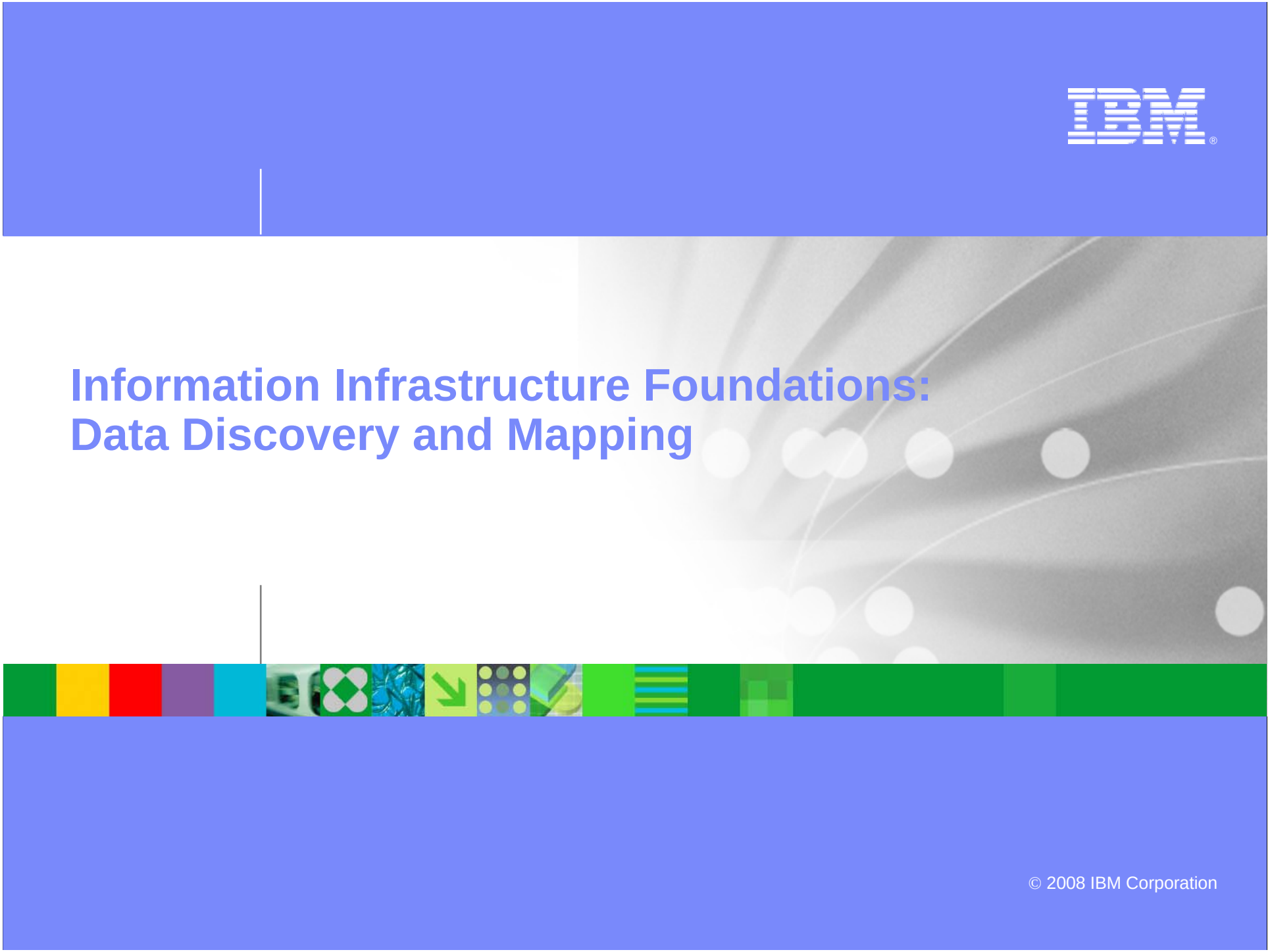
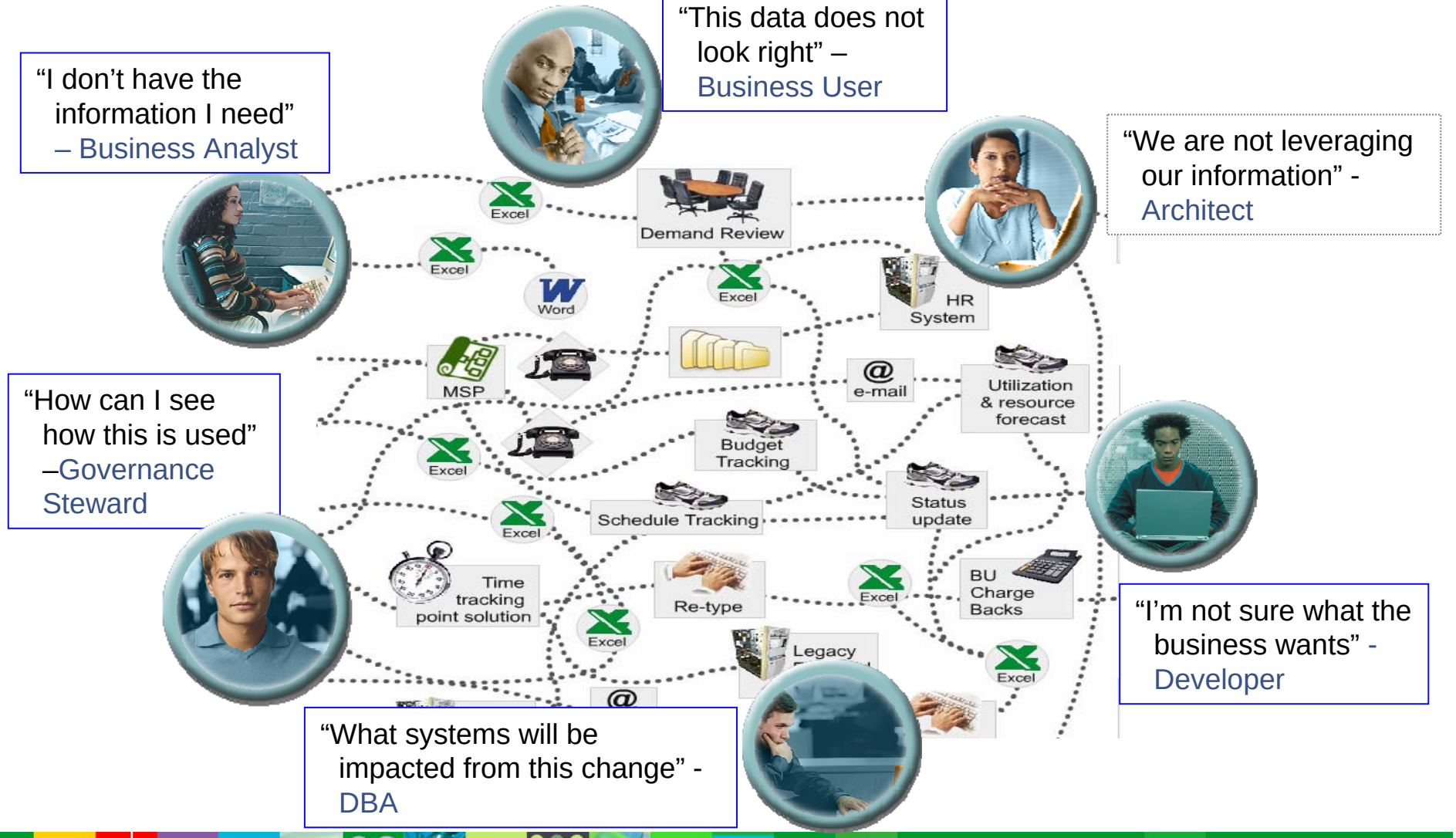
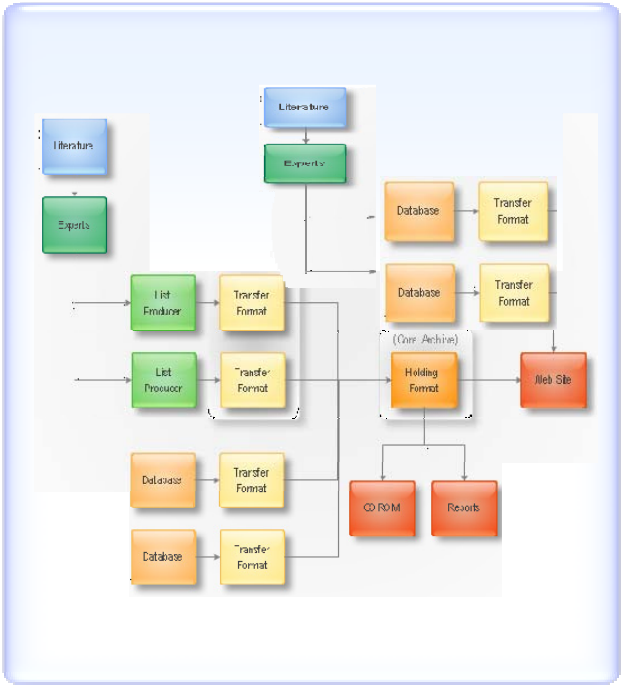


Information Infrastructure Foundations: Data Discovery and Mapping

Understanding Core Information Assets Is Critical



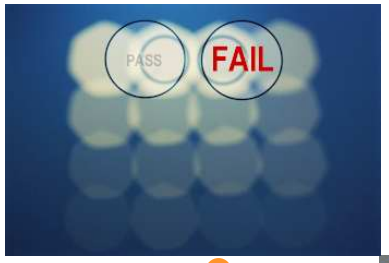
Understanding Core Information Assets



- **Which data is managed by which system?**
- **What relationships exist between data in different systems?**
- **How do I map data together for new uses?**
- **What should my data look like to allow me to use it to optimize my business?**
- **How does the business use the data? What meaning does it have to the business?**
- **Which systems are best sources for specific pieces of information?**
- **What data quality issues do I have?**

Impact of **NOT** Managing Core Information Assets

83% of data integration projects either overrun or fail



Inaccurate or incomplete data is a leading cause of failure in business-intelligence and CRM projects

Scrap and rework Increased $$$

Lack of consumer confidence

Lost opportunities

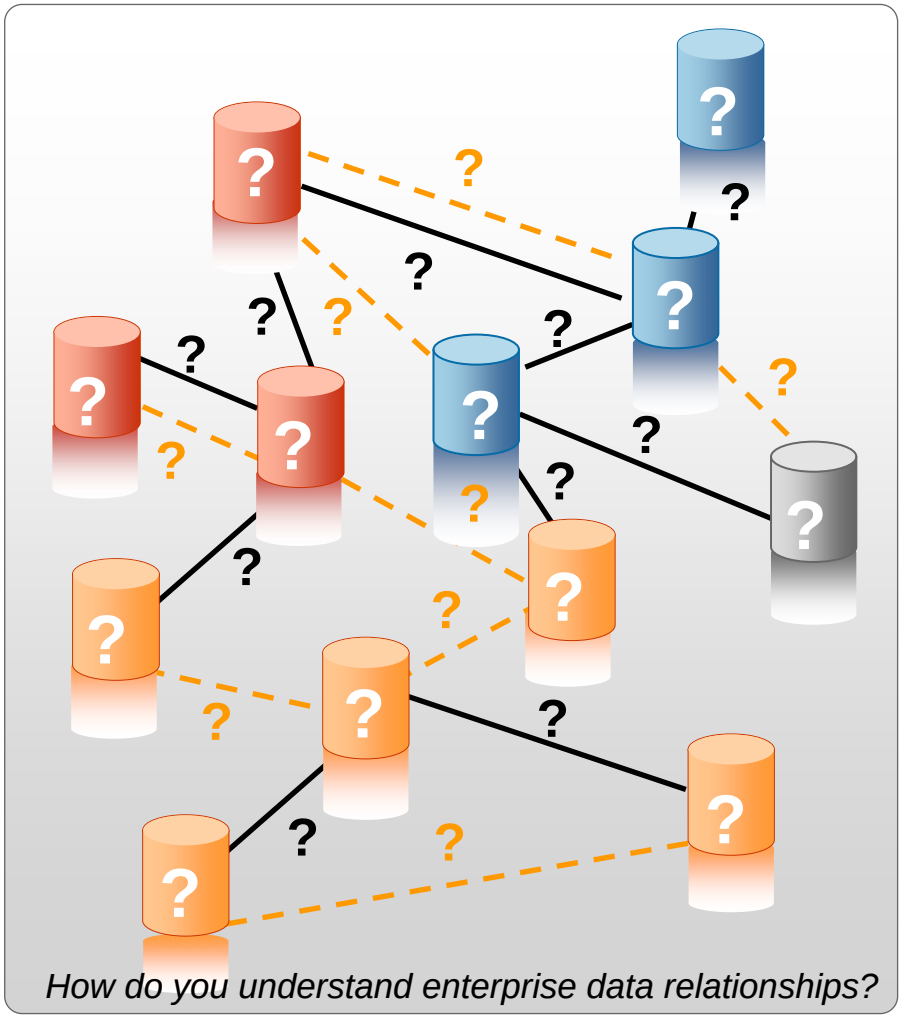
Low data quality costs companies $611 billion annually

25% of time is spent clarifying bad data

Undetected defects will cost 10 to 100 times as much to fix upstream

Challenge of understanding complex data landscapes

The first step for any successful information centric project



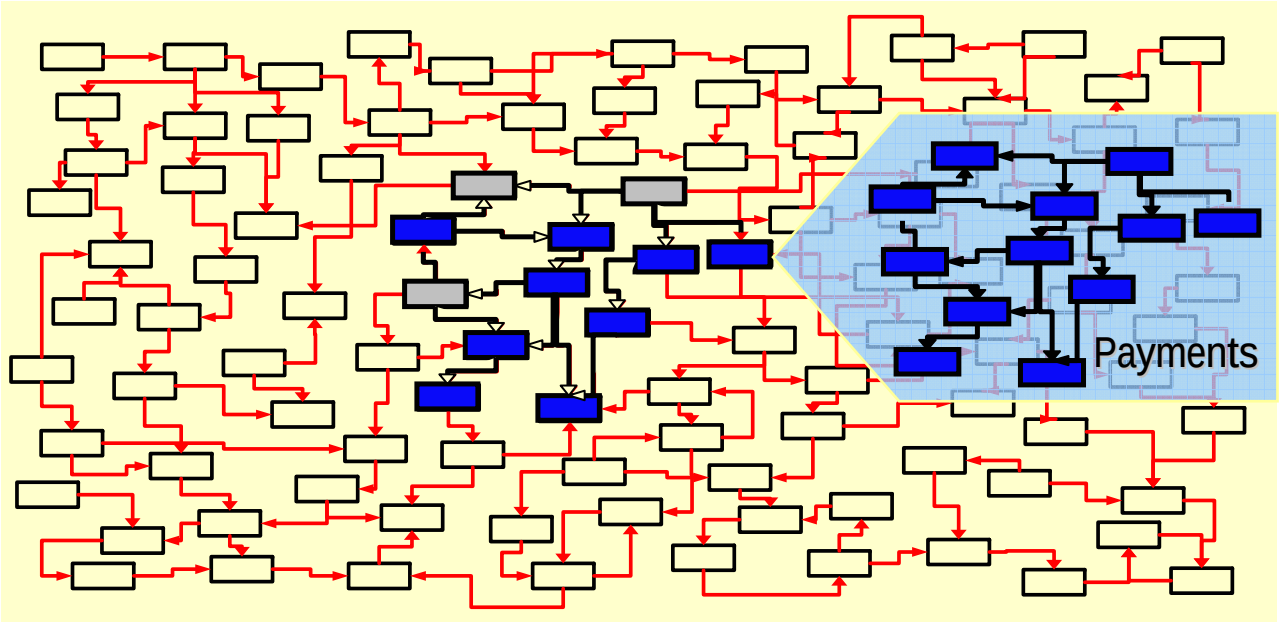
Challenge:

- Distributed heterogeneous sources
- No documentation on data structures
- No understanding of data relationships
- Lack of trusted data – unknown quality
- Limited understanding of confidential data elements

Cost Prohibitive Alternative Solutions:

- Manual spot checking of data
- Hand coding

What is Data Discovery?



Discovery of relationships between data elements, within and across systems – assembling a complete understanding of a business object

IBM acquired Exeros in May 2009 to enhance and extend the portfolio

Extending the Portfolio

Enterprise Projects

- Test Data Generation
- Application Retirement & Consolidation
- Data Archival
- Data De-identification
- Data Quality
- Data Integration
- Master Data Management
- Data Warehousing

Discovery
Extends & Completes
Discover Phase

InfoSphere Foundation Tools

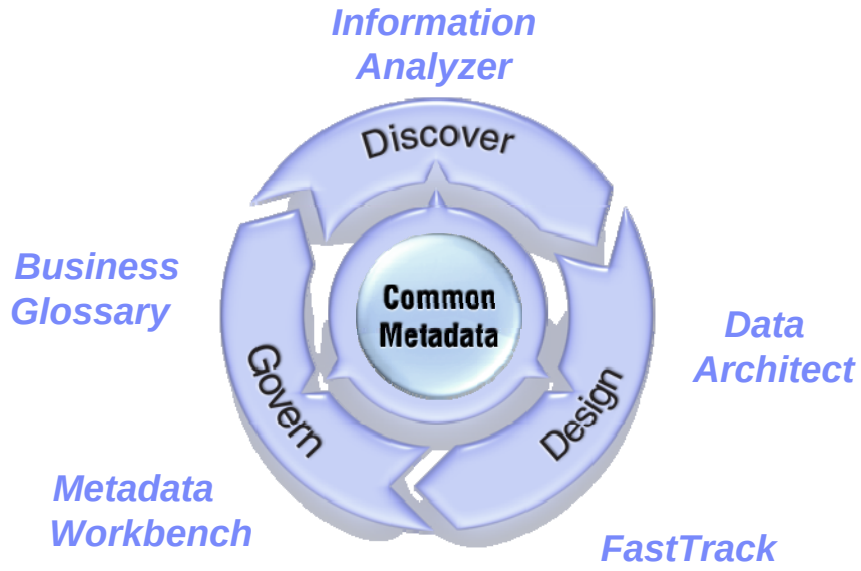
Manage Business Terms	Discover Data Relationships	*Design Enterprise Models*	Capture Design Specifications	Assess, Monitor, Manage Data Quality	Monitor Data Flows
Business Glossary	***New – Discovery***	***Data Architect***	***FastTrack***	***Information Analyzer***	***Metadata Workbench***

InfoSphere Foundation Tools

Software to help convert your information into a trusted strategic asset

Open tools that provide value to any data integration, business intelligence, or data warehouse projects…

Only IBM has invested to provide the breadth of capabilities to define and govern your information…



- ***Discover, understand and relate the data you have to your business***
- ***Design your trusted information structure***
- ***Govern your information over time***

Who InfoSphere Foundation Tools are For

Target Audience

- Data/Business Analysts
- Subject Matter Experts
- Architects
- Governance Stewards

Data/Business Analysts

Subject Matter Experts

Architects

Governance Stewards

What are they working on?

- Information-centric projects:
  - BI & Data Warehousing
  - Master Data Management
  - Application Implementation, Consolidation or Migration
  - Information Architecture
  - Governance Initiatives

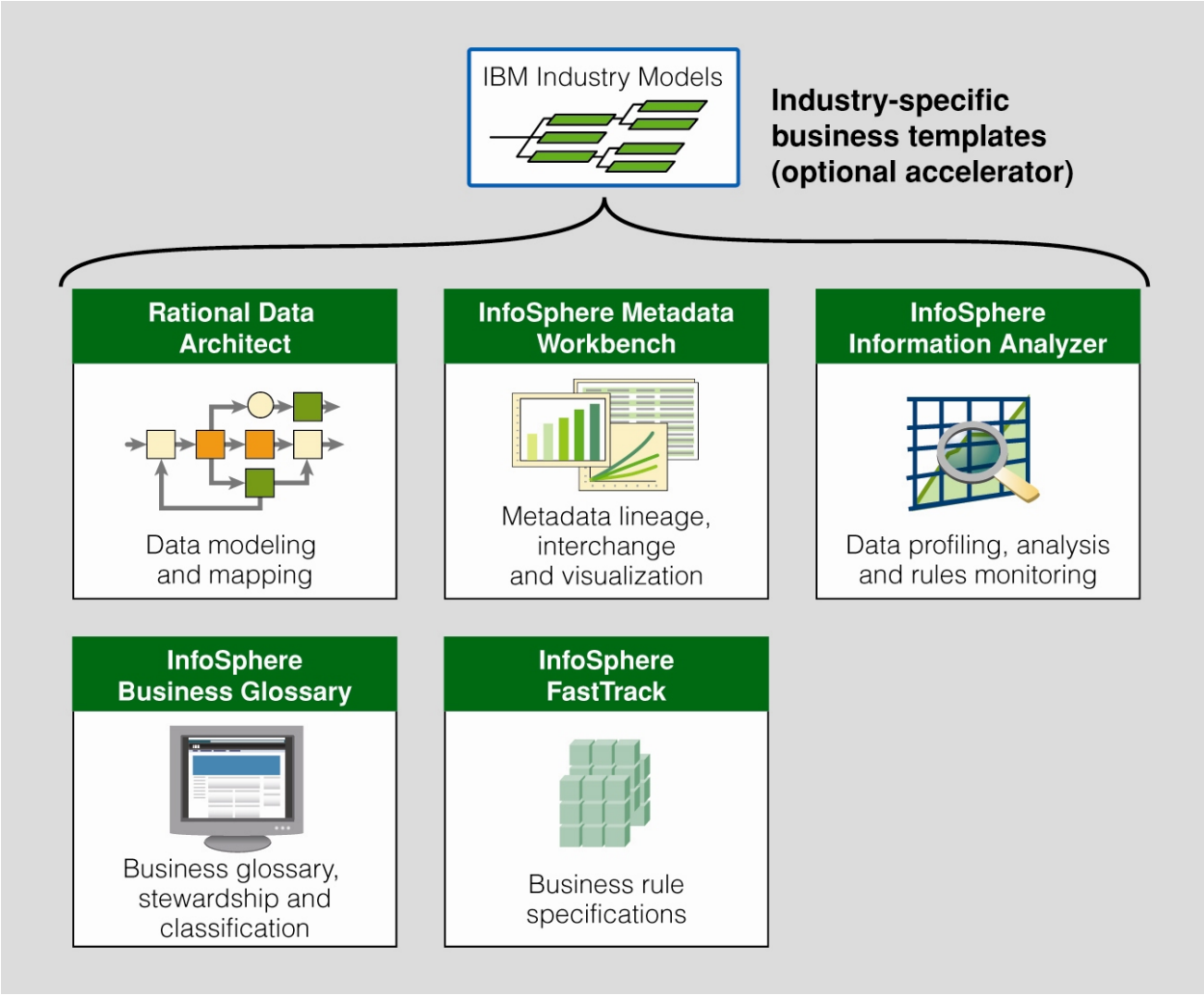
What do these roles do today?

- Manage information manually in disconnected tools, documents, and spreadsheets

What is wrong with what they do today?

- **Time consuming** – churn between business & IT
- **Imprecise & error prone** – manual processes not thorough enough
- **No collaboration** – different roles work in silos
- **Lacks audit trail** – no ongoing record
- **Redundancy** – duplication of effort & storage

InfoSphere Foundation Tools Includes



InfoSphere Foundation Tools – InfoSphere Business Glossary

- ***Starting point* for designing information governance**
- **Web-based authoring, sharing & management of business metadata**
- **Enhanced collaboration**
  - Aligns the goals of the business with the efforts of IT
  - Accelerate project delivery with information sharing
- **Define relationship between business definitions and IT assets**
- **Drive information trust across enterprise applications**
  - From any application, click and automatically search glossary
- **Establish accountability**
  - Creation of stewards and assignment of responsibilities

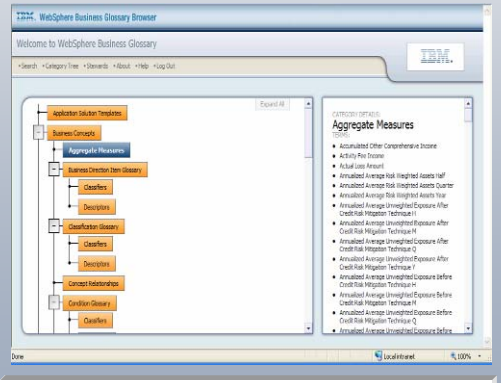
Governance Stewards

Any User

Design & Govern

InfoSphere Business Glossary

Create and manage business vocabulary and relationships and related to physical sources



Business Context View

InfoSphere Foundation Tools - InfoSphere Information Analyzer

- ***First step*** **in defining an information infrastructure across enterprise ecosystem**
- **In-depth analysis of heterogeneous information – IBM or non-IBM sources**
  - Data-centric analysis of databases, files and enterprise applications for content, quality, and structure
  - Secure, detailed profiling of fields, and relationship analysis across fields and across sources
- **On-going measurement and baseline reporting of information quality**
- **Know where information is managed across systems**
  - Capture fitness of sources and re-engineering requirements for downstream use

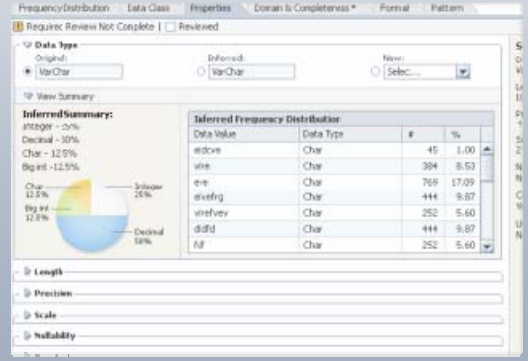
Subject Matter Experts

Business & Data Analysts

Discover & Govern

InfoSphere Information Analyzer

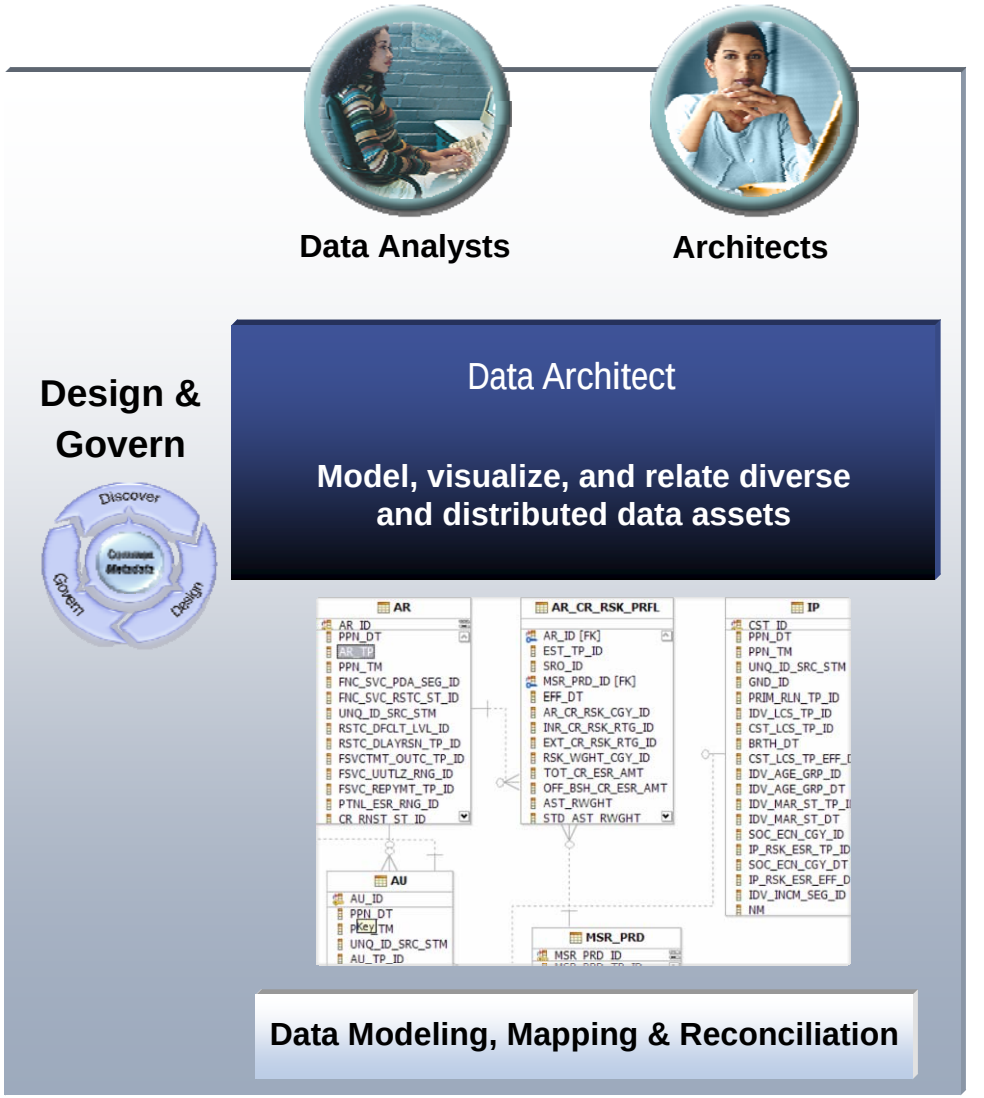
Analyze source data structures, and monitor adherence to integration and quality rules



Physical Understanding

InfoSphere Foundation Tools - Data Architect

- ***Design* information infrastructure *modeling assets***
  - Leverage profiling results from Information Analyzer
  - Create and optimize physical and logical data models
  - Design and deploy federated databases
- **Define enterprise standards to govern data models**
  - Analyze models for conformance
  - Compare and synchronize across models
- **Enhance collaboration**
  - Create glossary models and exchange with Business Glossary
  - SQL & XML generation capabilities



InfoSphere Foundation Tools - InfoSphere FastTrack

- ***Create a solid information infrastructure***
- **Design transformation rules to optimize business information**
  - Leverage Information Analyzer profiling results for most complete requirements
  - Flexible rules support textual descriptions or function logic
- **Centrally track design decisions for auditing**
- **Enhance collaboration**
  - Easy to use desktop interface or import work from Excel spreadsheets
  - Define and link InfoSphere Business Glossary terms to physical structures
- **Accelerate development of InfoSphere integration jobs**

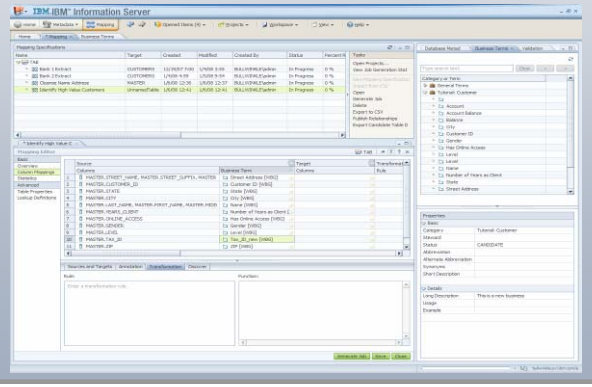
Subject Matter Experts

Data & Business Analysts

Design

InfoSphere FastTrack

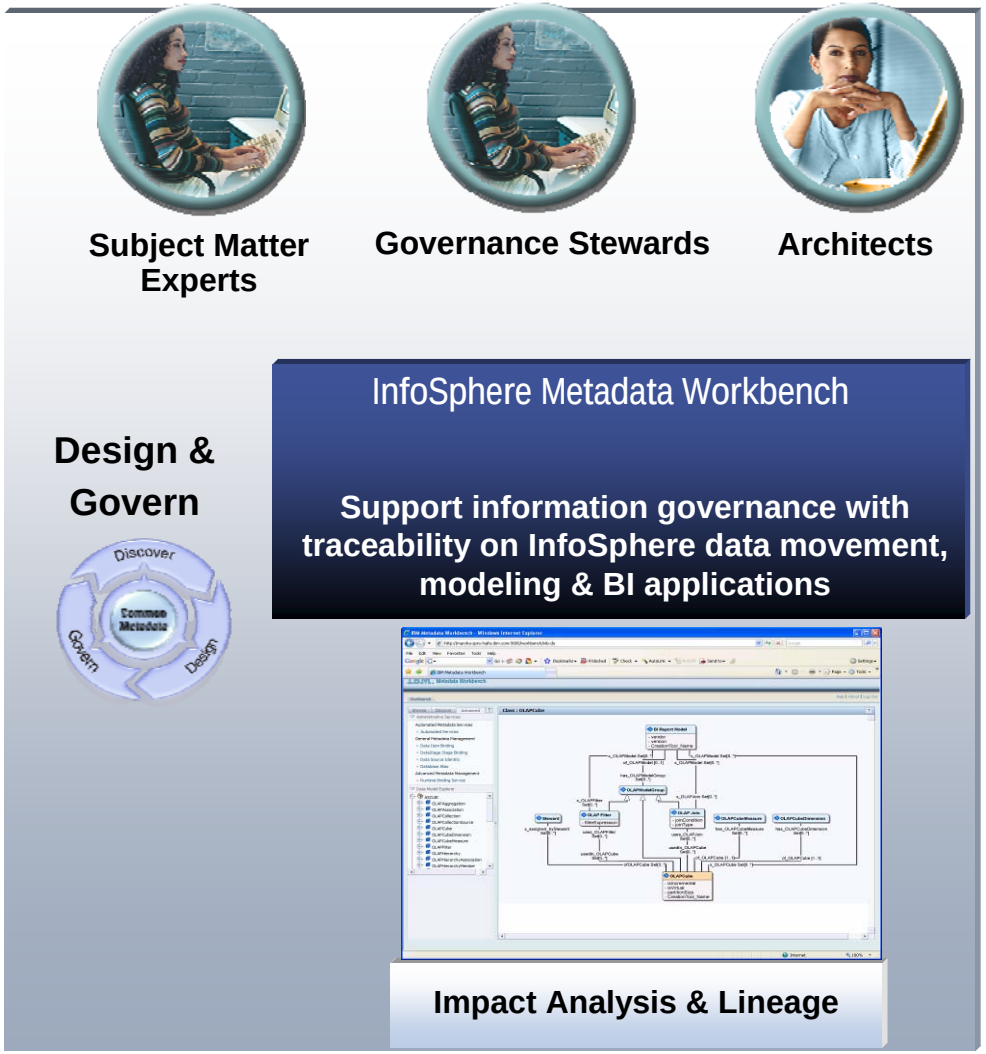
Create required transformation rules for source-to-target historical mapping documentation



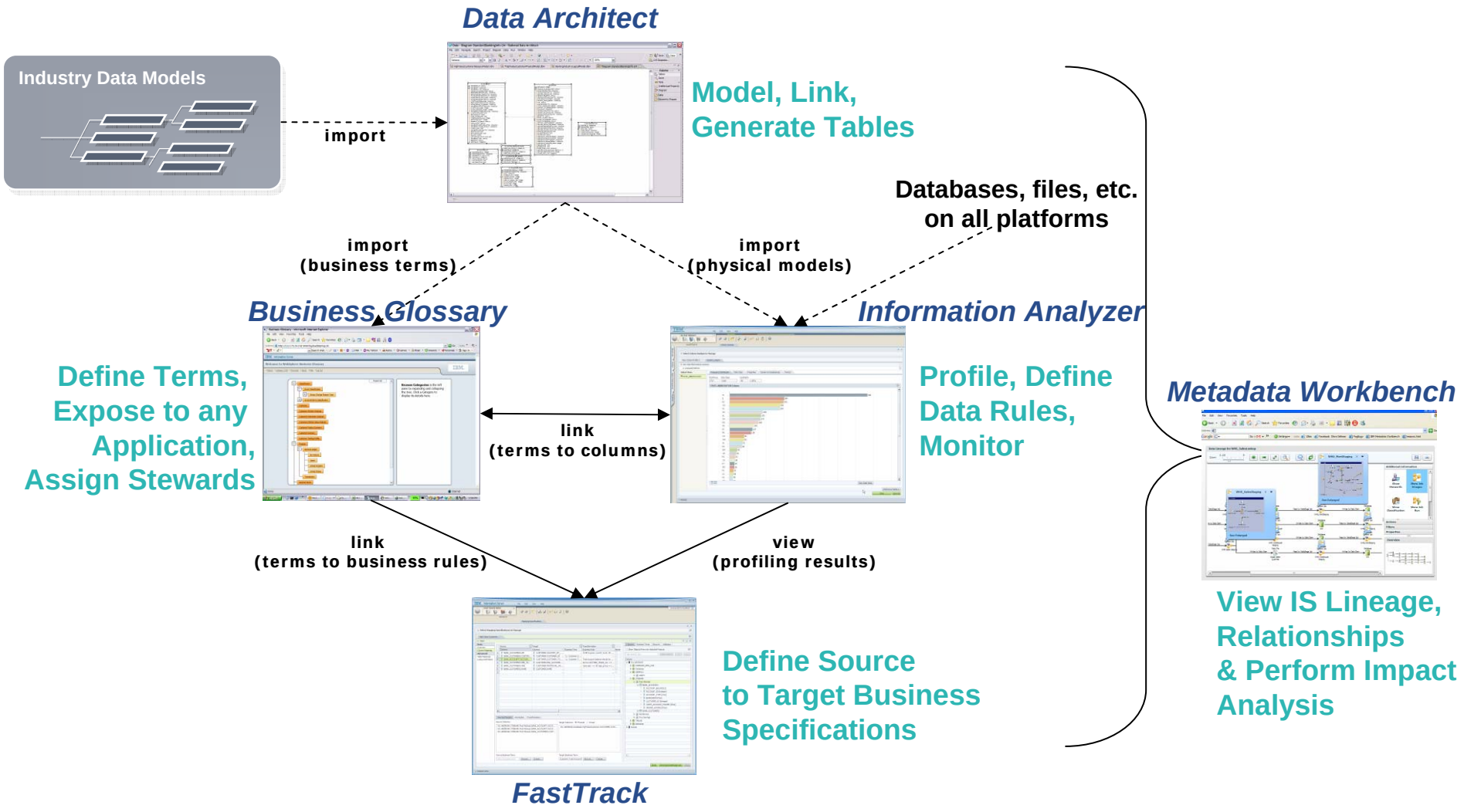
Link Data Rules with Business Meaning

InfoSphere Foundation Tools – InfoSphere Metadata Workbench

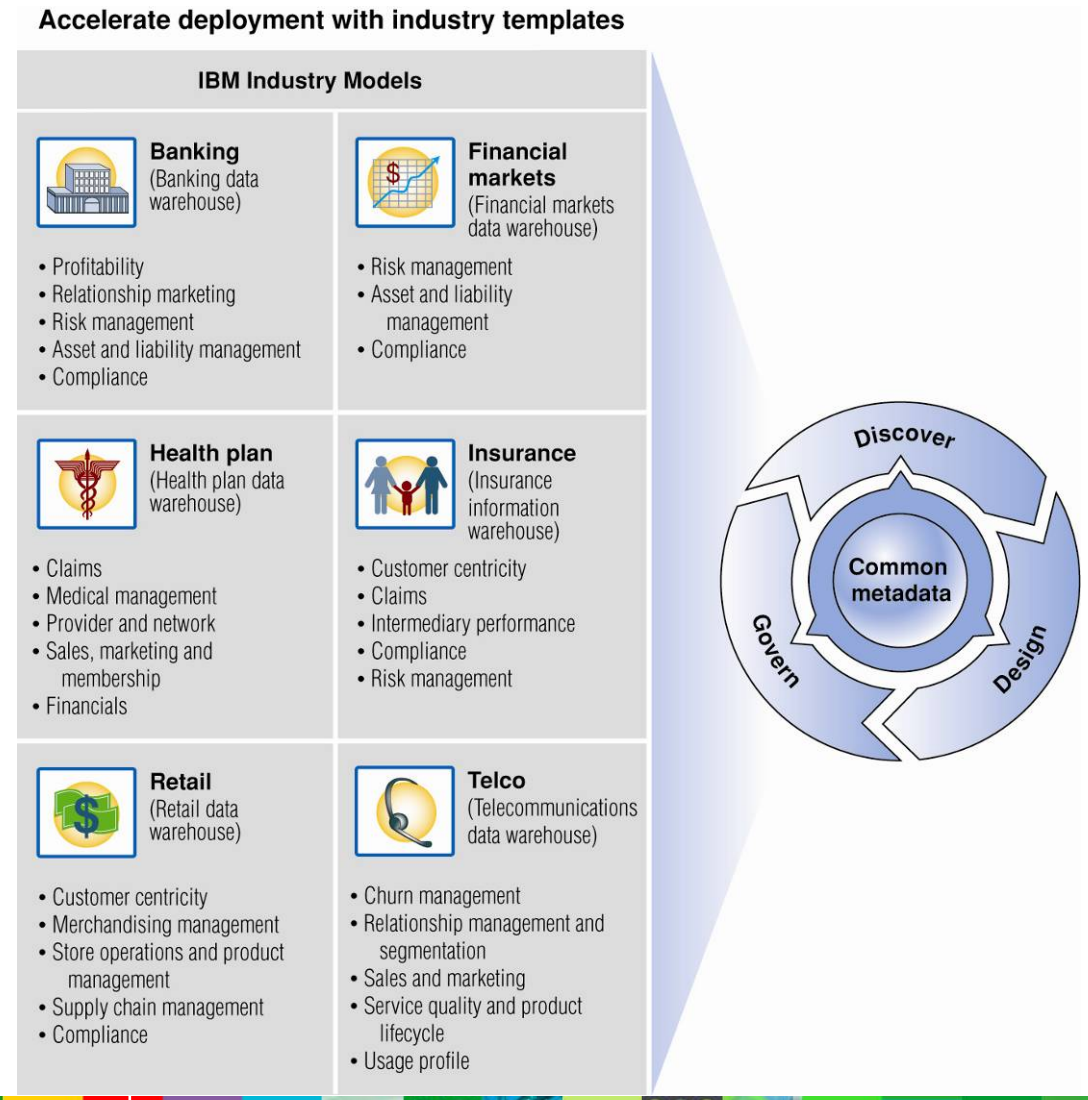
- **Understand your *end-to-end information infrastructure***
- **Proactively manage and administer your information infrastructure**
  - Web-based exploration of metadata relationships – InfoSphere, modeling and BI applications
  - Explore and analyze both graphically and textually
  - Perform searches and customized queries
- **Assess and mitigate change management risk**
  - Assess dependencies across InfoSphere and 3rd party tools
- **Support compliance and governance initiatives (eg Sarbanes-Oxley, Basel II)**
  - Trace data lineage of modeling, BI reports and InfoSphere objects



InfoSphere Foundation Tools Integration



Accelerate Foundation Tools Deployment with Industry Templates



- Provides massive **acceleration** for your **Information Agenda**
- Pre-defined proven industry models across **six industries**
- Robust data models form a complete **foundation for design**
- **Pre-populate Business Glossary** with full industry vocabulary

How the Products Work Together

Discovery

Understand Data Relationships

Automated relationship discovery

Accelerate and improve data quality monitoring and management

Consolidation prototyping capabilities

Build a better quality solution the first time

Transformation rules discovery

Save time finding implied logic in legacy applications

Automated business object identification and confidential data location

Achieve regulatory compliance faster

Information Analyzer

Assess, Monitor, Manage Data Quality

DataStage / QualityStage
MDM Server

ETL / Data Quality,
Master Data Management

FastTrack

Capture Design Specification

Optim

Data Privacy, Test Data Management
Data Growth & Application Retirement

How the Products Work Together – Metadata Integration

Discovery

Understand Data Relationships

Automated relationship discovery

Physical Data Assets (Sources, Tables, Columns, Data Connections)
Data Relationships (Primary Keys, Foreign Keys)
Summarized Analysis at the table and column level

Information Analyzer

Assess, Monitor, Manage Data Quality

Consolidation prototyping capabilities

Physical Data Assets (Sources, Tables, Columns, Data Connections)
Data Relationships (Primary Keys, Foreign Keys, Matching keys, mapping and transformation rules)
Summarized Analysis at the table and column level

DataStage / QualityStage
MDM Server

ETL / Data Quality,
Master Data Management

Transformation rules discovery

Discovered Transformation Rules & Transformation annotations
Source to Target Definitions
Physical Data Assets (Sources, Tables, Columns, Data Connections)
Data Relationships (Primary Keys, Foreign Keys)
Summarized Analysis at the table and column level

FastTrack

Capture Design Specification

Automated business object identification and confidential data location

Physical Data Assets (Sources, Tables, Columns, Data Connections)
Data Relationships (Primary Keys, Foreign Keys)
Summarized Analysis at the table and column level
Logical Data Assets (Discovered Business Objects)

Optim

Data Privacy, Test Data Management
Data Growth & Application Retirement