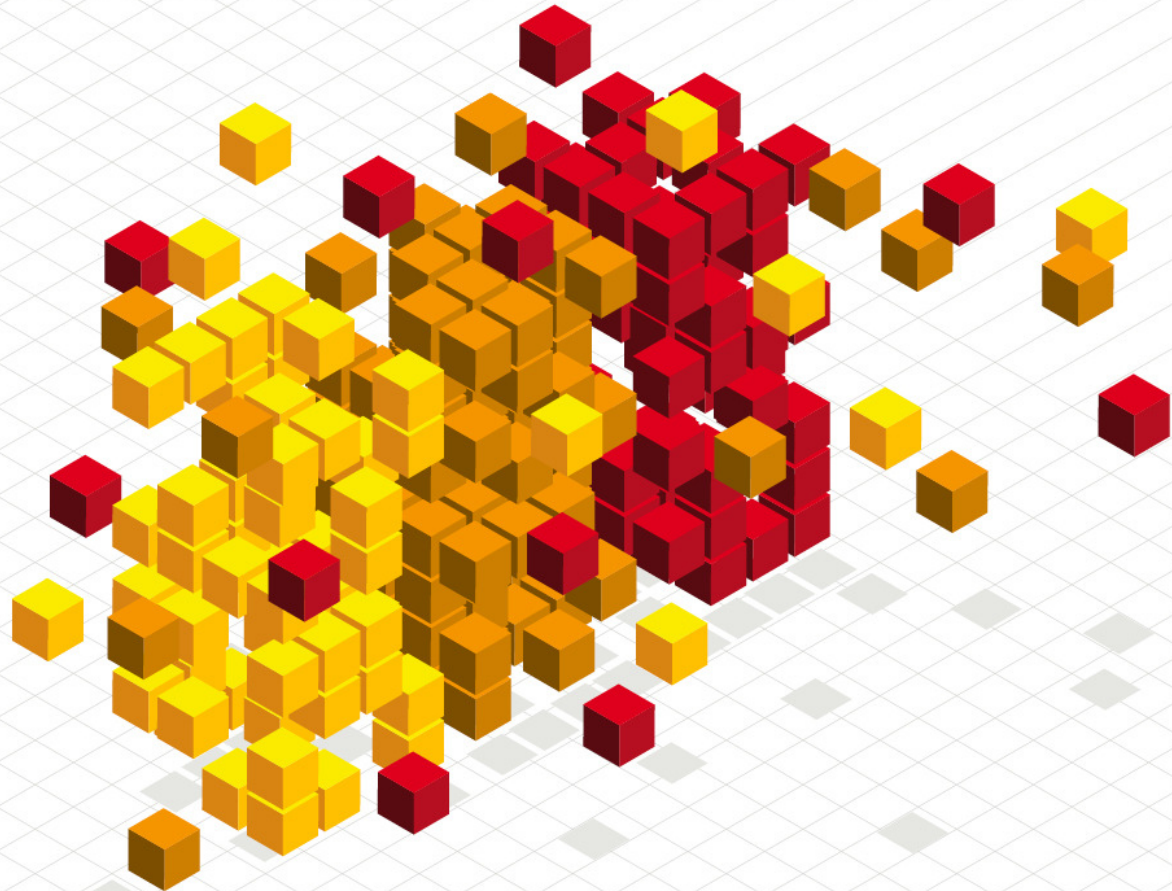


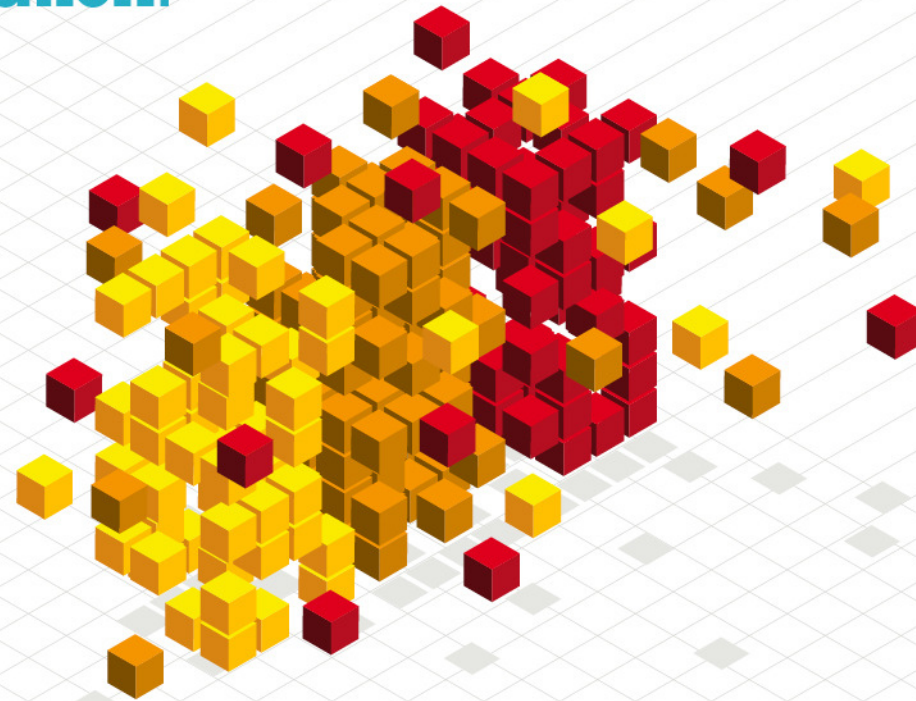
Smarter technology for a smarter planet:

**Trusted data is the foundation
for a smarter organisation.**



Smarter technology for a smarter planet:

**Trusted data is the foundation
for a smarter organisation.**



InfoSphere Information Governance: Getting Started with Foundation Tools

Roger Hecker

*Product Manager - InfoSphere Business Glossary & Metadata Workbench
IBM Software Group, Information Management*



Challenges around governing the usage, sharing and processing of massive amounts of electronic data

Lack of trusted information

- No alignment of definitions across business and IT
- No clear understanding of data sources & relationships
- No standardized quality rules or threshold metrics
- Lack of control over test data environments

“By 2013, 25% of the companies in highly regulated industries will create and staff positions in accounting, human resources, compliance and audit and law that deal explicitly with the management of information via technology.”

– Gartner, Inc., “Organizing for Information Governance”

Debra Logan, November 2009

Growing infrastructure and resource management costs

- No policies for management of data growth – performance degradation
- Redundancy of data
- Disparate, complex applications and more users

“...An [Information Management] strategy should incorporate life-cycle information governance practices...[for] business optimization, agility, and transformation.

– Forrester Research, “Refresh Your Information Management Strategy to Deliver Business Results”

Rob Karel & James G. Kobielus, August 2009

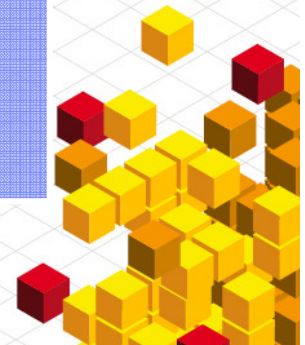
Lack of common security and privacy approach

- Risk of security breaches, compliance, audit failures
- No ability to assess areas of vulnerability and prevent unauthorized intrusion
- Lack of an overall protection strategy (relational / non-relational data and access controls)

“If you are going to protect your company's most valuable asset—your data—you will begin to view data security as a component of a more comprehensive information governance strategy.”

– Hurwitz & Associates,
“Why you need an information governance strategy for 2010”

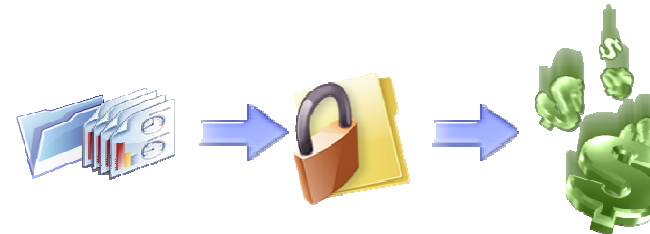
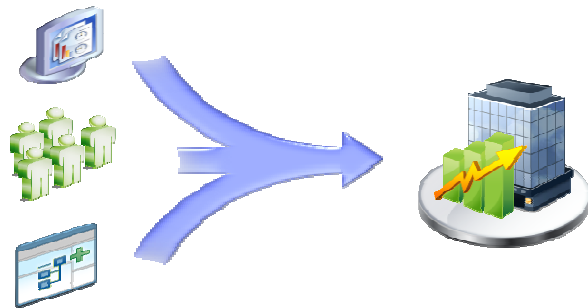
Marcia Kaufman, December 2009



Information Governance Creates Order out of Information Chaos

Information Governance is the exercise of decision rights to optimize, secure and leverage data as an enterprise asset.

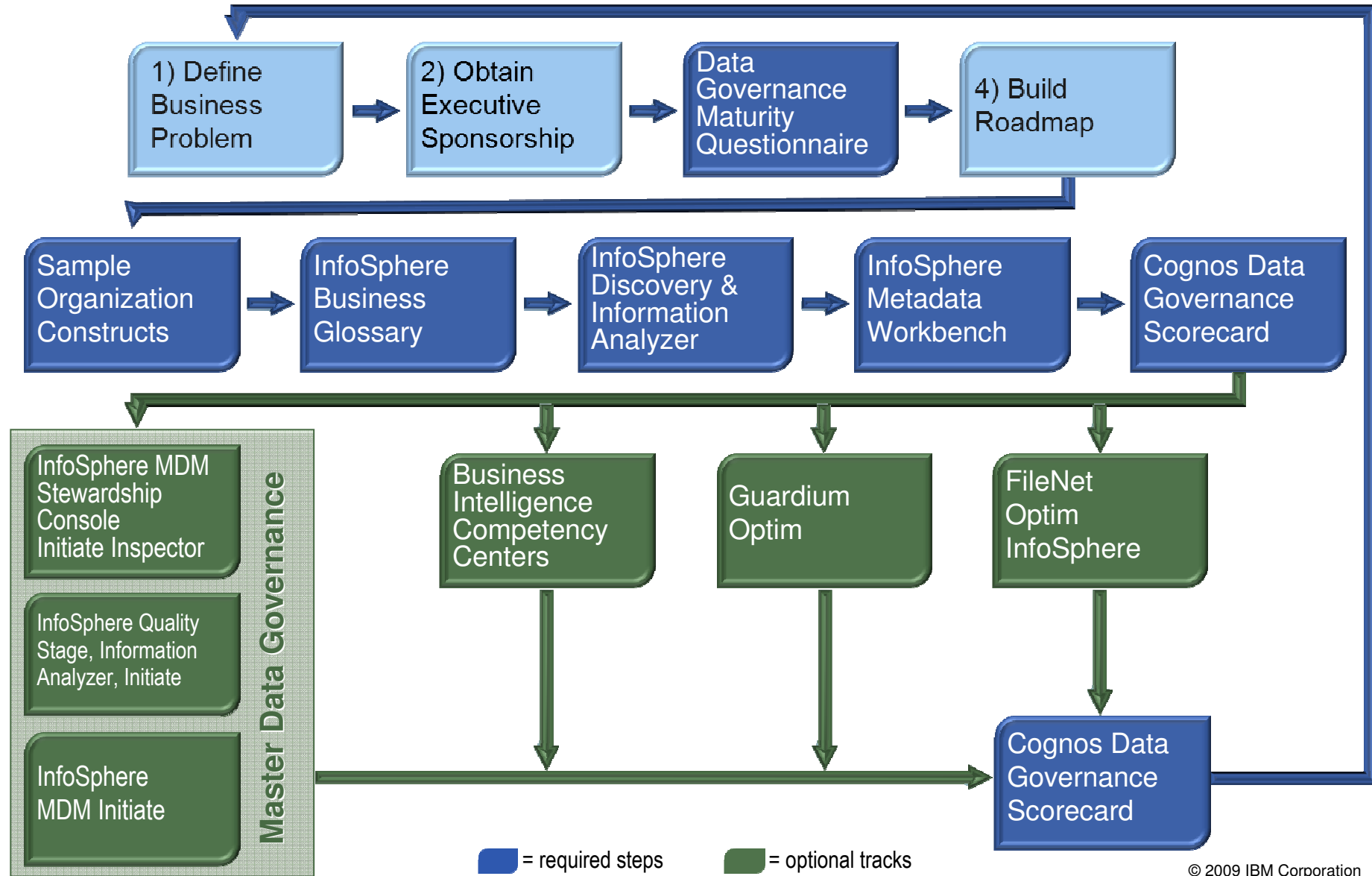
- Orchestrate people, process and technology toward a common goal
- Promotes collaboration
- Derive maximum value from information
- Leverage data as an enterprise asset to drive opportunities
- Safeguards information
- Ensure highest quality
- Manage it throughout lifecycle



Governing the creation, management and usage of enterprise data is no longer an option. It is:

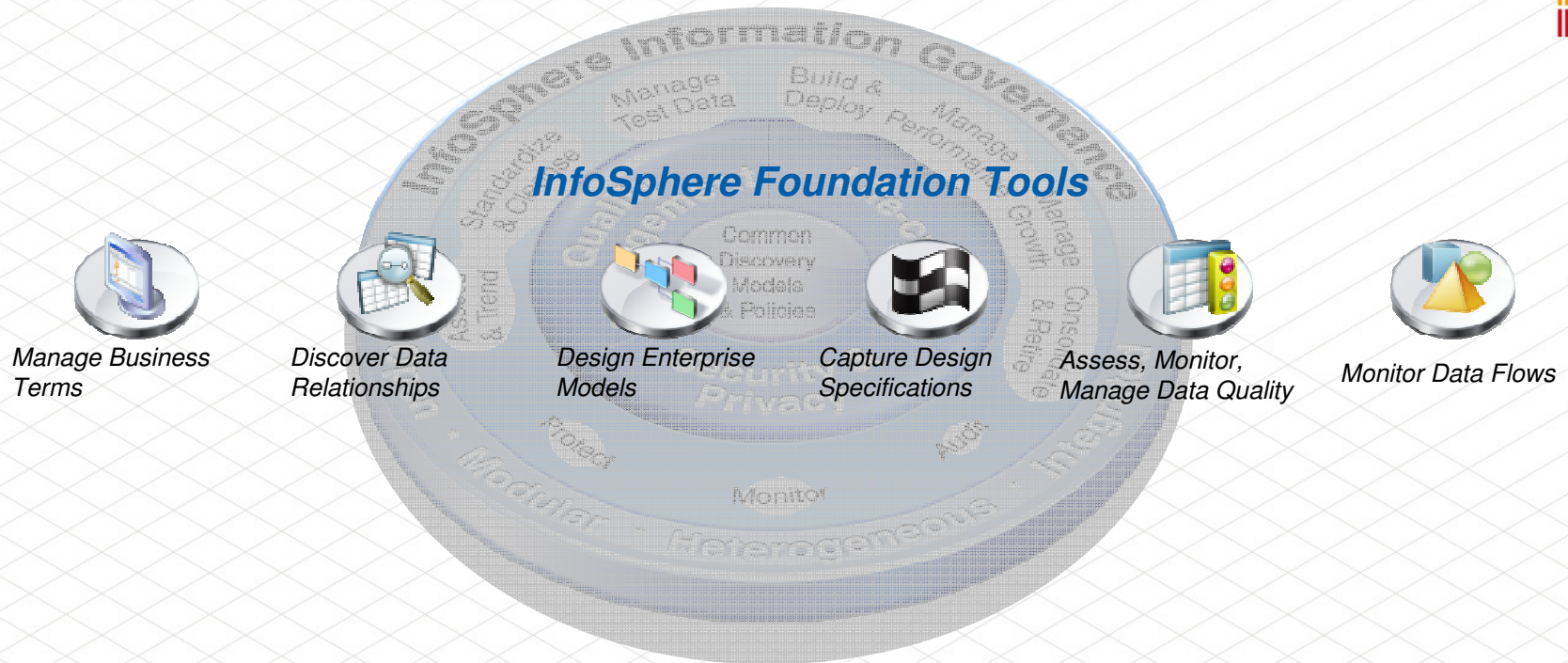
Expected by your customers ♦ Demanded by the executives ♦ Enforced by regulators/auditors

IBM Data Governance Unified Process – Software Components

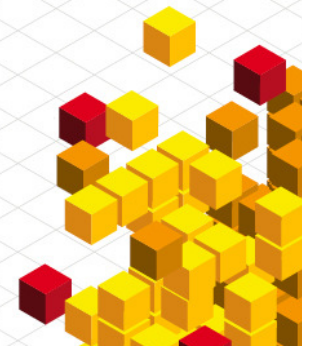


InfoSphere Foundation Tools

Put your arms around your data ... for Trusted Information



- Jump start any data integration, MDM, data warehousing, business intelligence or information governance project
- Discover your data across systems
- Design your trusted information structure
- Govern your information over time



Information Governance and InfoSphere Foundation Tools



Key Governance Capability

- Specify Information Flows
- Establish a common set of terms and definitions
- Understand the data I have in my organization
- Understand the way that data flows and how it is connected
- Assign meaning to data assets / Assign ownership to data assets
- Monitor data quality

Foundation Tools product

- ✓ InfoSphere Blueprint Director
- ✓ InfoSphere Business Glossary
- ✓ InfoSphere Discovery and InfoSphere Information Analyzer
- ✓ InfoSphere Metadata Workbench
- ✓ InfoSphere Business Glossary, Information Analyzer, Metadata Workbench
- ✓ Information Analyzer, InfoSphere Business Glossary, InfoSphere Metadata Workbench



Establish Information Project Blueprint: Blueprint Director

Capabilities

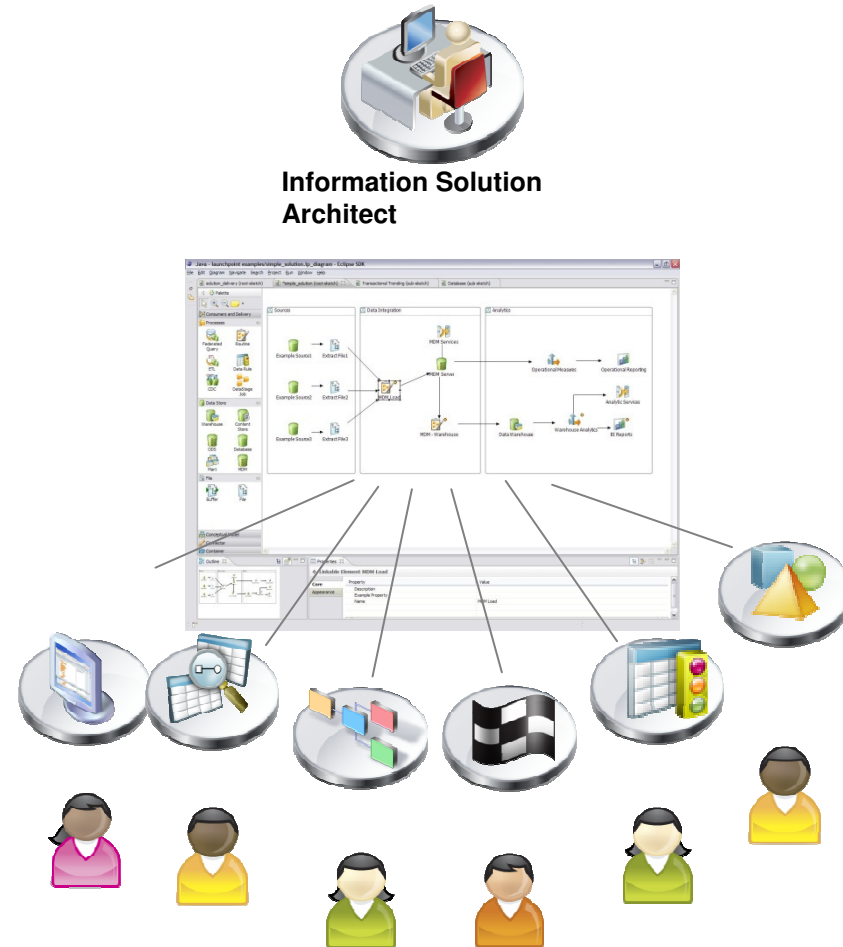
- Create and manage actionable blueprints of your information project
- Accelerate information centric projects by leveraging templates with methods – e.g. warehousing, master data – and customize your best practices
- Navigate through your information project by managing the information roadmap & its evolution over time

Benefits

- Establish business-driven development by:

Aligning business & IT views

Creating a consistent end-to-end design from business requirements



Information project leadership team
(stakeholders, bus. Analysts, stewards, specialty architects, etc.)

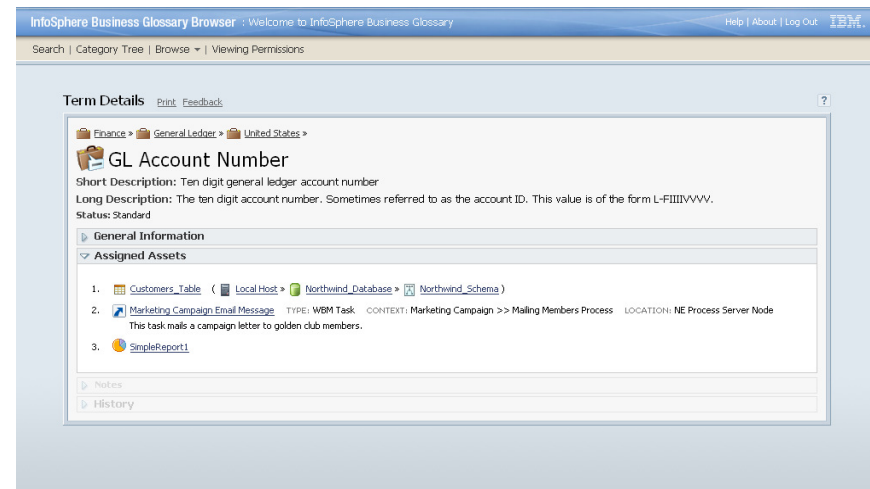
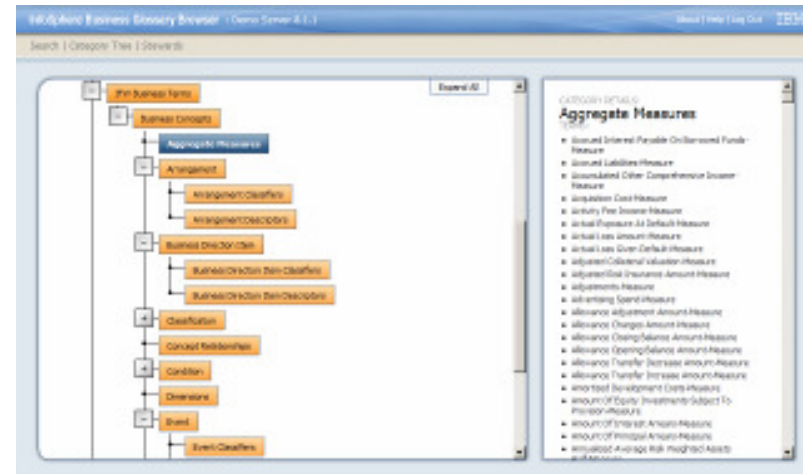
Establish Common Terms and Definitions: Business Glossary

Capabilities

- Facilitate **business & IT communications via a common business vocabulary**
- Web based **interface shared across enterprise business teams**

Benefits

- Aligns the efforts of **IT with the goals of the business**
- Improved productivity
- Improved collaboration



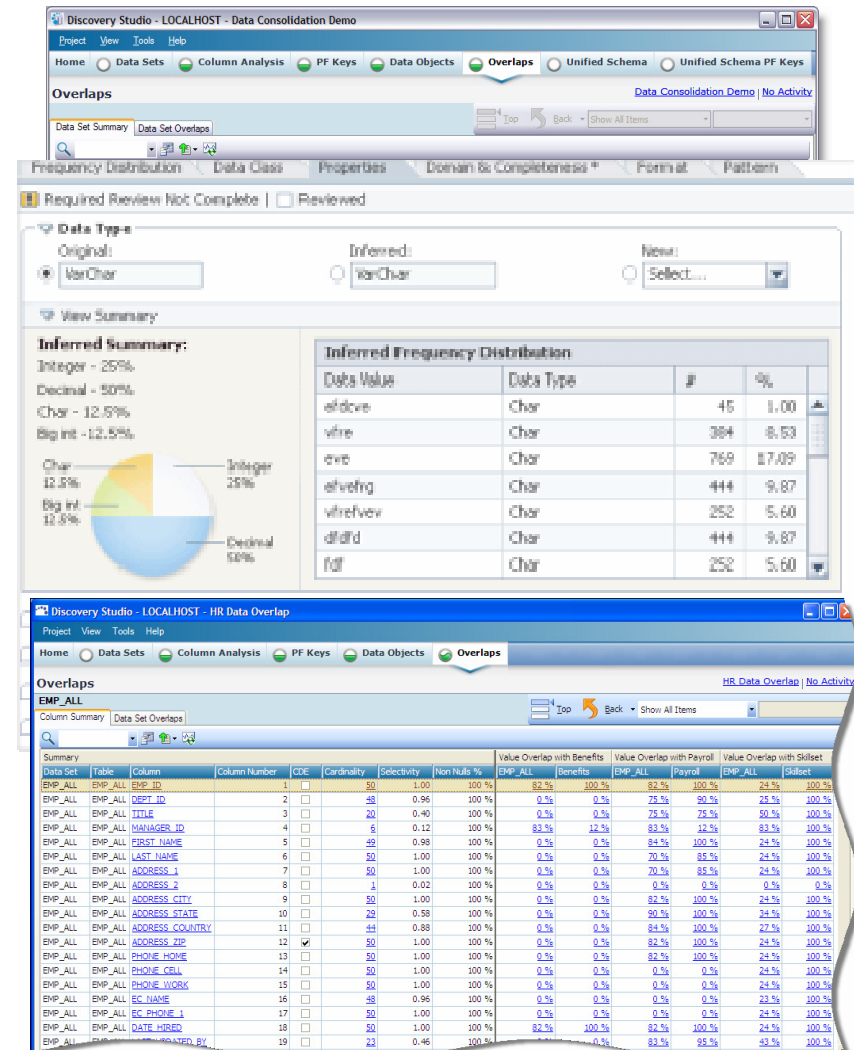
Discover and understand my data: InfoSphere Discovery, Information Analyzer

Capabilities

- Automatic data discovery
- Perform data quality assessment
- Define business rules to monitor data quality
- Establish stewards for governance of data quality
- Discover data transformation rules and heterogeneous relationships

Benefits

- Increase your time to value by automating the analysis of your data sources
- Identify data quality issues early to reduce project risks
- Monitor quality metrics over time for compliance





Understand data flows and relationships: Metadata Workbench

Capabilities

- Trace your data movement from sources to applications and business reports
- Understand relationships across your data flows
- Analyze the impact of changes to your data
- Centralized management of your IT assets

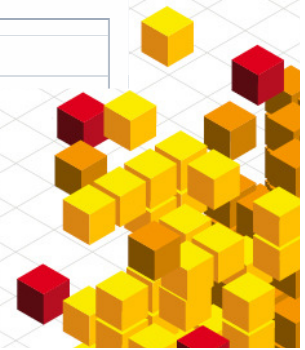
Benefits

- Govern your data by understanding where its coming from and where its being used
- Support audit and regulatory compliance requirements by reporting on your operational metadata

The screenshot displays the Metadata Workbench interface. The top section shows a data flow diagram with nodes representing data sources and destinations. The right pane lists assets identified in the image, including 'SalesTransaction' and 'SalesTransaction_02'. Below the diagram, the 'DataStage Job' details are shown for 'SalesTransaction'. The job description is 'Reconcile all current transactional information into the operational data store, comprising activity of the past 12 months.' The job contains several stages: 'Operational_Sales', 'Operational_Sales', 'Sales_Report', 'Sales_Report_Source', 'Writing_Sales_Figures_To_XML', and 'Writing_Sales_Figures_To_XML'. The 'DataStage Job Usage' section shows a list of job runs, including 'SalesTransaction 2007-01-24 15:21:12' and 'SalesTransaction 2007-01-24 15:26:44'.

Job	SalesTransaction
Job	SalesTransaction
Project	LAURIE DEMO\TAB
Folder	\\Jobs\JACME_Sales
Description	Reconcile all current transactional information into the operational data store, comprising activity of the past 12 months.
Data Steward	Linker-Nestli
Contains DataStage Stages	Operational_Sales, Operational_Sales, Sales_Report, Sales_Report_Source
Contains DataStage Links	Writing_Sales_Figures_To_XML, Writing_Sales_Figures_To_XML
Contains DataStage Local Containers	None

Operational Job Locator Information	SoftwareExecutable (Job)='SalesTransaction'
Job Runs	SalesTransaction 2007-01-24 15:21:12, SalesTransaction 2007-01-24 15:26:44, SalesTransaction 2007-01-25 12:06:21
Previous DataStage Job	None/Concurrent



Assign meaning and ownership to data assets: Business Glossary, Metadata Workbench, Information Analyzer

Capabilities

- Allows creation of stewards & assignment of their responsibilities for terms & assets.
- Link business terms to information assets

Benefits

- Provides business context to information technology assets
- Establishes responsibility and accountability in accordance with data governance policies

The screenshot displays two overlapping windows from the IBM InfoSphere Business Glossary Browser and Metadata Workbench.

The top window, titled "InfoSphere Business Glossary Browser", shows "User Details" for Ms. Jackie Smith. Her details include:

- Job Title: IT Steward
- Organization: IIS
- Email Address: jsmith@mycompany.com
- Location: Room 1204
- Business Address: 200 Elm Street Any City, USA 90000
- Office Phone Number: 555 - 1212
- Mobile Phone: 555 - 1212
- Fax Number: 555 - 1212

 Below the user details is a section for "Managed Assets" with a list of 9 items:

- EWS
- EWS_REPORTS
- EWS_REPORTS
- Financial Glossary Pack
- Insurance Glossary Pack
- PROD_MRT (EWS)
- ProductionRunReport v1.5
- SALES_MRT (EWS)
- Sales Person Report v1.3

The bottom window, titled "InfoSphere Metadata Workbench", shows details for a "Database Table: CUSTOMER". The metadata includes:

- Name: CUSTOMER
- Imported From: ODBCConnector 3.5
- Short Description: Customer DWS Data
- Long Description: Bank Customer Data Warehouse
- Term: Customer ID
- Steward: Richard Keith
- Database: BANKDATA
- Schema: BANK
- Columns: 2 more instance(s) exist. View all 27 results.
 - ACCOUNT_BALANCE
 - ACCOUNT HOLDER_ID
 - ACCOUNT_ID
 - ACCOUNT_TYPE
 - ADDR1
 - ADDR2

Manage and Monitor Data Quality: Information Analyzer, Metadata Workbench, Business Glossary



Capabilities

- Ongoing monitoring of your data quality – beyond Initial load
- Monitor based on sophisticated rule logic
- Report on data quality issues
 - Data Filters/Rules, Filter Templates, Metrics
 - Exception Tables and Reports
 - Metrics, Scorecards and Trend Analysis and Reports

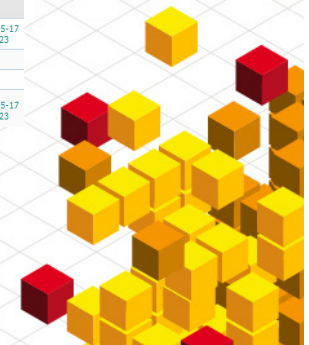
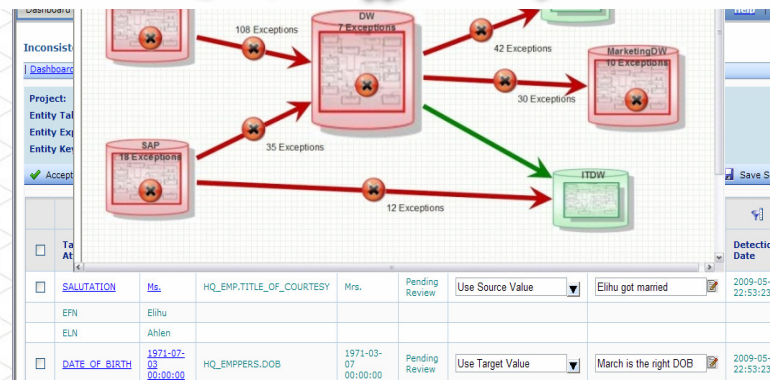
Benefits

- Govern your data quality
- Ensures the lasting value of your data integration investment
- Ensures superior data quality over time

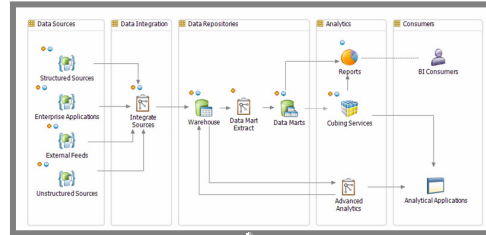
Inferred Summary:	
Integer	25%
Decimal	50%
Char	12.5%
Big int	12.5%
Char	
Integer	

Inferred Frequency Distribution			
Data Value	Data Type	#	%
refid,emp	Char	10	0.00
whs	Char	384	6.13
divt	Char	769	17.09

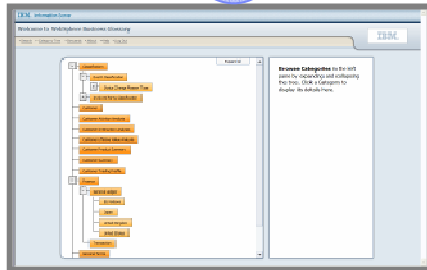
Type	Timestamp	Validity	Validity		Trend	Contact	Sample
			Severity	# Pass			
Run	01/01/09	✘	10.8%	792	208	✘	Mim Foster
Run	12/01/08	✔		904	96	✘	Mim Foster
Run	11/01/08	✔		926	74	✘	Mim Foster
Run	10/01/08	✔		500	0		Mim Foster
Run	09/01/08	✔		500	0		Mim Foster



1 Start from consistent blueprint, leveraging “best” practices

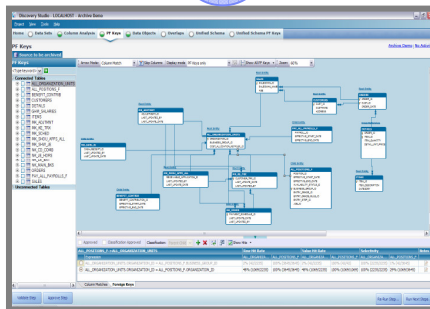


2



Document KPIs and associated business terms

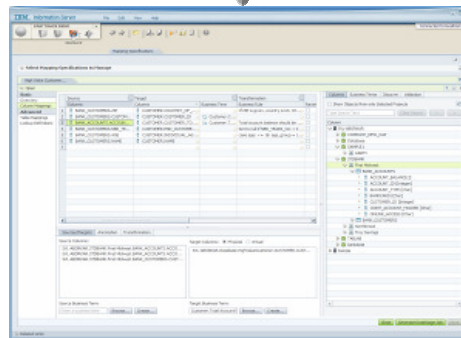
3



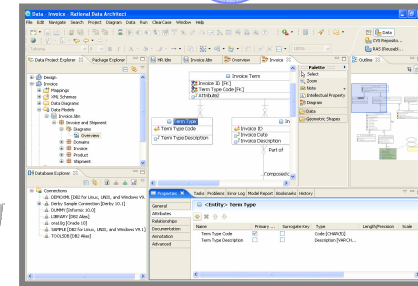
Identify where KPI's exist, relationship and level of quality

6

Create new transformation rules & document

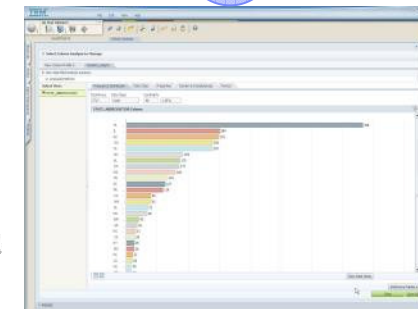


4



Create or Modify Data Model

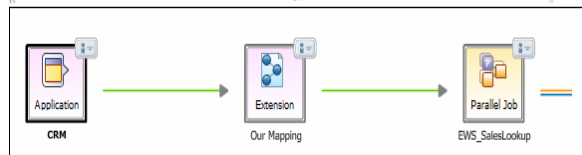
5



Assess & monitor your data quality over time

7

Report & Govern on metadata assets



Information Governance Resources

Recommended reading:

IBM Data Governance Unified Process

Sunil Soares – Director, Information Governance

<https://i2.infoprint.com/sales/catalogs.nsf/agswgcatalogint?openagent>

IBM Data Governance Council

<http://www.ibm.com/ibm/servicemanagement/us/en/data-governance.html>

Smarter technology for a smarter planet:

**Trusted data is the foundation
for a smarter organisation.**

