



Collibra

Data Governance Software

Data Governance in Service Economies

handling quality, policy & risk of big data sharing across organisational boundaries

Dr. Pieter De Leenheer, Co-founder & Research Director



Vrije
Universiteit
Brussel

www.collibra.com @collibra contact@collibra.com

Context and Necessity

- **Services** are increasingly
 - knowledge-intensive relying on millions of data points from
 - Partners
 - Third parties
 - Customers
 - co-produced in federated, decentralised, multi-tier settings
- intersection of disciplines
 - Algorithms: *Big Data Analytics*
 - Infrastructure: *Internet of Things*
 - Service Innovation Methods: *Living Labs*
 - sufficient.....? **No**



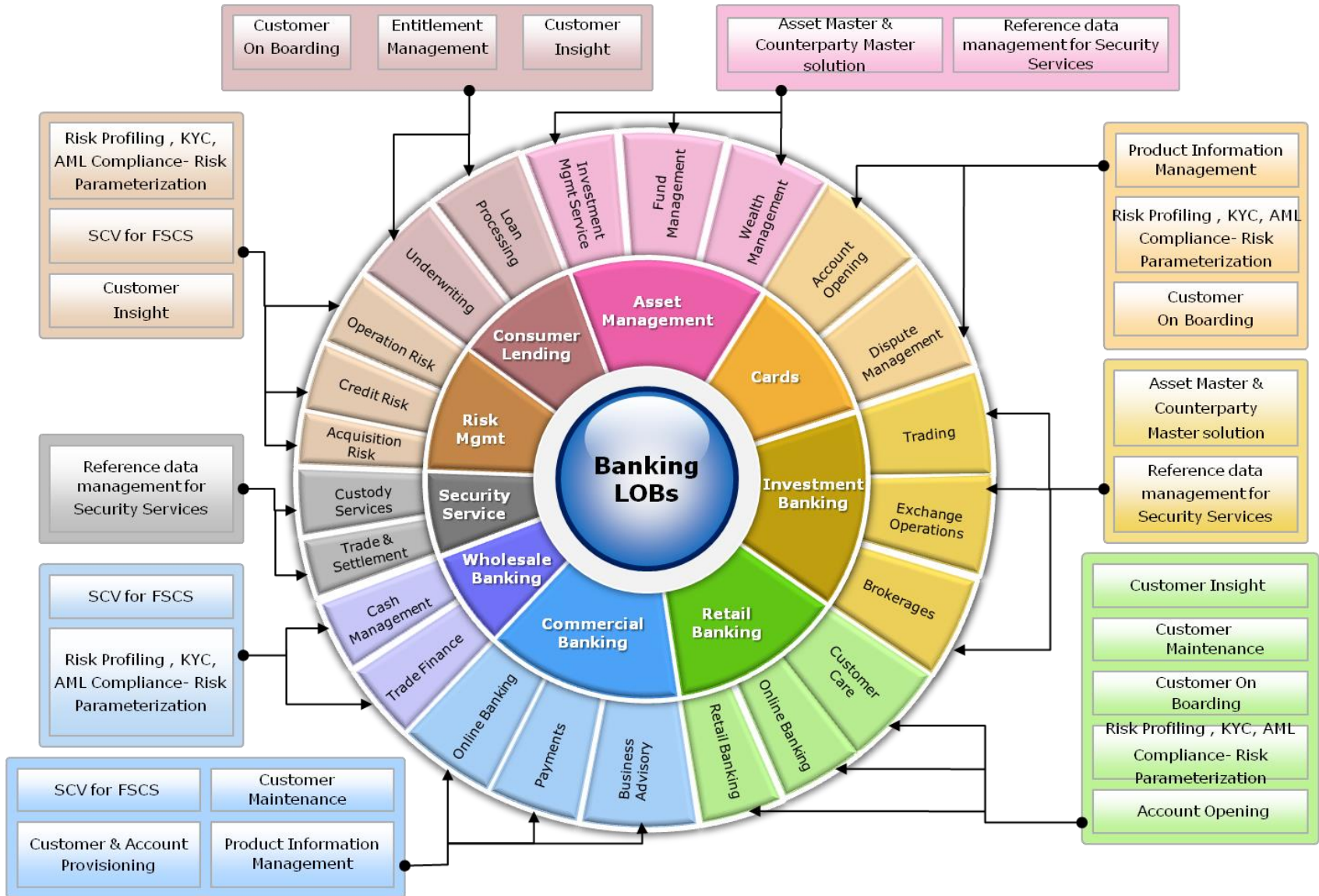
Data Management Challenges

- **Service** = data sharing agreement across organisation silos, policies, regulations, semantic assumptions
- **No clear line between data ownership and control:**
 - responsibilities are ignored
 - for each data point : exposure to risk regarding quality and policy compliance?
 - ask Alice, she knows

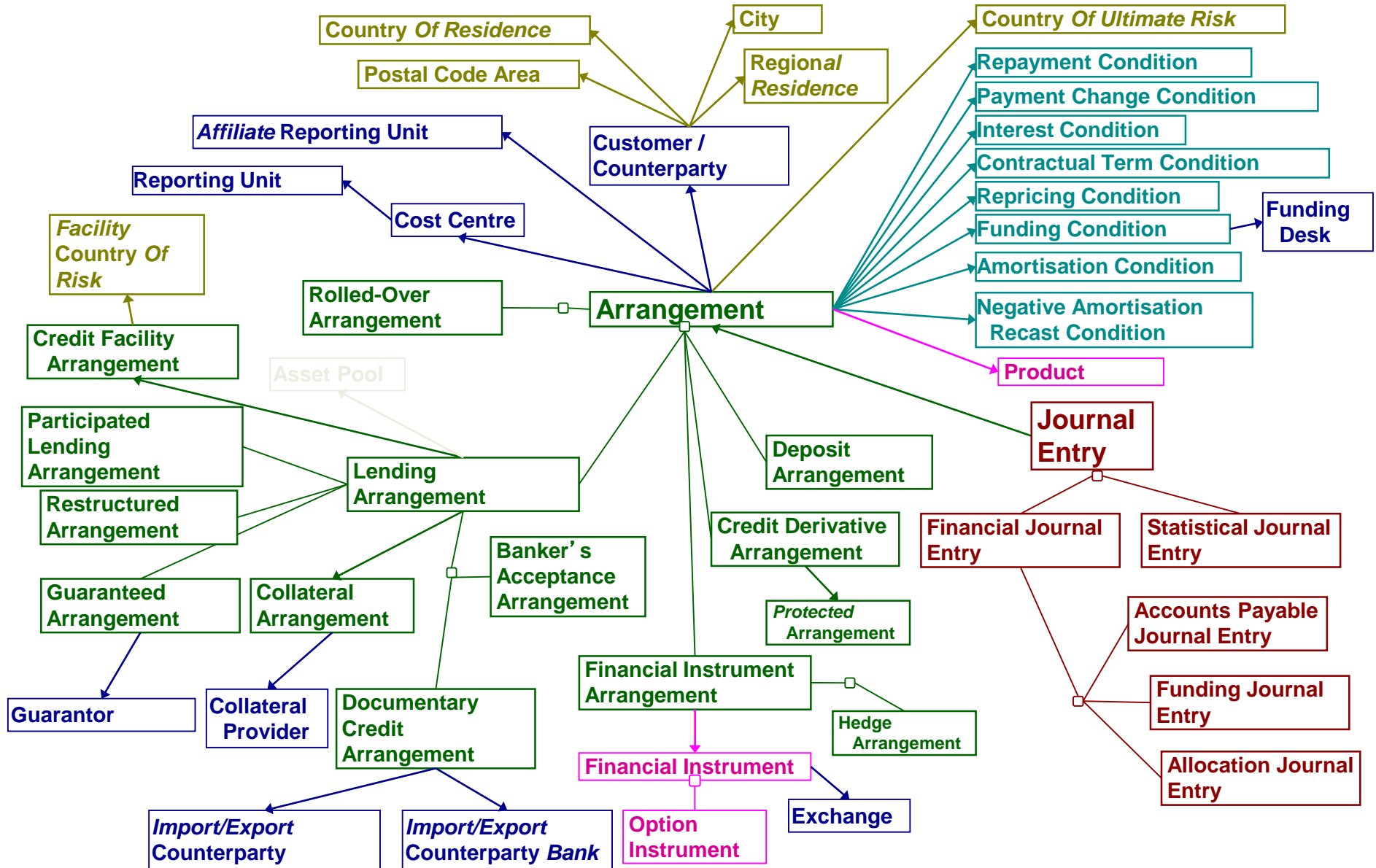
*Regulatory compliance risks continue to persist and remain a solid driver for governance, risk and compliance technologies. However, **more hype is being generated by external risks posed by third parties, suppliers and customers.***



The Wheel of Fortune: 360 View on Customer



...the reality: huge risk to create a 360 view without policy



Data Stewardship & Governance

- **Ownership**

=> Power + Control



Data Stewardship & Governance

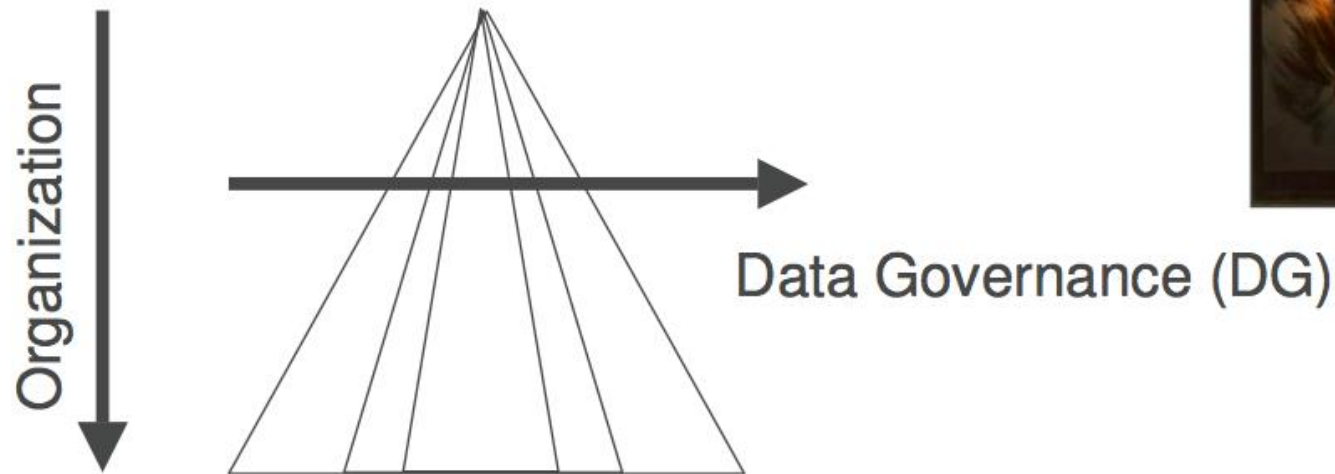
- **Ownership + Responsibility** ==> **Power + Control**
 1. (global) data stewardship
 - Req. 1: people who *define* data policy
 2. (systematic) data governance
 - Req. 2: processes that *enforce* data policy
- Now let's build software for it...

“New Information infrastructure technologies must enable organizations to define, organize, share, integrate and govern data and content to create business value”

(Gartner Hype Cycle on  Information Infrastructure Tech 2013).

Yet contradicting forces...

Borrowed from Dirk Coutuer (ING)



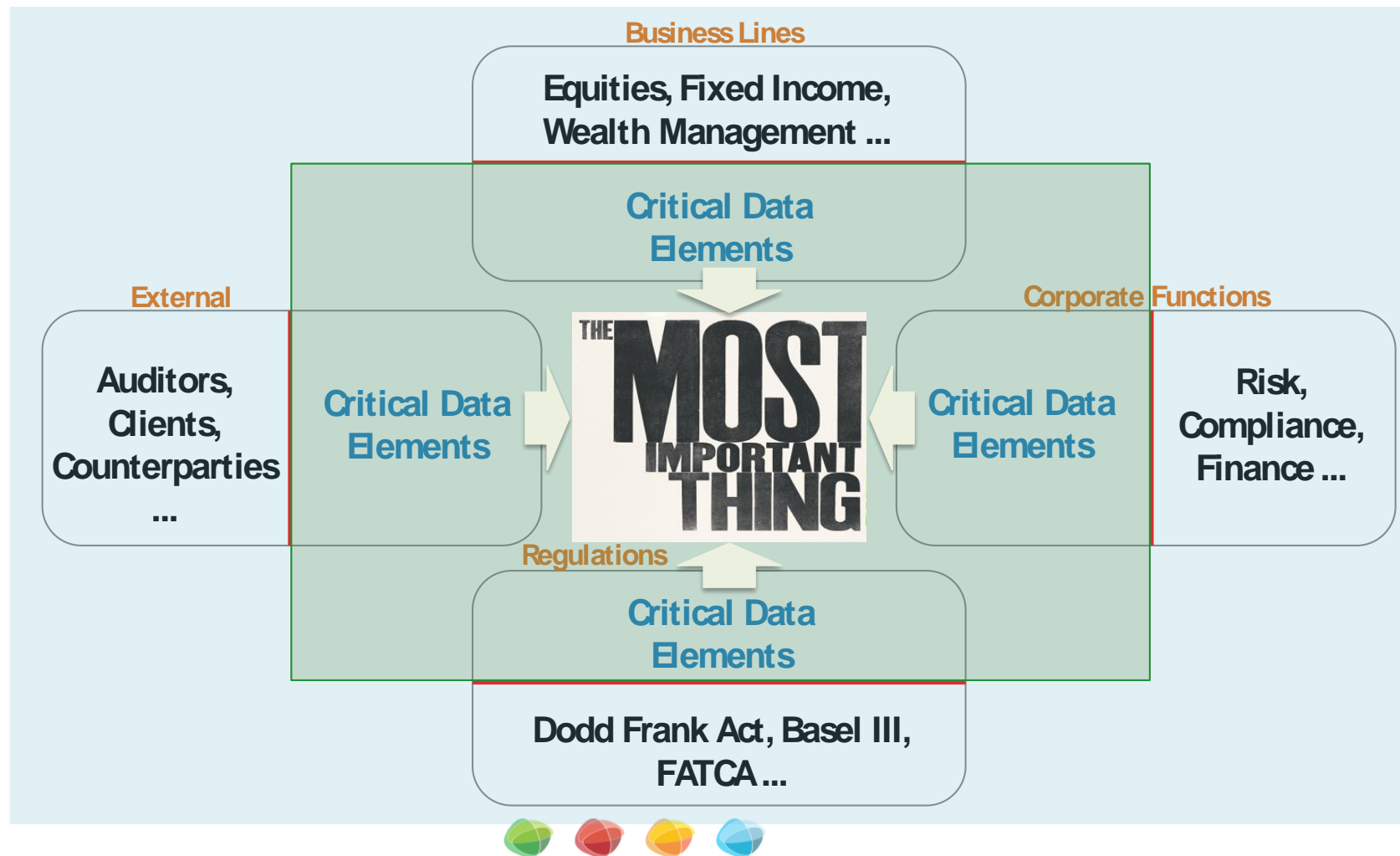
Even if there is an agreement on the added value of DG, the temptation is to continue to work (only) within the own silo

Agreement <> Commitment



..and some data points are more equal than others -> need for policy scoping too

Borrowed from Predrag Dizdarevic (Element 22 NYC)



Can technology *globalise* and *systematise* data policy *scoping*, definition and enforcement which is by nature a human **process**?



Collibra

Data Governance Software

Process-driven Data Governance

Collibra

2008
VC-backed Spin-off VUB
25 FTE + partners
HQ: Brussels
Offices: Wroclaw, NYC



Data Governance
Software Company

Reference Customers

STANDARD
& POOR'S



novo nordisk®

Johnson & Johnson

vmware®

VDAB



HSBC



colt

PriorityHealth



entsoe
Reliable Sustainable Connected



smarter / faster / further

BARRY CALLEBAUT

KYIVSTAR
THINKING OF YOU



eandis



Products



Data Stewardship
Manager

Business Semantics
Glossary

Reference Data
Accelerator

Data Governance
Center

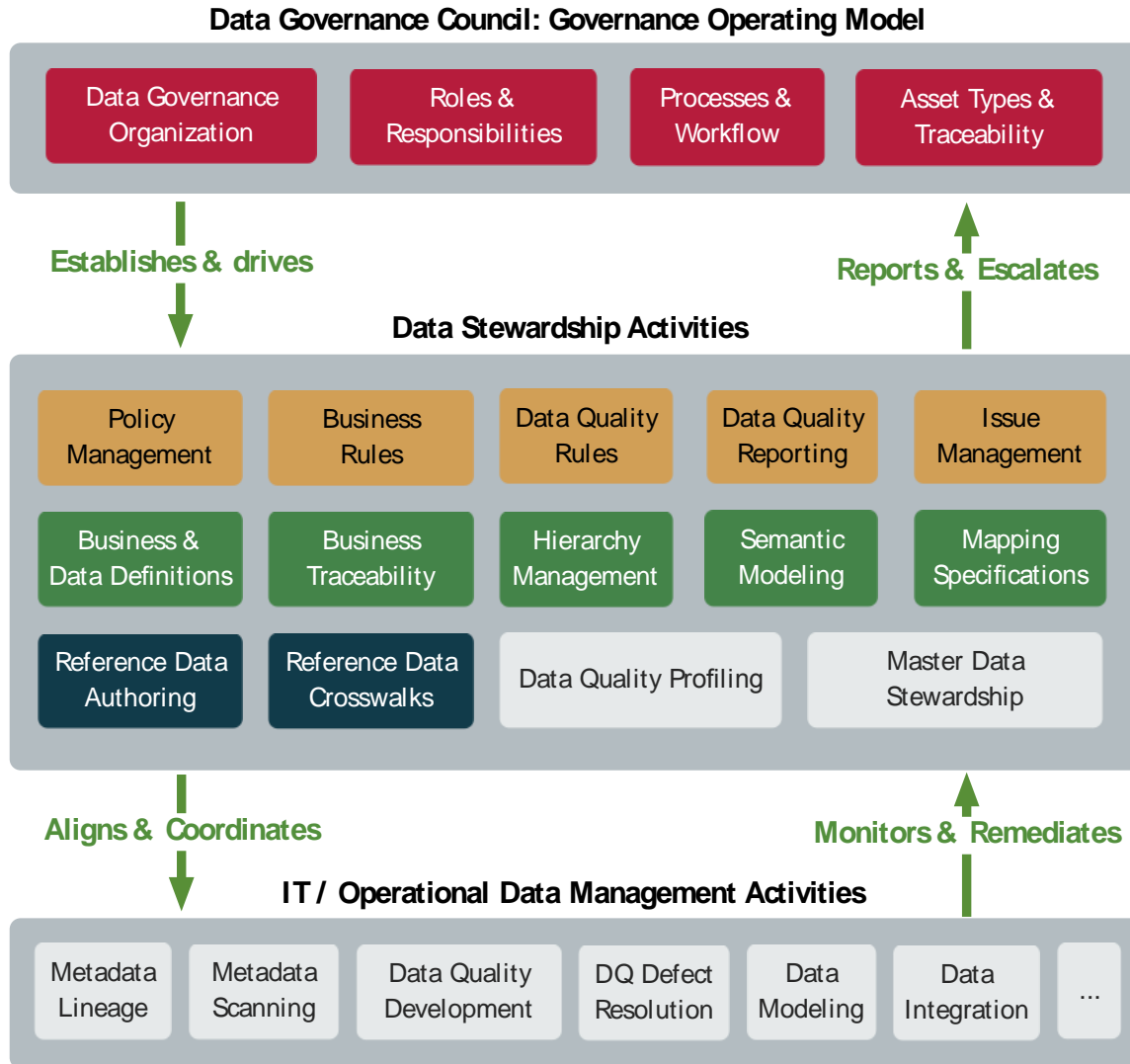


Load, Define & Enforce

Load...

Scope,
select,
define

enforce



- Collibra Platform
- Collibra Data Stewardship Manager (DSM)
- Collibra Business Semantics Glossary (BSG)
- Collibra Reference Data Accelerator (RDA)
- Other Data Management Vendor products



5 Modeling Concepts in DGC Operating Model

Domains logically group assets (according to their function, project, or knowledge area) and are owned by exactly one community. It has a domain type that specifies which asset types can be created in the domain.

Eg., Customer Domain groups all assets related to customer relationship management

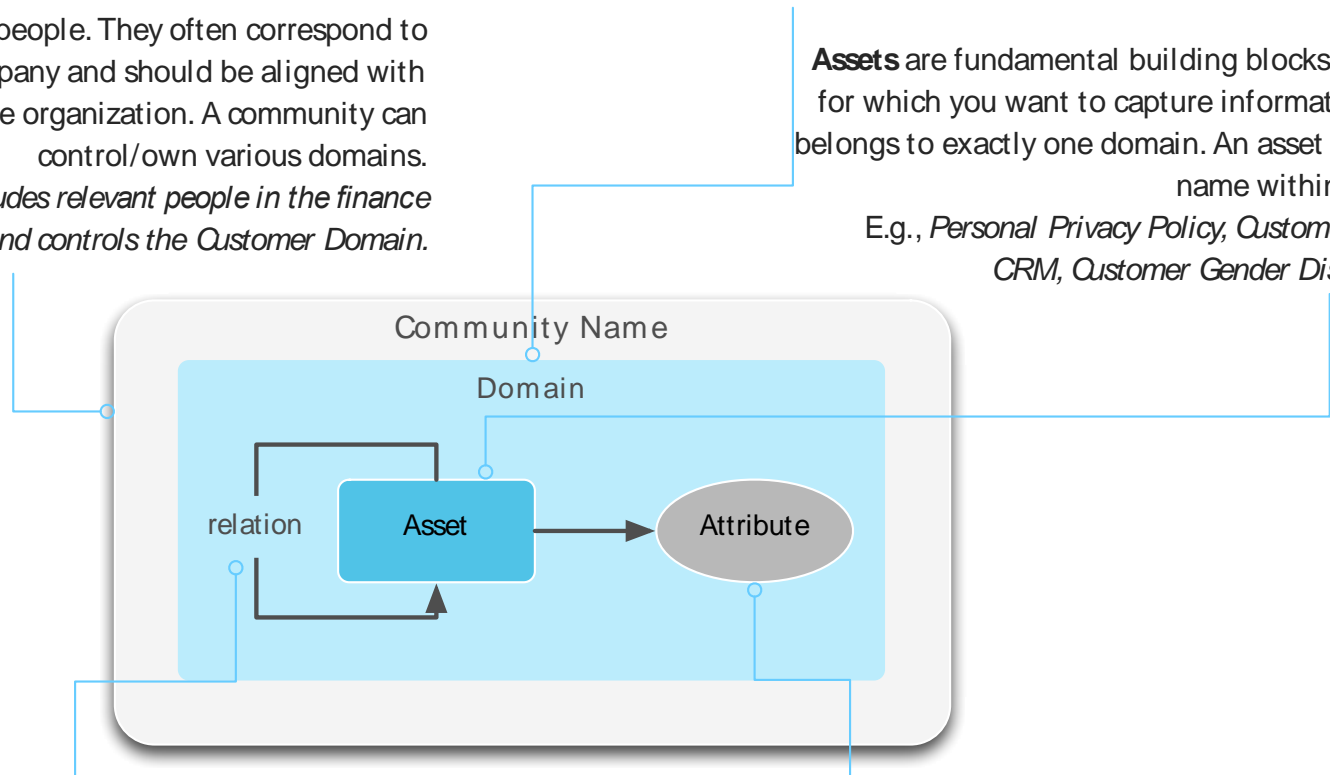
Eg., Enterprise Rules and Policies Domain collects all valid policies and rules in the organisation

Communities are groups of people. They often correspond to functional divisions in a company and should be aligned with the company's governance organization. A community can control/own various domains.

Eg., Finance Community includes relevant people in the finance function, and controls the Customer Domain.

Assets are fundamental building blocks or resources for which you want to capture information. An asset belongs to exactly one domain. An asset has a unique name within its domain..

E.g., Personal Privacy Policy, Customer, ISO 3166, CRM, Customer Gender Disclosure Issue



Relations semantically relate 2 assets

E.g., between assets "Customer" and "CRM": "Customer has system of record / is system of record for CRM"

E.g., between assets "Customer" and "Gender": "Customer has gender / gender of Gender"

Attributes are literal values such as strings or numbers that do not form an asset on their own right.

E.g., the Description attribute for asset "Customer" is "Person that placed at least one order for at least one product with Bank and Insurance"

DGC Asset Types

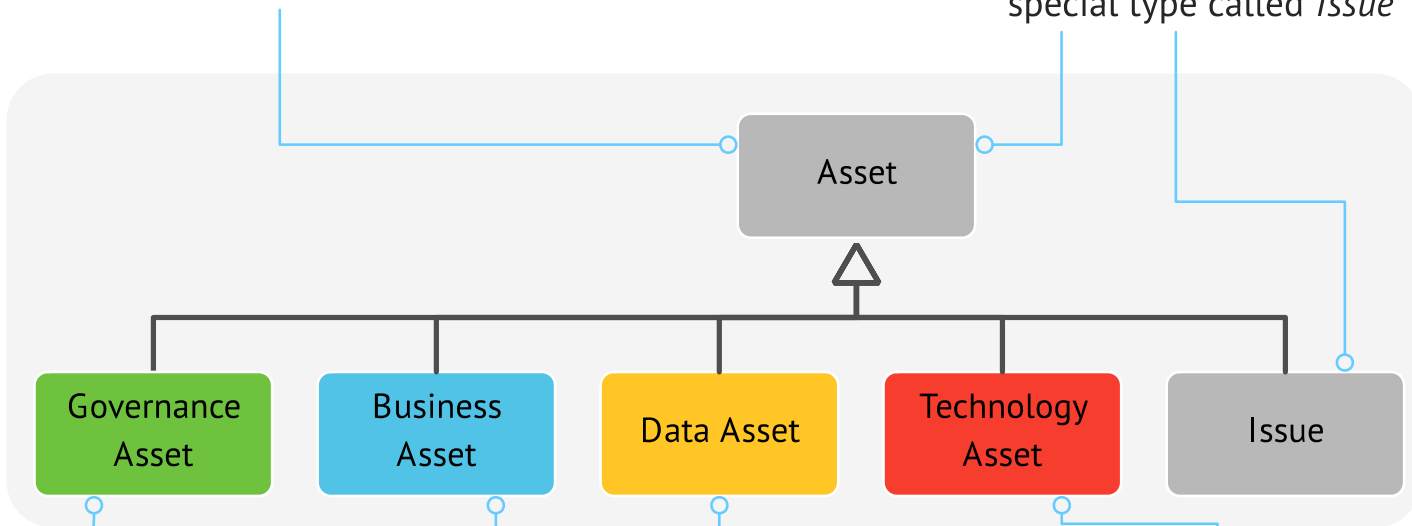
Asset Types allow you to formally specify what type an asset is, as a kind of template. They are assigned to one or more Domain Types.

E.g., Business Term is type for “Customer” and “Gender”

E.g., Code Value is type for “CG_NA”;

E.g., System is type for “CRM”

We distinguish between 4 main types of asset, and 1 special type called *Issue*



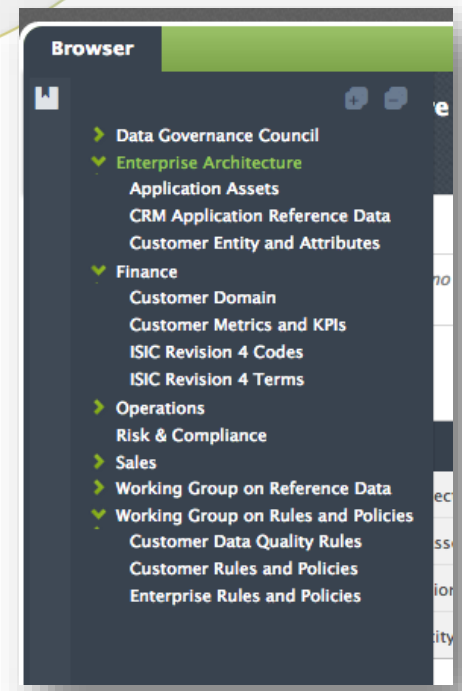
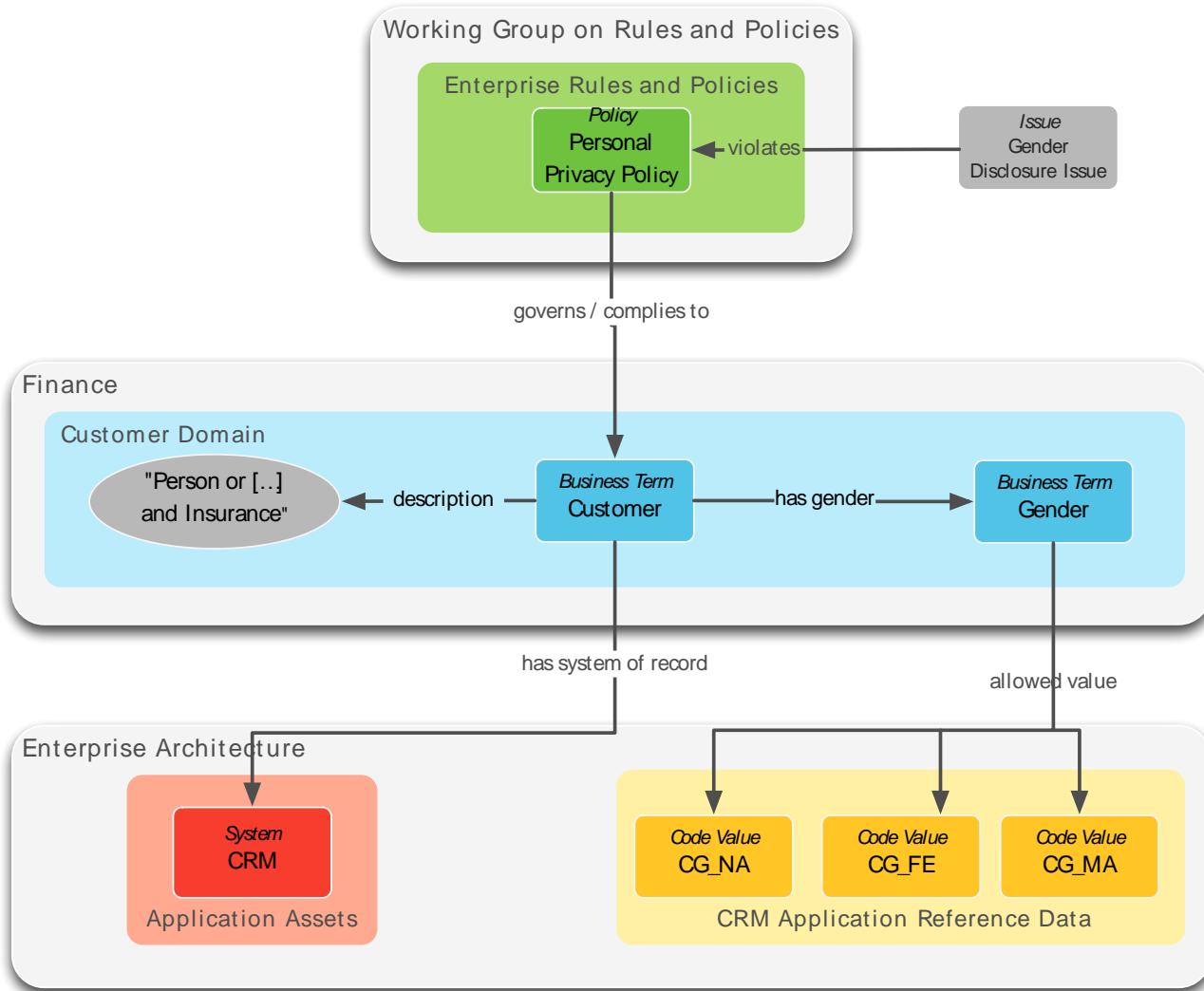
includes asset types such as *Policy* and *Rule*

subsumes asset types such as *Business Term*, *KPI*, and *Report*

includes asset types such as *System* and *Database*

subsumes asset types such as *Code Value*

Traceability of Assets across Domains



Assigning types to assets, relations, domains gives meaning; and brings a better understanding of different viewpoints on DG



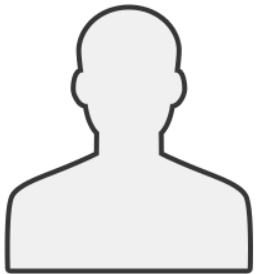
Collibra

Data Governance Software

Use-cases

Business Glossary

1



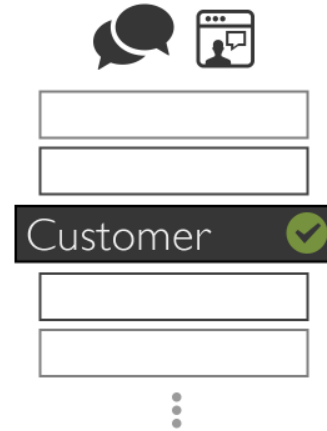
Import

2



Improve

3



Publish

4



Use





Announcements

Welcome! This Glossary is NetApp's single source of truth for business definitions. It is a vital step toward achieving our vision of commonly understood, consistent, trusted data throughout the enterprise.

1
2

Support

- Contact Us
- FAQs
- Domains & Stewards List
- Edit announcement
- Edit labels
- Edit completion metric
- Edit/Add Link Attachments

All Users



Index of Domains

Our key business concepts, each under the watchful eye of a Data Steward.



Full Index of Terms

All business concepts are defined here, organized alphabetically.



Glossary Metrics

Key measures that highlight the current status of the Glossary.

Data Stewards



My Steward Page

Login page for Data Stewards, granting read/write access to their domains.



Recent Changes

Audit trail for all updates made to the Glossary.



Glossary Metadata

Standard classifications applied across the Glossary.

You have

0 open tasks

Recent Changes

1 month ago



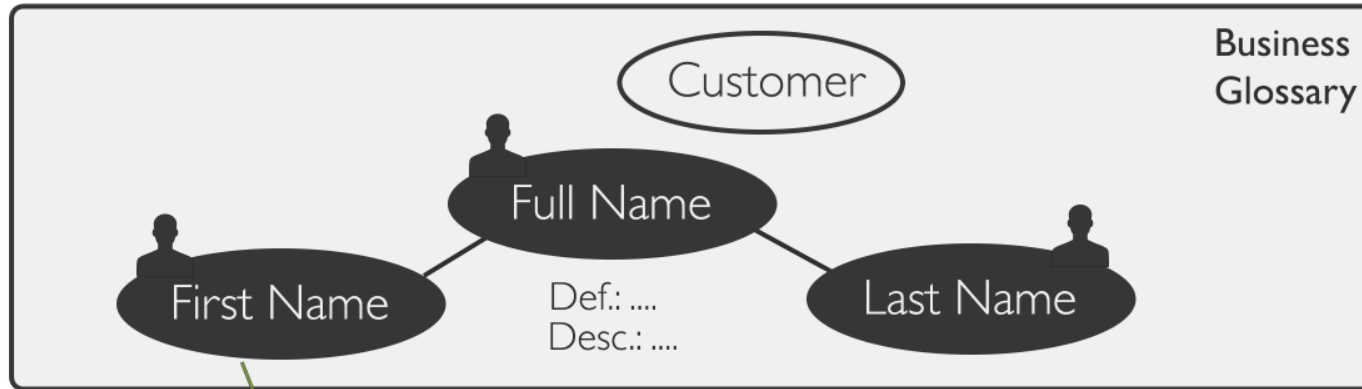
Support is 3-submitted for approval.



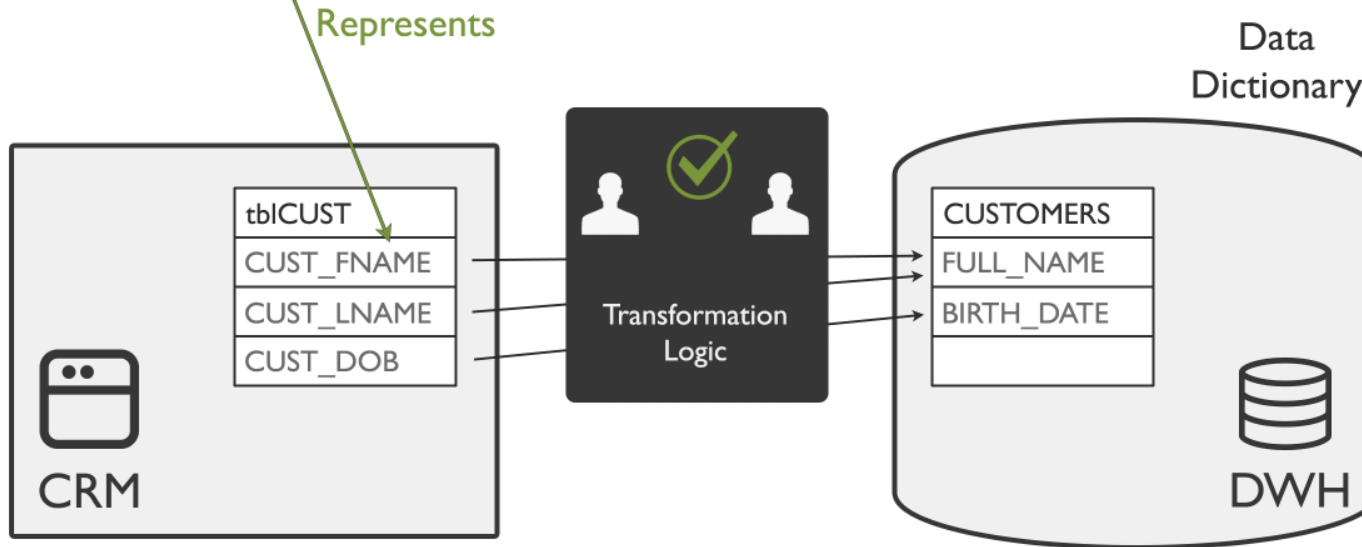
A term **Support 2** was added to support.



Data Dictionary



1
Define



2
Map

3
Visualize





Master Data Governance



Dashboard



Business Glossary



Reference Data



Data Stewardship



Settings



Account

Browser | Data Governance Documentation > ... > ... > |

Search...



Lab/Office new

Type: **Field** Status: **Accepted**

- Edit
- Delete
- Start working on this document
- Simple Approval
- Vote
- Approve Asset
- Favorite
- Import
- Export
- Share
- Print

Overview

+ Add

Responsibilities

Traceability

Definition

MARA-LABOR is a standard SAP field. It is often referred to as lab/office. It has always been in use at Barry Callebaut to indicate the brand of materials. Edit Delete
[Karen_Aelbrecht](#) 3/1/13

SAP Tag

Lab/Office Edit Delete
[Karen_Aelbrecht](#) 3/1/13

Content Rules

Principles Edit Delete
The lab/office field represents the brand that is used to sell a product. For most products, the brand is shown on the packaging.
To differentiate brands that are used in different markets (regional or business unit), a brand can be split. In such a case the business unit or region is mentioned in the value description.



Master Data Governance



Dashboard



Business Glossary



Reference Data



Data Stewardship



Settings



Account

Browser | Data Governance Documentation > ... > ... > |

Search...



Lab/Office new

Type: **Field** Status: **Accepted**

- Edit
- Delete
- Start working on this document
- Simple Approval
- Vote
- Approve Asset
- Favorite
- Import
- Export
- Share
- Print

Overview

+ Add

Responsibilities

Traceability

Definition

MARA-LABOR is a standard SAP field. It is often referred to as lab/office. It has always been in use at Barry Callebaut to indicate the brand of materials. Edit Delete
[Karen_Aelbrecht](#) 3/1/13

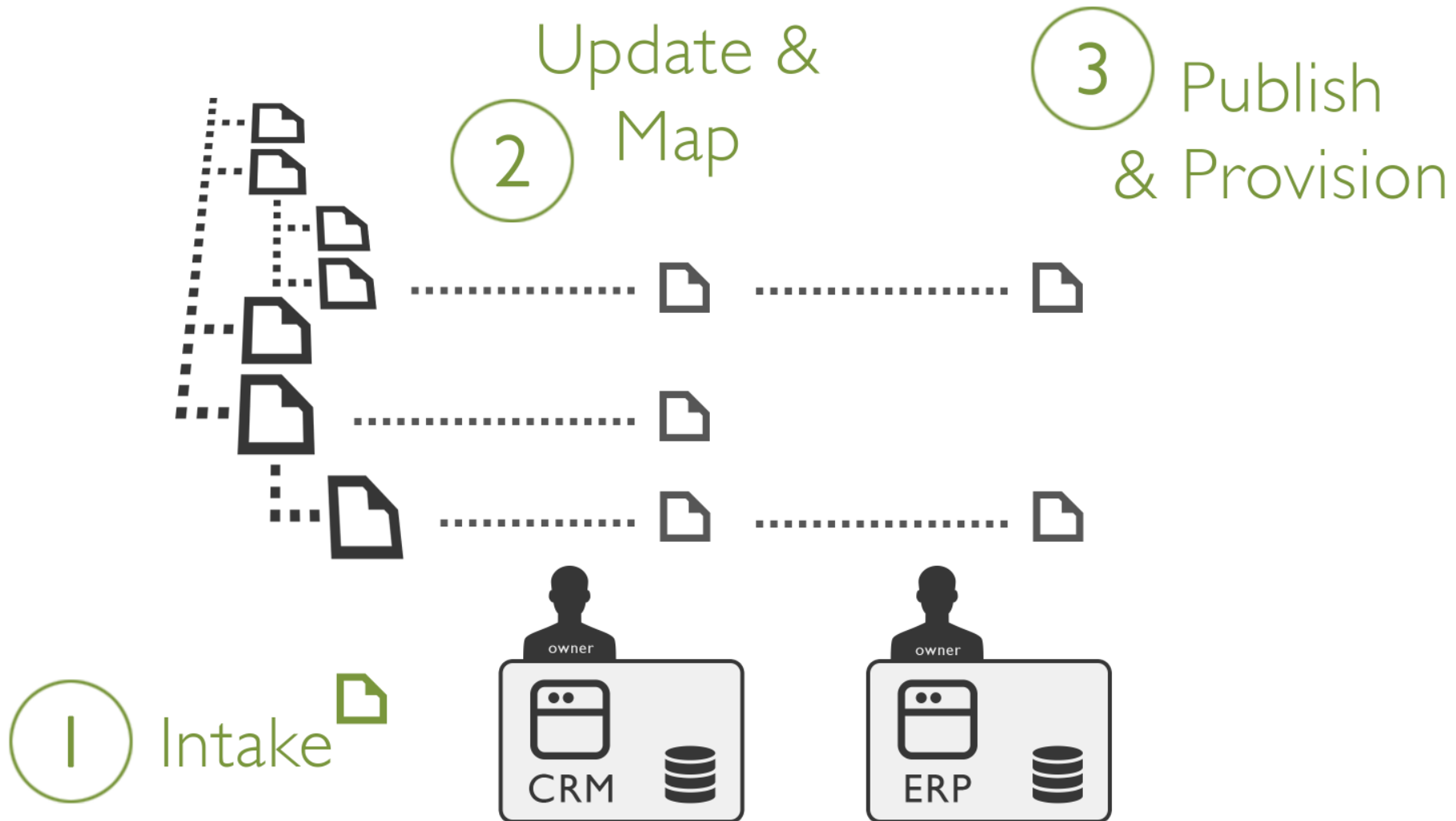
SAP Tag

Lab/Office Edit Delete
[Karen_Aelbrecht](#) 3/1/13

Content Rules

Principles Edit Delete
The lab/office field represents the brand that is used to sell a product. For most products, the brand is shown on the packaging.
To differentiate brands that are used in different markets (regional or business unit), a brand can be split. In such a case the business unit or region is mentioned in the value description.

Reference Data



Issue Management

1 Report



2 Resolve



3 Notify



Triage & Review



B Escalate

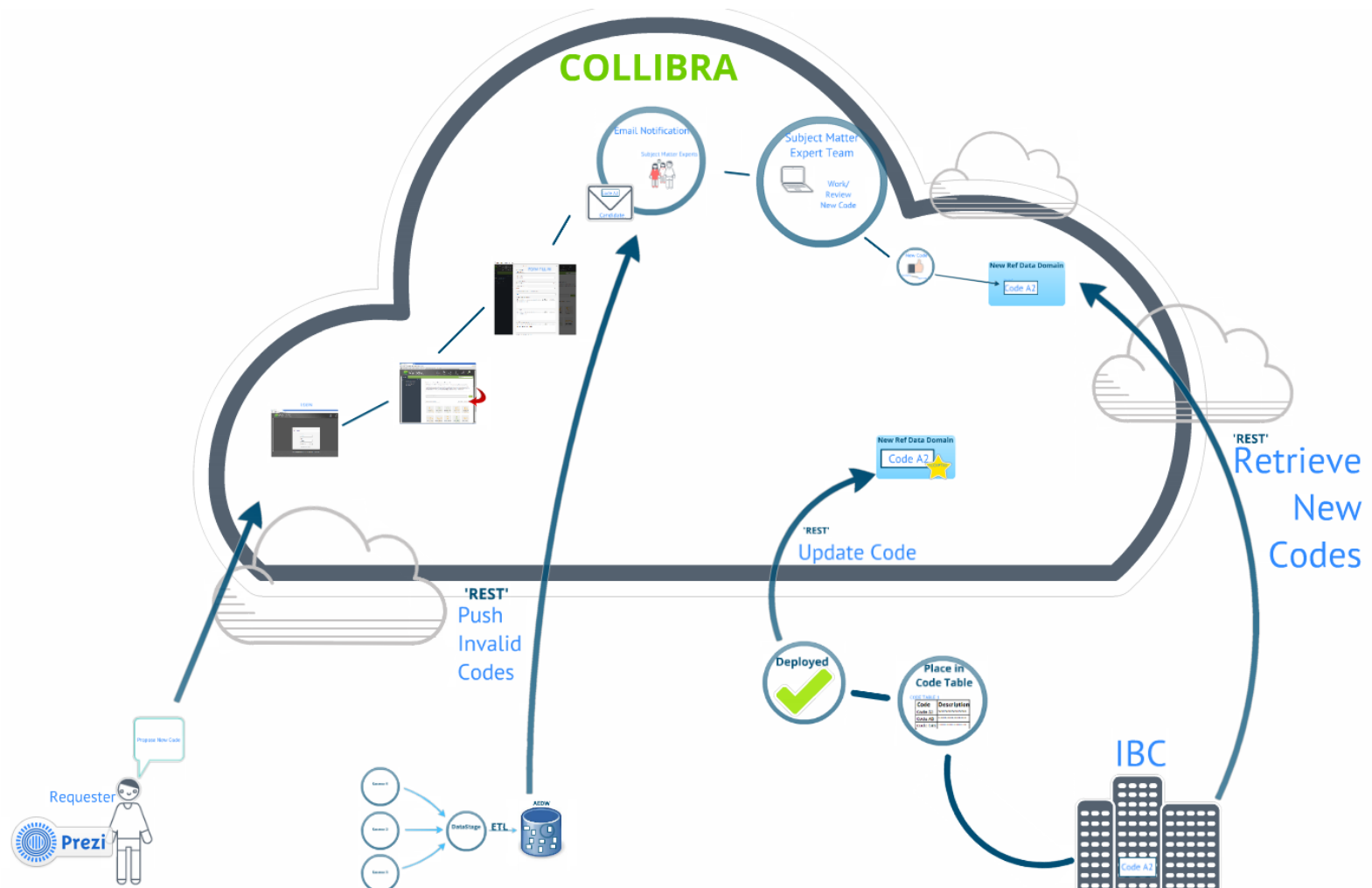


C Assign



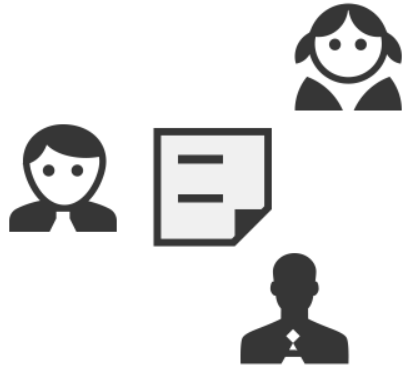
Reference Data & Issue Mgt at Health Insurance Co.

- <http://prezi.com/ve1ws8jmpqcn/workflow/>



Policy Management

1



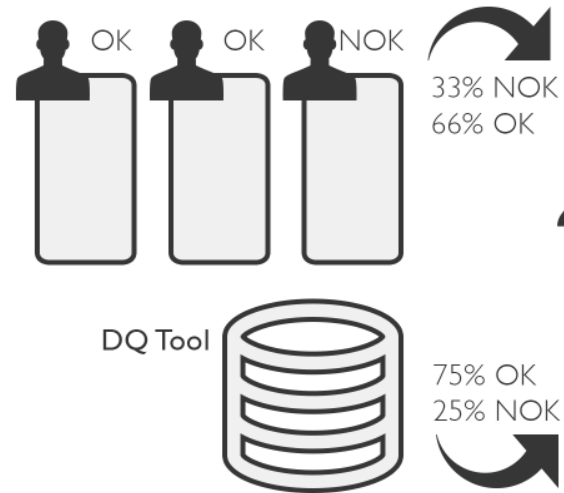
Define

2



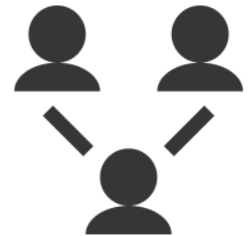
Apply

3



Comply

4

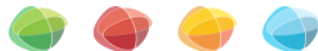


Act



Conclusion

- **Services are data-intensive**
- **Their coproduction requires data sharing across organisation policies / modelling assumptions / regulations**
- **Data Stewardship highlights responsibility aspect of Data Power**
- **Data Governance programs enforces data policy and regulations**
- **Data Governance Technologies are promising to overcome these issues that hamper service innovation**





Collibra

Data Governance Software

Questions & Feedback?

www.collibra.com @collibra contact@collibra.com