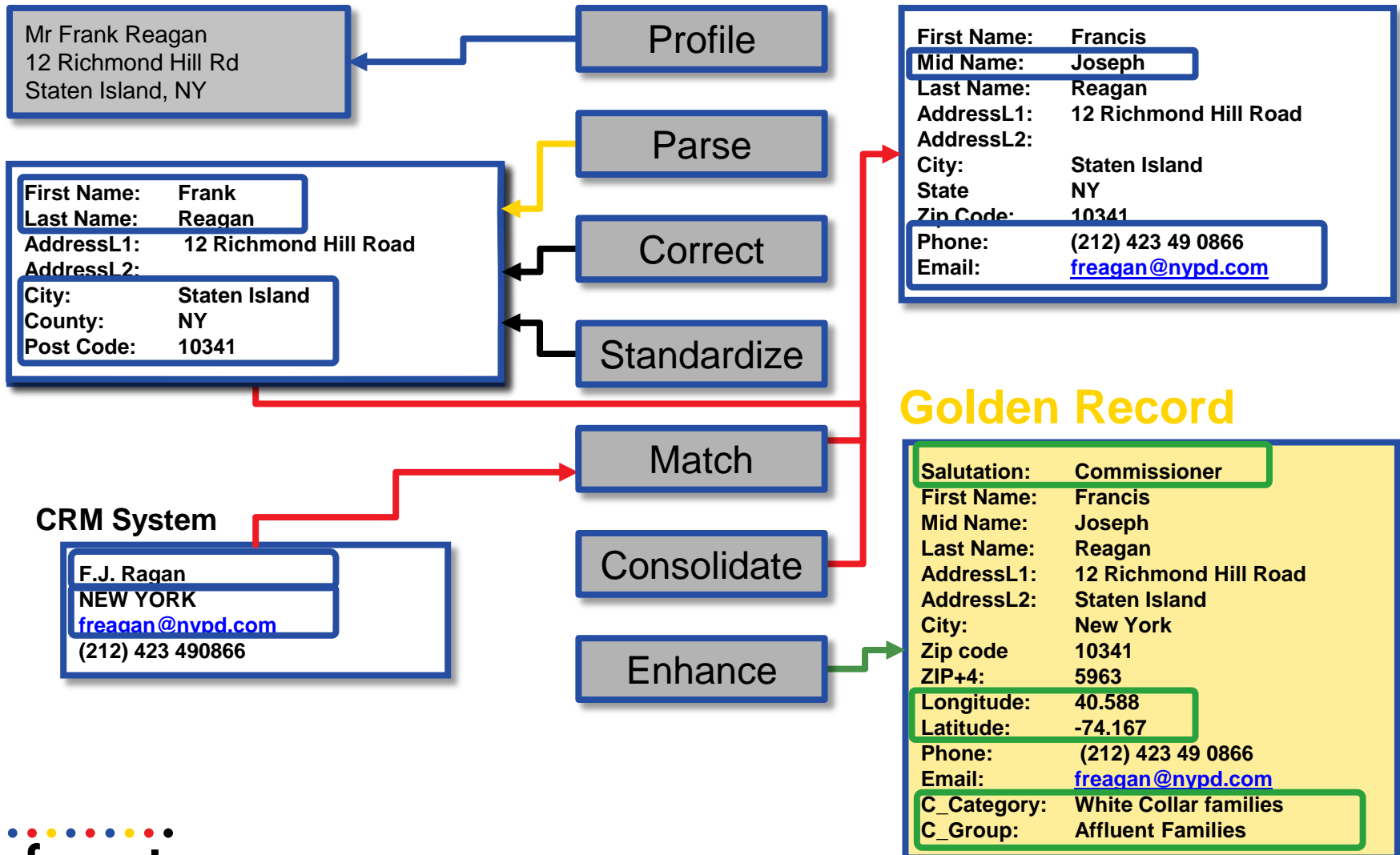
John Smith
654-65-8945
4563-3456-9876-6342
100 Cardinal way
Redwood city

Glen Carter
900-45-2643
4563-XXXX-XXXX-6342
342 54th Street
New York



What is Data Quality?

Let's walk through an example





informatica

Put potential to work.™