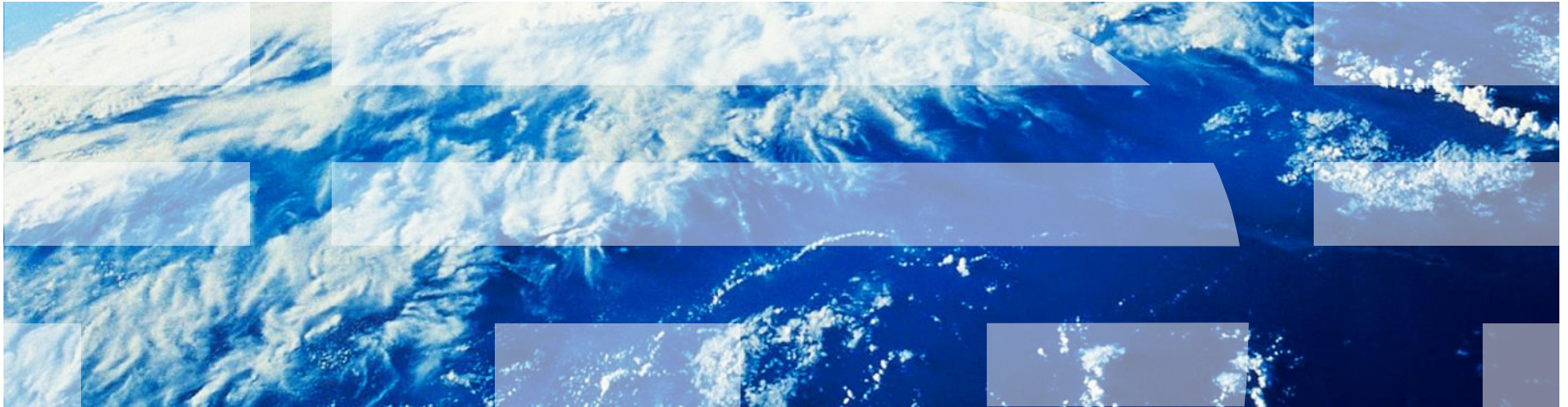


Value of IBM InfoSphere Information Server

Information Governance Catalog



Today's topics

- The overarching concept of Information Governance
- Overview of InfoSphere Information Governance Catalog (IGC)
- Information Governance Catalog Glossary Assets
 - Terms
 - Categories
 - Information Governance Policies
 - Information Governance Rules
- Governance and Stewardship
- Workflow
- Understand the Information Governance Catalog Glossary Interfaces
- Understand the Information Governance Catalog Assets
 - Queries
 - Collections and Labels
 - Lineage

IBM Analytics Portfolio: Power Behind and Across Your Data



Data
Science
& Cognitive
Learning

*Embrace new
ways to
develop
insights and
streamline
operations*



Unified
Governance

*Enable better
insight and
compliance
across all data*



Hybrid Data
Management

*Unify the
approach to data
and content on
the path to cloud*



Data Analytics
& Visualization

*Empower all to
make data-driven
decisions quickly
and easily*



Open Source: Hadoop, Spark & more
Commit to openness—for speed and innovation

Users Need Data That Is Accurate, Reliable and Trustworthy Across All Use Cases



Public



Enterprise
Data



Web/Mobile
Data



Social



IoT



Data Engineer



Data Scientist



Developer



Business Analyst



Data Steward

Businesses are challenged to solve key questions about the integrity of their data:

- Where is my data?
- Do I have the right data and context?
- How do I get a complete 360 view of my customers?
- How do I make data more trustworthy to my consumers?
- How do I better understand my data and fight against fraud?
- Who has access to this data?
- Do I understand the risk of using the data “as is”?

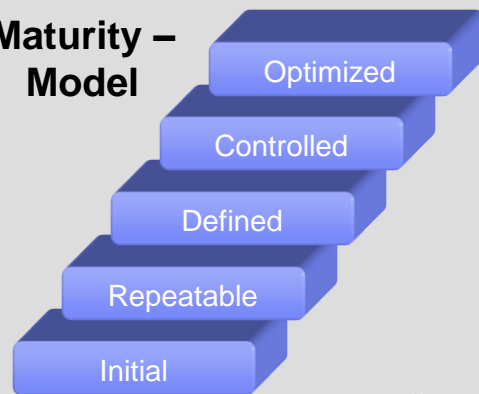
The Information Governance Community prepares definitions, approaches and concrete deliverables

Information Governance is ...
*the orchestration of **People, process** and **technology** in order to unleash corporate data and raise the **value** derived from it*

Information Governance Framework



Maturity – Model



Continuous improvement
of Best Practices and
Maturity Modell

Information Governance Maturity Model

IBM forms with ca. 40
other organization the
„**Information
Governance Council**“



Current trends and
requirements like
BCBS 239 and GDPR

Information
Governance in the era
of „**Big Data**“

Suggestions for future
risk prevention **based
on the experience
from the financial
crisis**

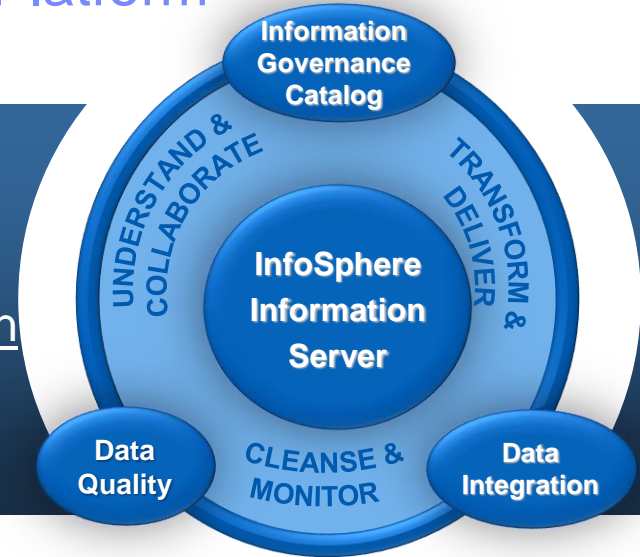
**Information
Governance
Framework**

Information Governance is based on eleven interconnected disciplines



Information Server: The Data Governance Platform

Integrating and transforming data and content to deliver accurate, consistent, timely and complete information on a single platform unified by a common metadata layer



Understand & Collaborate

„What is the meaning of...“?
„Where is ...?“

- Catalog technical metadata & align w/ business language
- Manage (big) data lineage
- Compliance reporting



Cleanse & Monitor

„Is the data correct?“ „What data am I looking at?“

- Analyze & validate w/ enhanced classification
- Cleanse & standardize
- Define, manage & monitor data rules + exceptions



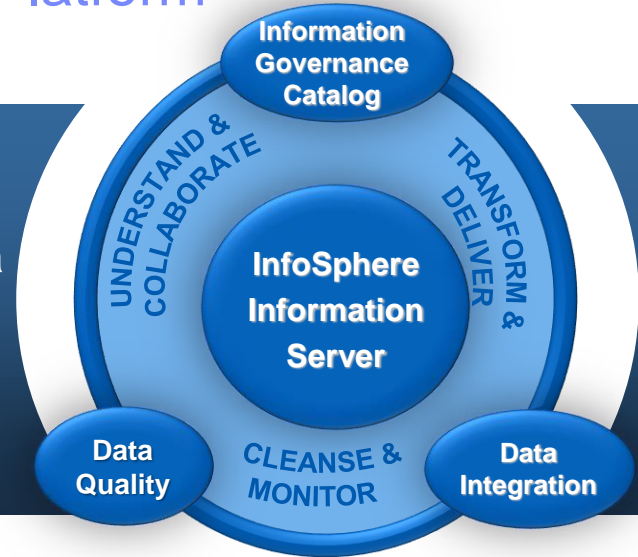
Transform & Deliver

„Can I get data out of SAP / DWH, ... into my Hadoop / DB, ...?“ „Can I cleanse that data?“

- Massive scalability
- Power for any complexity
- Deliver in batch and/or real-time with change capture

Information Server: The Data Governance Platform

Integrating and transforming data and content to deliver accurate, consistent, timely and complete information on a single platform unified by a common metadata layer



Understand & Collaborate

- Information Governance Catalog
- Business Glossary Anywhere
- Information Governance Dashboard and Reports



Cleanse & Monitor

- Information Analyzer
 - Quality Stage
- Stewardship Center

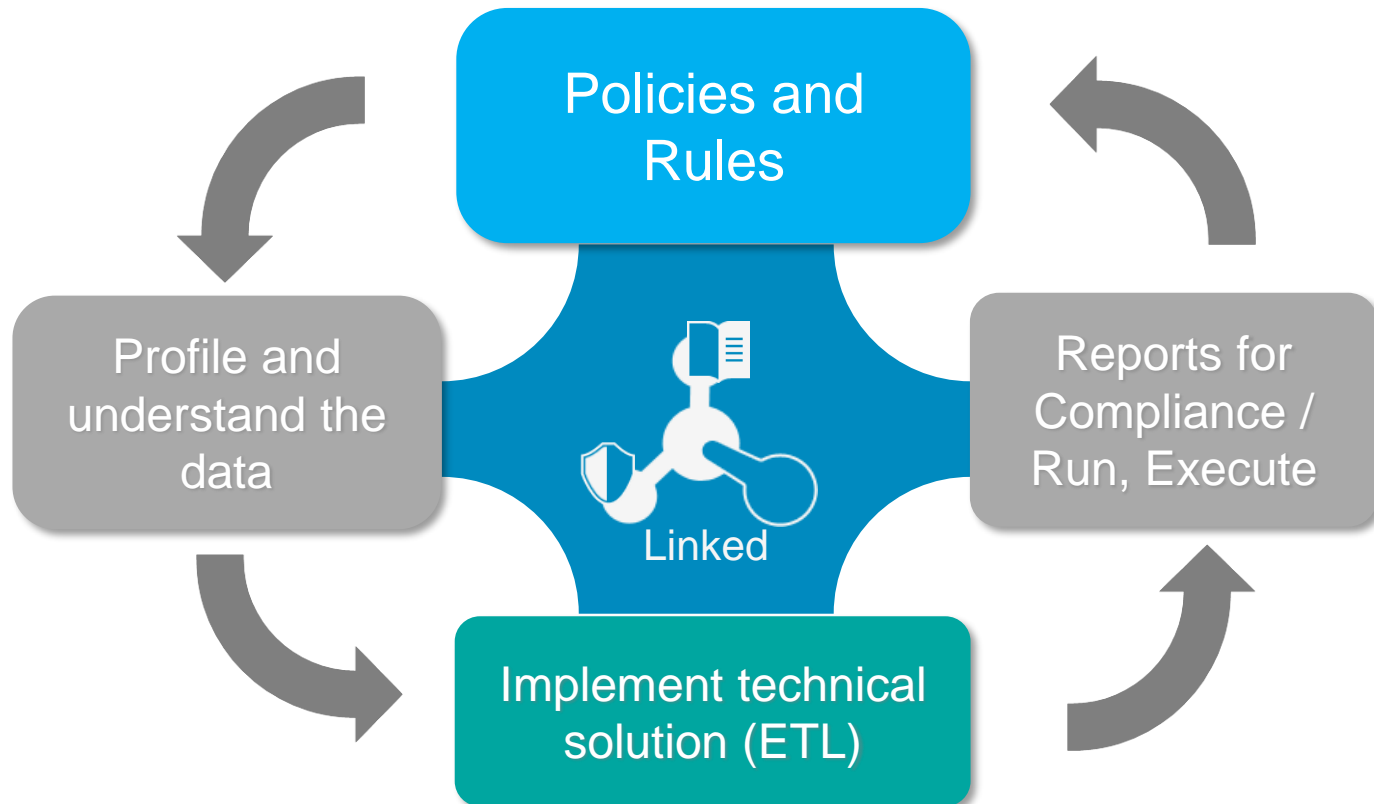


Transform & Deliver

- DataStage
- Big Integrate
- Enterprise Packs

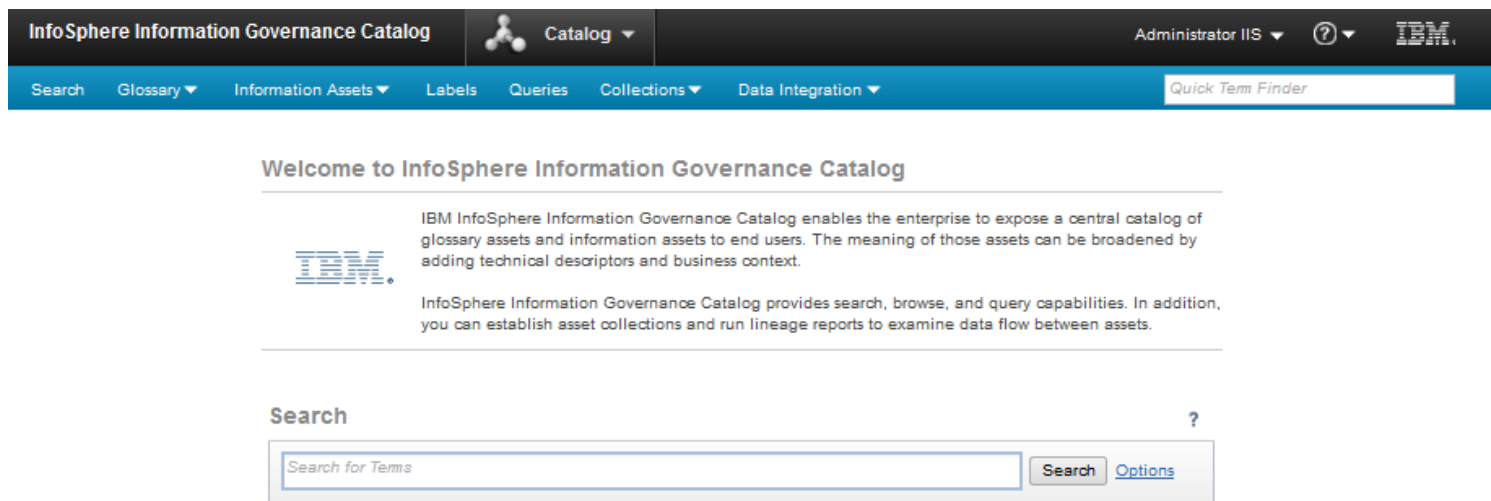
- Deployment options (on-prem, cloud, download & go) • common connectivity • shared metadata • security • common execution engine with flexible deployments (native on YARN)

IBM InfoSphere Information Server – Data Governance Lifecycle

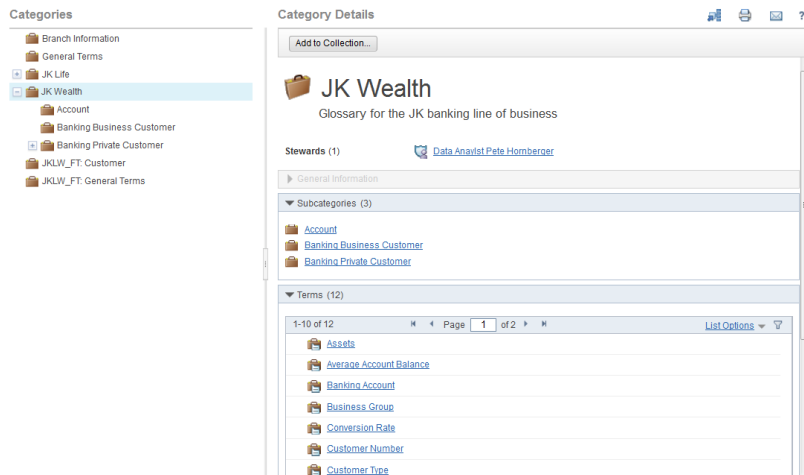


Information Governance Catalog

- A Web-based tool for business users that enables
 - The creation and management of a controlled vocabulary
 - Creation and management of a business taxonomy
 - Collaborative authoring of business metadata
- A reference for learning about the information assets of the enterprise
 - Meaning
 - Dependencies
 - Usage
 - Quality
 - Ownership/Responsibility
 - Data movement and lineage



Establish a common business language and manage business perspectives about information and align those views with the IT perspective.

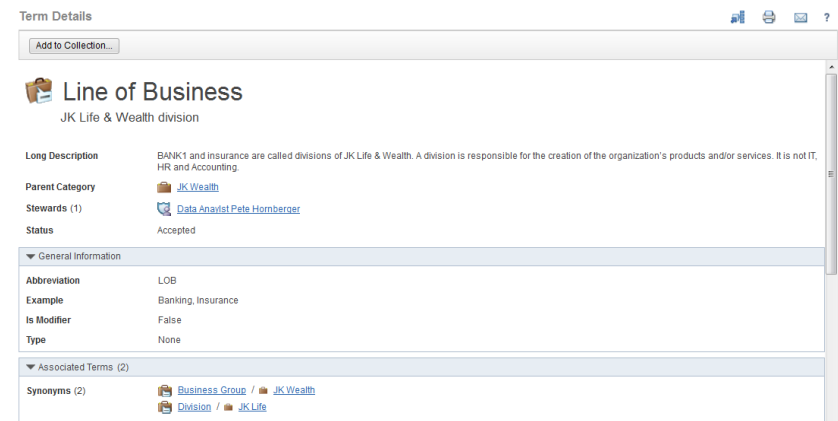


Terms

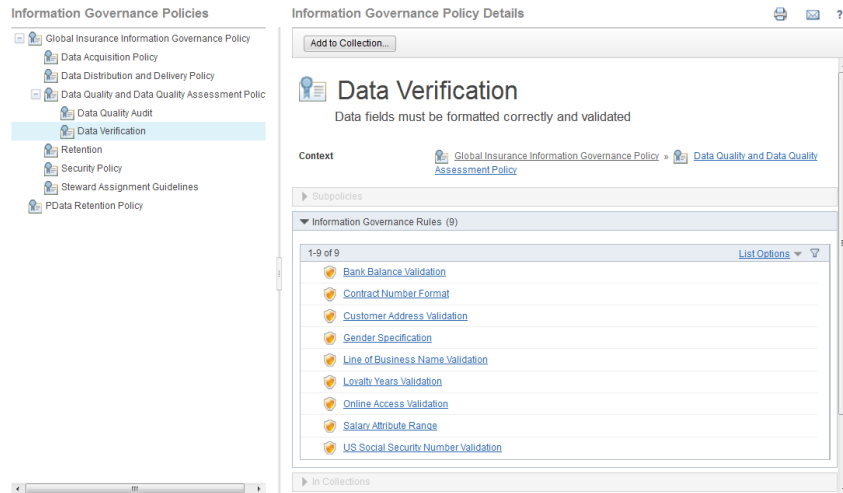
- A term is a word or phrase that describes a characteristic of the enterprise
- The properties of a term give it meaning and differentiate it from other terms
- IT assets can be assigned to terms to provide business context to that asset
- Relationships between terms can be defined to describe how a term can be related to another term, referenced by one or more categories other than its parent category, and can be replaced by another term

Categories

- Categories provide the logical structure for the business glossary
- Categories can contain other categories and terms
- Categories can reference terms that it does not contain
- Relationships between terms and categories



Establish governance policies and governance rules to describe the characteristics for making information resources compliant with corporate objectives.

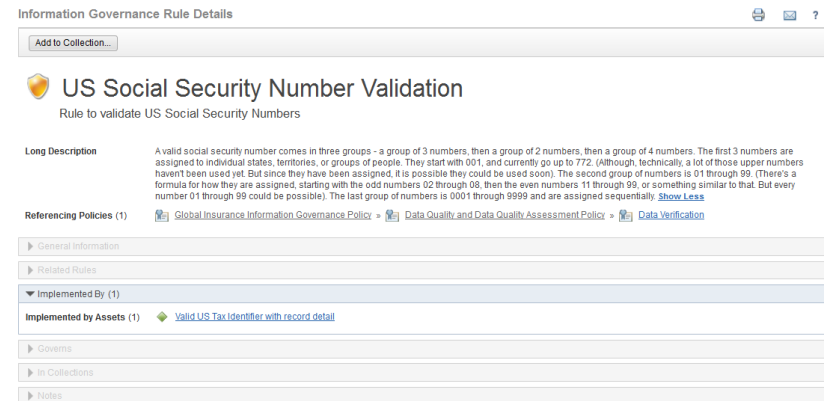


Governance Rules

- An information governance rule is a description of the criteria that is used to determine whether information assets are compliant with business objectives
- Information governance rules can be used as specifications for data governance operational rules
- Information governance rules can govern assets, an asset behaves or should behave as specified by the information governance rule
- An information governance rule can be related to one or more other information governance rules

Governance Policies

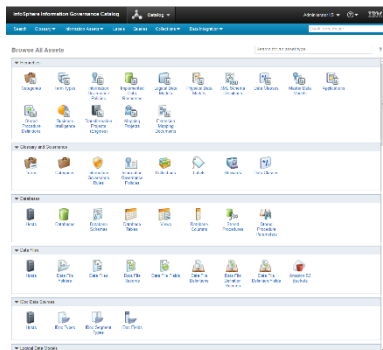
- An information governance policy is a description of a governance subject area
- Information governance policies and information governance rules describe the characteristics for making information resources compliant with corporate objectives
- Information governance policies can be organized in a hierarchy based on their meaning and relationships to one another
- Information governance policies can specify contained sub-policies, referenced rules, stewards, and other properties



CREATED BY: QA Release Manager Phillip Wood CREATED ON: Jul 3, 2014 10:58:59 AM MODIFIED BY: Data Analyst Pete Homberger MODIFIED ON: Aug 17, 2014 5:34:08 PM

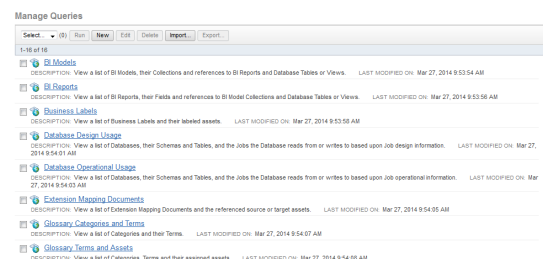
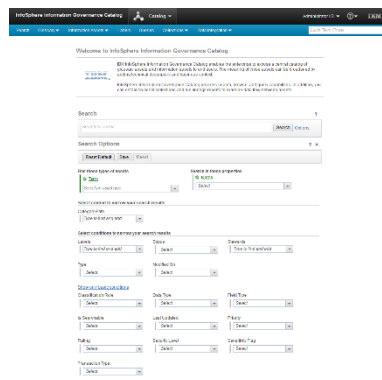
Exploration of assets in the catalog

- Explore assets in the catalog:
 - Explore glossary assets, information assets, and external assets, as well as stewards, and other catalog content
 - Provide simple and advanced search and robust querying
 - Display a graphical view of asset relationships
 - Track the history of changes to terms

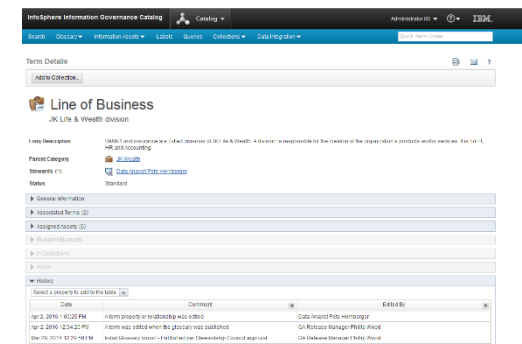


Explore

Search



Queries



Term History

With InfoSphere Information Catalog, you can...

- Enhance existing metadata with descriptions and associations:
 - Associate related assets, stewards, labels, and custom attributes to assets

▼ Associated Terms

Is a Type Of (0)

Has Types (0)

Is Of (0)

Has A (0)

Synonyms (2)

1-2 of 2 [List Options](#)

☐ Business Group / JK Wealth

☐ Division / JK Life

Preferred Synonym

Related Terms (0)

Replaces (0)

Replaced By

Assigned Terms (0)

Assigned to Terms (0)

▼ Header

* Name

Short Description

Long Description BANK1, BANK2, and insurance are called divisions of JK Life & Wealth. A division is responsible for the creation of the organization's products and/or services. It is not IT, HR and Accounting.

* Parent Category

Referencing Categories (0)

Labels (0)

Stewards (1)

1-1 of 1 [List Options](#)

☐ Data Analyst Pete Homberger

* Status

Governed by Rules (0)

Associate Related Assets

Stewards, Labels, Custom Attributes

▼ Assigned Assets

Assigned Assets (6)

1-6 of 6 [List Options](#)

☐ CST_LOB_FACT / IS-SERVER.IBM.COM » JKLW_EDW » JKLW_EDW

☐ CST_MST_FACT / IS-SERVER.IBM.COM » JKLW_EDW » JKLW_EDW

☐ JK_BANK1 / IS-SERVER.IBM.COM » JKLW_DBS

☐ JK_BANK2 / IS-SERVER.IBM.COM » JKLW_DBS

☐ JKLW_Customer_Report_4 / is-server.ibm.com » Public Folders » JKLW Customer Analysis Package

☐ LOB / IS-SERVER.IBM.COM » JKLW_EDW » JKLW_EDW

Assigned External Assets (0)

► Blueprint Elements

Assigned Assets

Assign business terms to technical metadata

Business Term

Termineinlage

Termineinlagen sind kurz- bis mittelfristige Geldanlagen bei Kreditinstituten, bei denen die Laufzeit oder Kündigungsfrist mindestens einen Monat beträgt.

Ausführliche Beschreibung Je nachdem, ob der Kunde mit der Bank eine feste Laufzeit oder eine bestimmte Kündigungsfrist für seine Geldanlage vereinbart, unterscheidet man zwischen Festgeld und Kündigungsgeld.

Übergeordnete Kategorie Banking » Einlagen

Status Kandidat

► Allgemeine Informationen

▼ Zugehörige Terme (1)

Synonyme (1) [Termingeld](#) / Banking » Einlagen

▼ Zugewiesene Assets (1)

Assigned Assets (1) [Termineinlage](#) / WIN7-X64 » SAMPLE » BANK

Table

Termineinlage

Kontext WIN7-X64 » SAMPLE » BANK

Termen zugewiesen (1) [Termineinlage](#) / Banking » Einlagen

► Datenbankspalten (2)

▼ Implementierungsdetails (1)

Implements Design Tables or Design Views (1) [Termineinlage](#) / BANK

► Allgemeine Informationen

▼ Schlüssel, Indizes und Analyse (2)

Defined Primary Key (1) [EINL_ID](#)

Defined Foreign Key (1) [EINL_ID](#)

Link zu synonymem fachlichen Begriff

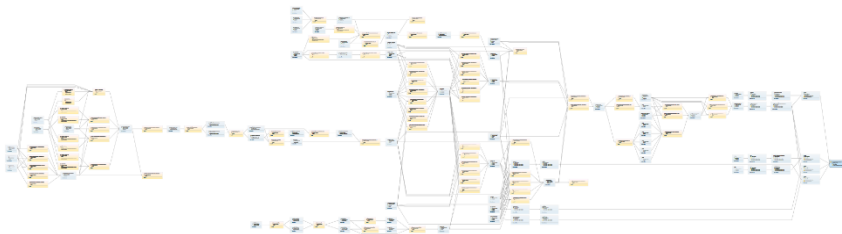
Link von fachlichem Begriff zu technischem Metadatum

Link von technischem Metadatum zu fachlichem Begriff

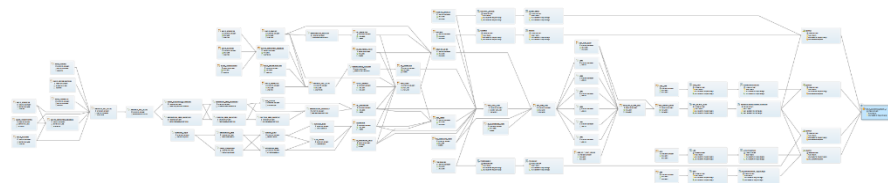
Link zu Modell-Entität

Analysis of Dependencies

- Analyze dependencies and relationships of key information assets and business intelligence reports:
 - Trace lineage from jobs and databases to business intelligence reports
 - Understand columns, database tables, and other assets
 - Perform lineage analysis to understand where data comes from or goes to by using shared table information, job design information, or operational metadata from job runs
 - Perform impact analysis to understand dependencies and the effects of changes to a column or job
 - Analyze operational metadata from job runs and report on rows that are written and read, and the success or failure of events

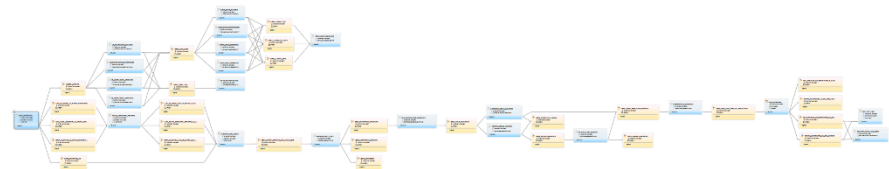


Technical and Operational Lineage



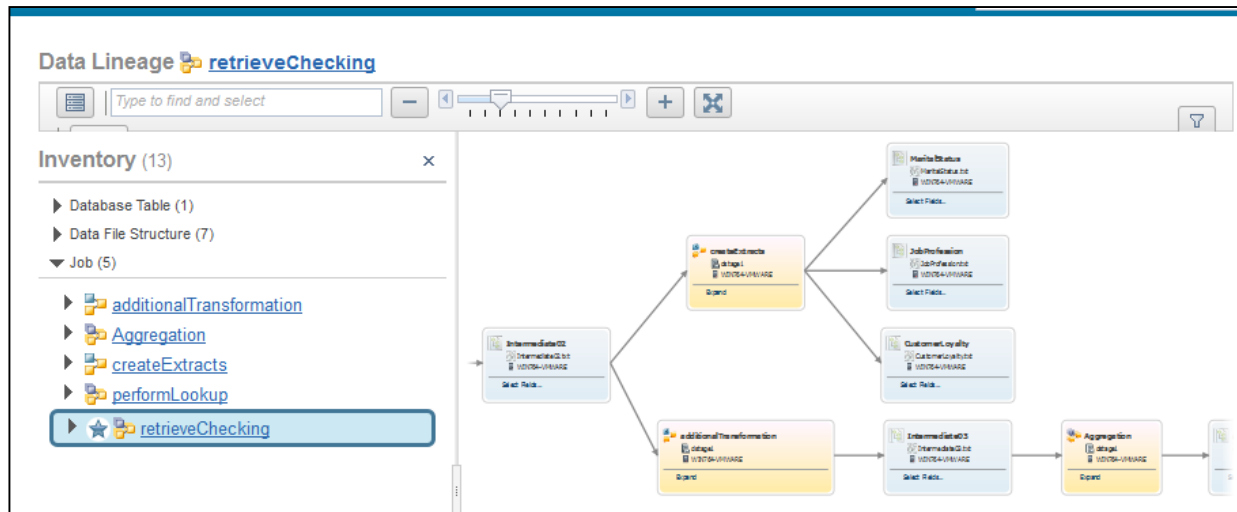
Business Lineage

Impact Analysis



With InfoSphere Information Catalog, you can...

- Manage the catalog metadata to obtain in-depth analysis reports:
 - Reconcile duplicate assets
 - Map databases to database aliases
 - Access runtime information to enrich reporting
 - Automate the discovery of relationships within and across data sources to accelerate project deployment
 - Run data lineage to create trusted information that supports data governance and compliance efforts



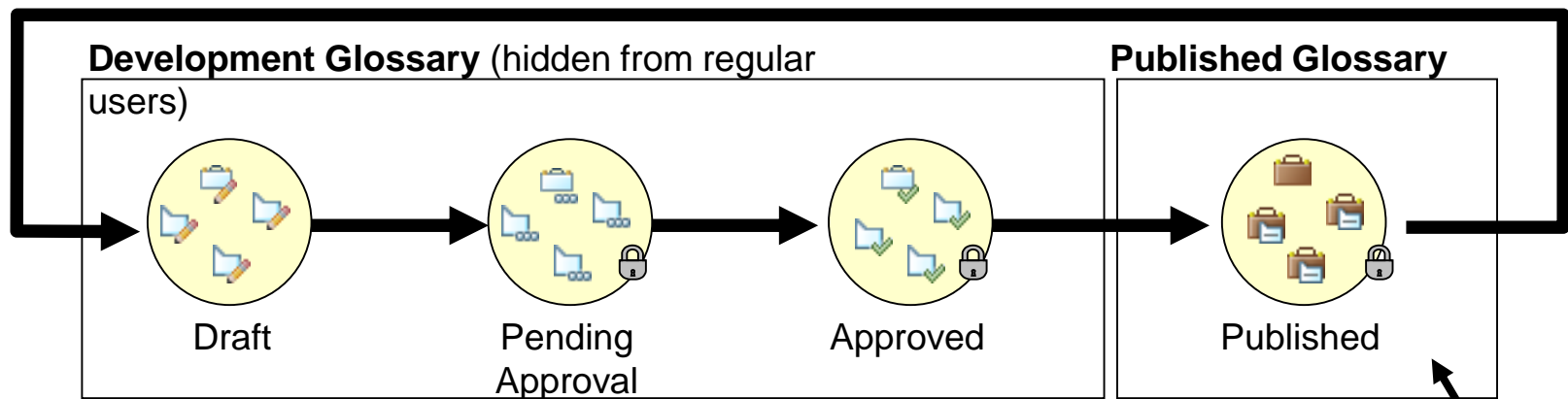
*Simple to
understand and
use with no
training required!*

Workflow based collaboration

- Manage the people and processes that govern your assets
 - Define the people responsible for creating, reviewing, and publishing information governance rules
 - Receive email notification of governance events that occur to your glossary and information assets
 - Approve a new or modified information governance rule, or return the rule to Draft status



Administrator enables workflow, designates participants, assigns areas of responsibility



Editor

Creates, Updates, Sends for Approval, Returns to Draft



Publisher

Approves, Returns to Draft, Publishes



User Searches, Browses, Explores

Governance

- Generally, information governance initiatives are deployed to promote the following goals
 - Increase consistency and confidence in decision making
 - Decrease the risk of regulatory fines
 - Improve data security
 - Provide consistent information quality across the organization
 - Maximize the income generation potential of data
 - Designate accountability for information quality
- The Information Governance Catalog is a cornerstone for governance
- Create, preserve, and disseminate knowledge about information (data)
 - What is information
 - How is information used
 - Who uses the information
 - Where is the information coming from
 - What happens to the information along the way
 - Where the information ends up

Data stewardship & governance

- Identifies
 - Who is responsible for this term?
 - Who is responsible for this category?
 - Who is responsible for this asset?
- Assign stewards to terms, categories or assets
- View contact information for steward

Steward Details

Jack Donaghy

Email Address jack@idw.com
User Name jack

▼ More Information

Given Name	Jack
Surname	Donaghy
Courtesy Title	Business Analyst

▼ Managed Assets (59)

Managed Assets (59) 1-10 of 59 Page 1 of 6 List Options

	Account Balance / JK Wealth > Account
	Account ID / JK Wealth > Account
	Account Name / JK Wealth > Account
	Account Title / JK Wealth > Account
	Account Type / JK Wealth > Account
	Actual Number Of Customers / JK Life
	Allow online access / JK Wealth > Account
	Allow teller machine access / JK Wealth > Account
	Assets / JK Wealth
	Average Account Balance / JK Wealth

Development glossary permissions

Apply to Category Apply Recursively

Customer

- Customer Attrition Analysis
- Customer Lifetime Value Analysis
- Customer Product Summary
- General Business Terms
- Transportation**
 - Air
 - Land
 - Sea

Management

- Workflow
- Development Glossary Permissions
- Published Glossary Permissions
- Stewards

CATEGORY: Transportation

☐ All users and groups have access
☒ Only selected users and groups have access

Select... 2 Selected

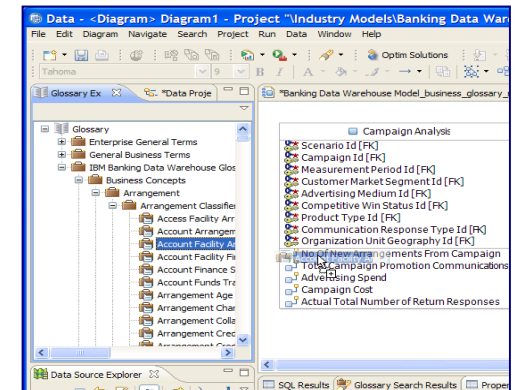
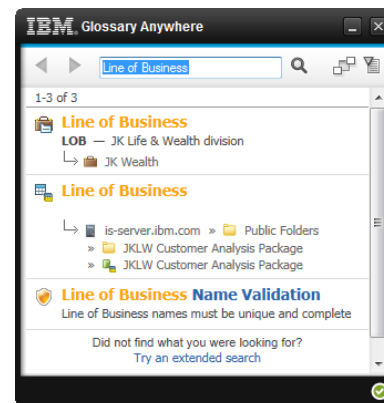
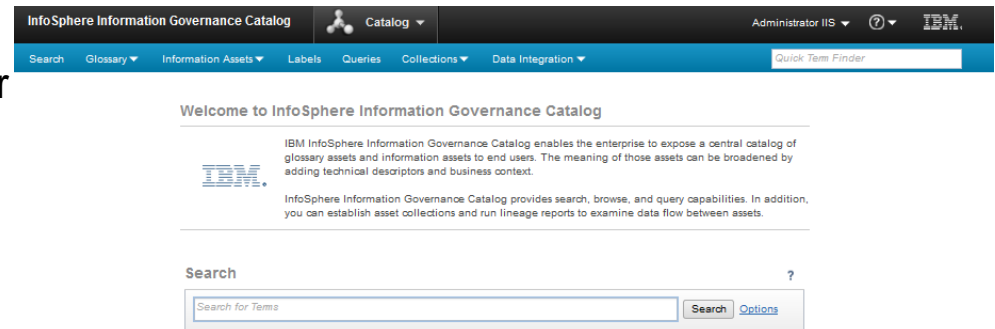
1-4 of 4 Page 1 of 1

<input type="checkbox"/>	Eddie Editor USER NAME: eddie
<input checked="" type="checkbox"/>	Edna Editrix USER NAME: edna
<input type="checkbox"/>	Pat Publisher USER NAME: pat
<input checked="" type="checkbox"/>	Paula Publisher USER NAME: paula

- Define access categories and avoid exposing sensitive information to unauthorized users

Information Governance Catalog - Glossary interfaces

- Information Governance Catalog Browser
- InfoSphere Glossary Anywhere (Read-Only)
- Other ways to access and edit the InfoSphere Business Glossary
 - InfoSphere Information Analyzer
 - InfoSphere FastTrack
 - InfoSphere Data Architect through InfoSphere Glossary Client for Eclipse
 - InfoSphere Information Governance Catalog REST Application Programming Interface (API)



Smart Hover

- Innovative technology for viewing metadata details without losing context of the object currently being viewed

Query Results: Job Design Information

Select... (0) Edit Add to Collection Detect Lineage Relationships

1-50 of 55 Page 1 of 2 List Options

Reads from Data Items Writes to Data Items

IDR2UC					
IDR2FILE_ACCOUNTS_BY_STATE_FileReaderJob					
JK_BANK2_CHECKING_LOYALTY_PROMO					
JK_STAGING_COMBINE_CST_INPUT_01_Extract_Data					
JK_STAGING_COMBINE_CST_INPUT_02_Investigate					
JK_STAGING_COMBINE_CST_INPUT_03_Standardize_Branches					

JK_STAGING_COMBINE_CST_INPUT_0...

Context IS-SERVER.IBM.COM » dstage1

Type Parallel

Image

Job Operational Information

dstage1	COMBINE_CST_INPUT
dstage1	COMBINE_CST_INPUT

Collections

- “Sets” of related assets
- Provides a direct “focus” for subject matter experts and teammates

The screenshot displays the IBM InfoSphere Information Server interface. On the left, a sidebar shows navigation options: 'Collections' (selected), 'Data Integration', 'My Collections', 'Collections Shared with Me', 'All Collections', and 'Create Collection...'. The 'My Collections' panel lists three collections: 'ATP009 2014', 'Finance', and 'RiskAggregation'. A line points from 'RiskAggregation' to the 'Collection Details' panel on the right.

The 'Collection Details' panel for 'RiskAggregation' shows the following information:

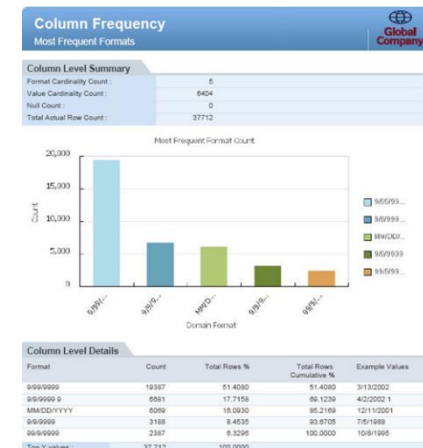
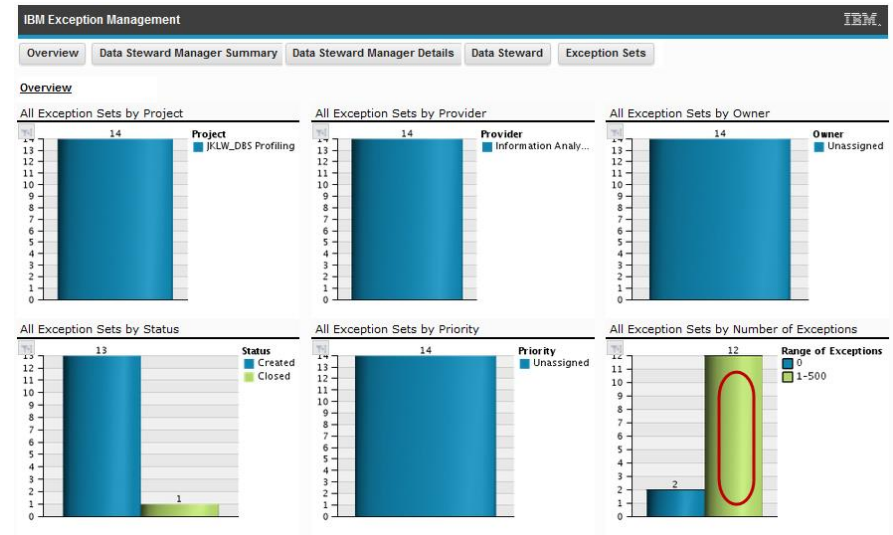
- Collection Details:** Edit, Add to Collection..., Delete
- RiskAggregation**
- General Information (3):**
 - Owners (2):** Administrator IIS, Data Analyst Kalpana Vijayakumar (circled in red)
 - Visibility:** Designated users and groups
 - Viewers (1):** Business Analyst Richard Smith
- Assets (7):**
 - Assets (7): Select... (0) Edit Data Click
 - 1-7 of 7
 - accident code / main » sub1
 - address / main » sub1
 - BANK ACCOUNTS / WIN764-VMWARE » DC

Main benefits of the Information Governance Catalog

- Affords a higher degree of trust by showing what happens to information
- Improves IT responsiveness and minimizes risk by allowing IT personnel to see the complete impact of any change
- Comply with regulations that require information traceability (Sarbanes-Oxley Act, Basel II)
- Improves the utility and manageability of IBM InfoSphere Information Server

InfoSphere Information Governance Dashboard - Reporting

- Governance Reporting Domains
 - Governance Performance
 - Data Quality
 - Information Recency
 - Metrics for Terms
- Drill Through Reporting
- Customizable
- **Report templates:** Over 80 out-of-the-box analysis reports as foundation
- **Control Output:** Schedule execution; determine expiration policy; establish access controls
- **Custom Processing:** API and CLI options to extract and utilize results in your environment and applications



InfoSphere Information Governance Catalog – Govern the data

- Facilitate business and IT communications by creating and managing a common business vocabulary
- Provides business context to information technology assets
 - Increased trust and acceptance of information by delivering definitions in context
- Establishes responsibility and accountability according to data governance policies
- Link business terms to technical assets
- Access glossary content and function through open interfaces and in a variety of tools
 - Import/Export
 - Representational State Transfer (REST), Eclipse, BG Anywhere
 - InfoSphere, Cognos, Rational, ...

The image displays several screenshots of the IBM InfoSphere Information Governance Catalog interface, highlighting key features:

- InfoSphere Information Governance Catalog**: The main dashboard with a search bar and navigation tabs for Search, Glossary, Information Assets, Labels, Queries, and Collections.
- Glossary Anywhere**: A floating window showing a search for "Line of Business" with results including "LOB - JK Life & Wealth division" and "JK Wealth".
- Taxonomy**: A view showing a hierarchical structure of categories, including "Branch Information", "General Terms", "JK Life", "Insurance Customer Details", "Address Information", "Insurance Policy", and "JK Wealth".
- Smart Hover**: A tooltip showing details for "JK Wealth", including "Banking Private Customer", "Short Description", "Parent Category", and "Stewards".
- Business Terms**: A list of terms such as "Assets", "Average Account Balance", "Banking Account", "Business Group", "Conversion Rate", "Customer Number", "Customer Type", "High Value Client", "High Value Customer", and "Line of Business".

InfoSphere Information Governance Catalog, ctd.

- Cross-suite view of Information Server assets
 - Assess and mitigate change management risk
 - Proactively manage complex data integration environments
- Support information governance with traceability on data movement, modeling and BI
 - Proactively manage complex data integration environments
- Assess dependencies across Information Server and non-IBM tools
 - Assess and mitigate change management risk
- Perform searches and build customized queries

The image displays three screenshots of the IBM InfoSphere Information Governance Catalog interface:

- Top Screenshot:** Shows the 'Information Governance Policies' sidebar on the left and the 'Data Verification' details pane on the right. The details pane lists 'Information Governance Rules' such as 'Bank Balance Validation', 'Contract Number Format', 'Customer Address Validation', 'Gender Specification', and 'Line of Business Name Validation'. A green box labeled 'Policies and Rules' highlights this section.
- Middle Screenshot:** Shows the 'Manage Queries' section with a list of query types: 'BI Models', 'BI Reports', 'Business Labels', 'Database Design Usage', and 'Database Operational Usage'. Each entry includes a description and a 'LAST MODIFIED ON' timestamp. A green box labeled 'Queries' highlights this section.
- Bottom Screenshot:** Shows the 'Data Lineage' view with a search bar and a list of 'Inventory (13)' items. The main area displays a complex lineage diagram with nodes like 'additionalTransformation', 'IntermediateTable', 'JobDefinition', 'Customization', 'IntermediateTable2', 'Aggregation', and 'Export'. A green box labeled 'Lineage' highlights the central part of the diagram.