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

VU UNIVERSITY AMSTERDAM Faculty of Sciences

Semantics Technology & Applications Research LAB

Vrije Universiteit Brussel

Context and Necessity

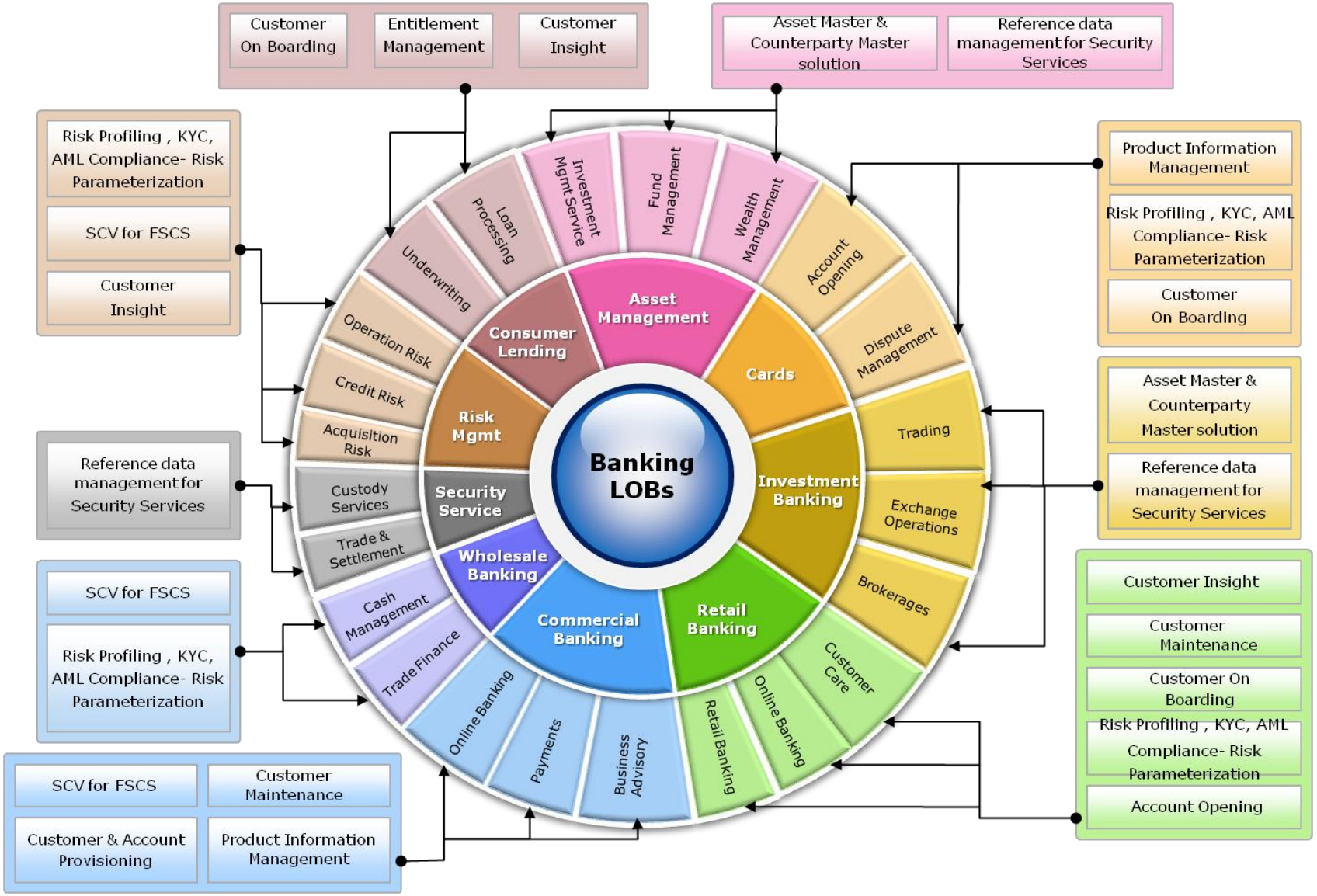
- **Services** are increasingly
  - knowledge-intensive relying on millions of data points from
    - Partners
    - Third parties
    - Customers
  - co-produced in federated, decentralised, multi-tier settings
  - intersection of disciplines
    - Algorithms: *Big Data Analytics*
    - Infrastructure*: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

Data Management Challenges

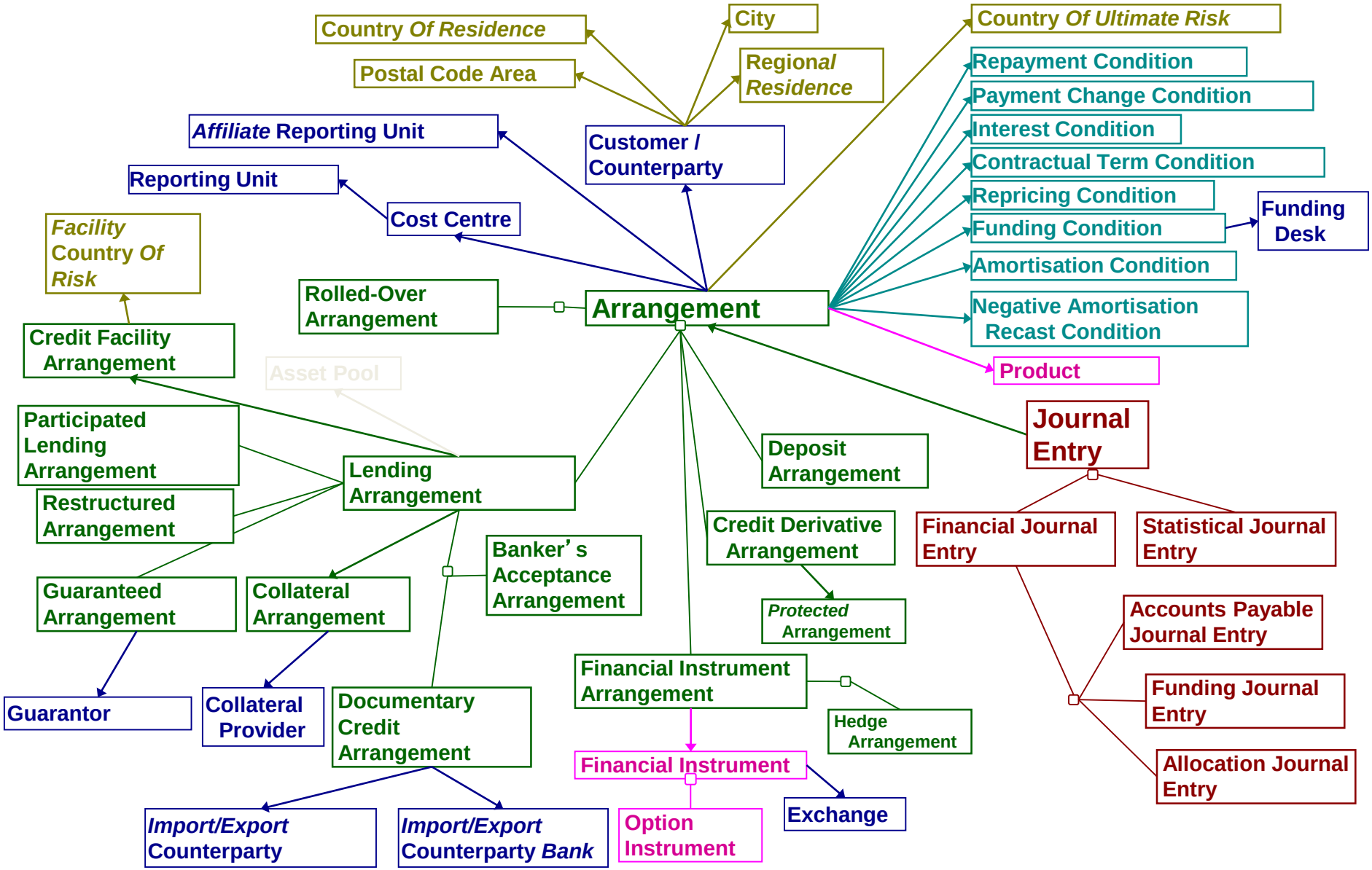
- **Service =** data sharing agreement across organisation silos, policies, regulations, semantic assumptions
- **No clear line between data ownership and control:**
  - responsibilities are ignored
  - for each data point : exposure to risk regarding quality and policy compliance?
    - ➢ ask Alice,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



Data Stewardship & Governance

- **Ownership** **=> Power + Control**

Data Stewardship & Governance

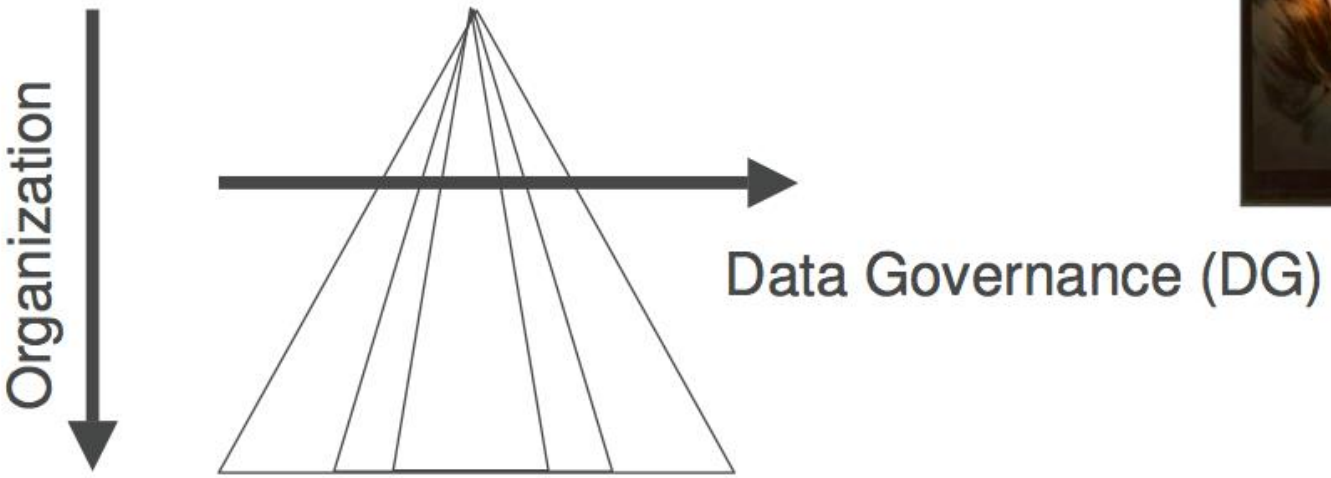
- **Ownership + Responsibility => Power + Control**
  1. (global) data stewardship
     - Req. 1: people who *define* data policy
  2. (systematic) data governance
     - Req. 2: processes that *enforce* data policy
- Now let's build software for it…

"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)

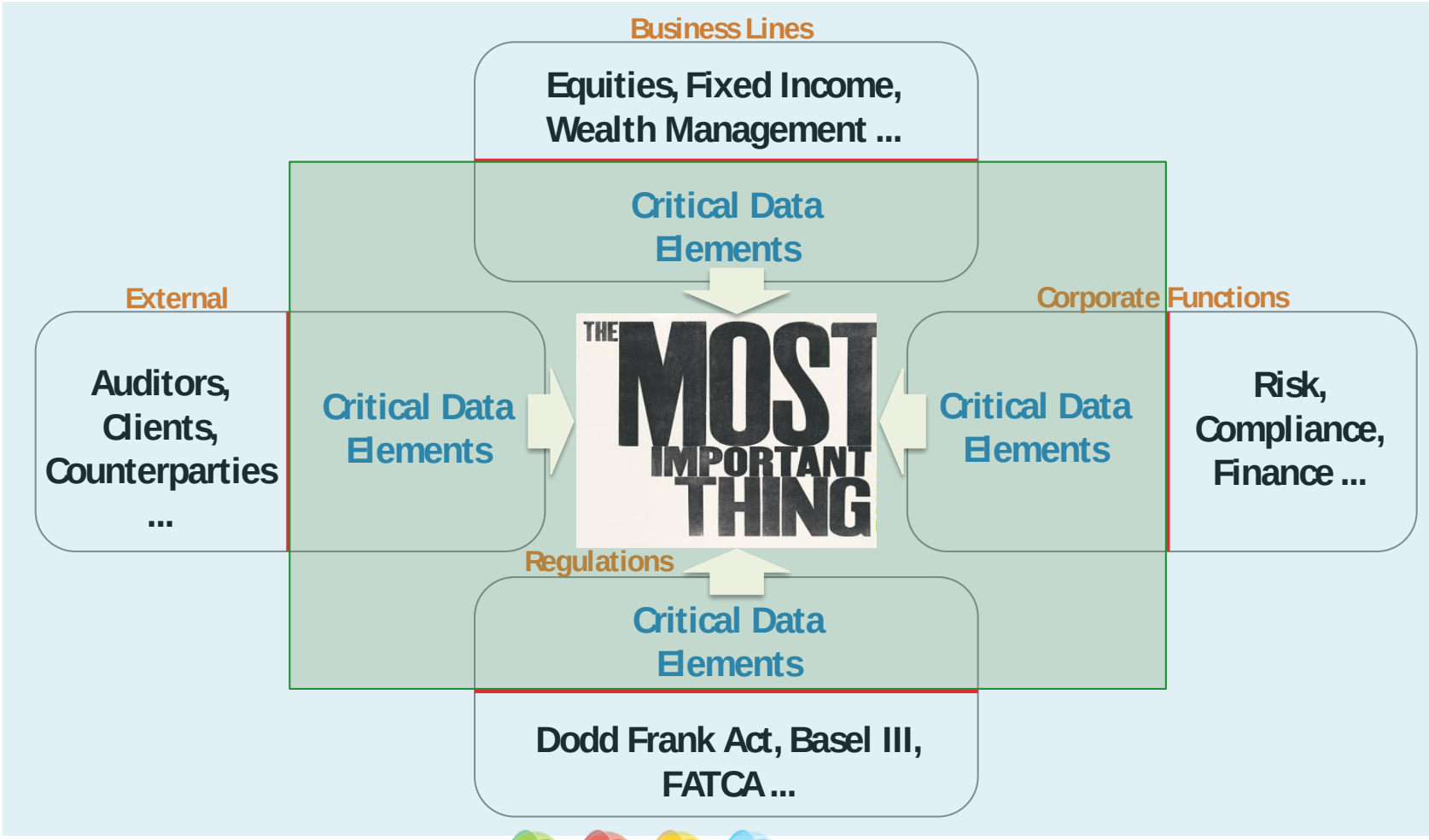


Even if there is an agreement on the added value of DG, the temptation is to continue to work (only) within the own silo

Agreement <> Commitment

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

Borrowed from Predrag Dizdarevic (Element 22 NYC)



Collibra
Data Governance Software

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

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

Collibra

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

Data Governance
Software Company

Reference Customers

STANDARD &POOR'S
ING
novo nordisk
Johnson & Johnson
vmware
VDAB
NetApp
HSBC
colt
smarter / faster / further
PriorityHealth
entsoe
ASPEN
BARRY CALLEBAUT
KYIVSTAR
THINKING OF YOU
EWI
eandis

Products

Data Stewardship Manager

Business Semantics Glossary

Reference Data Accelerator

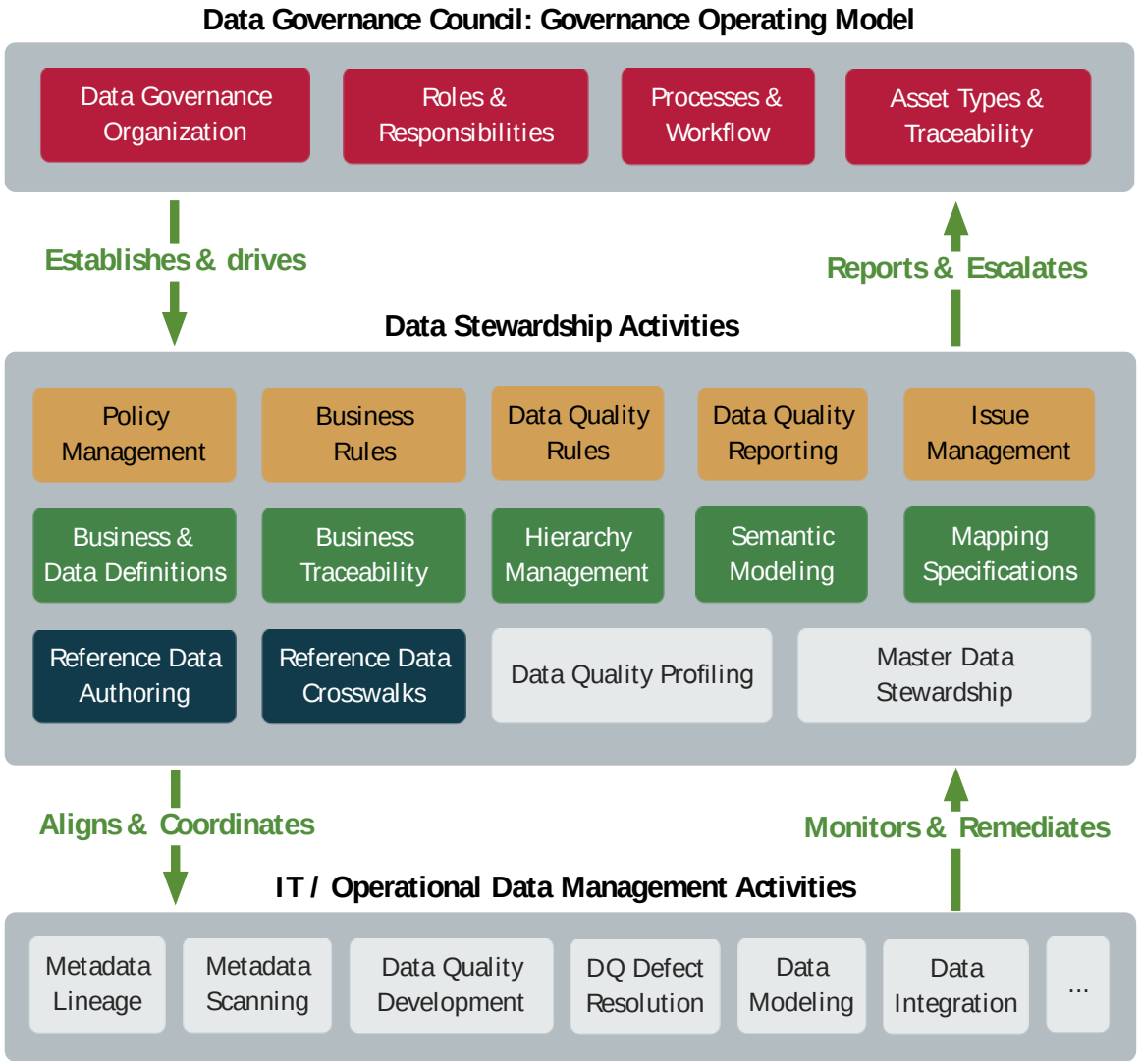
Data Governance Center

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**)

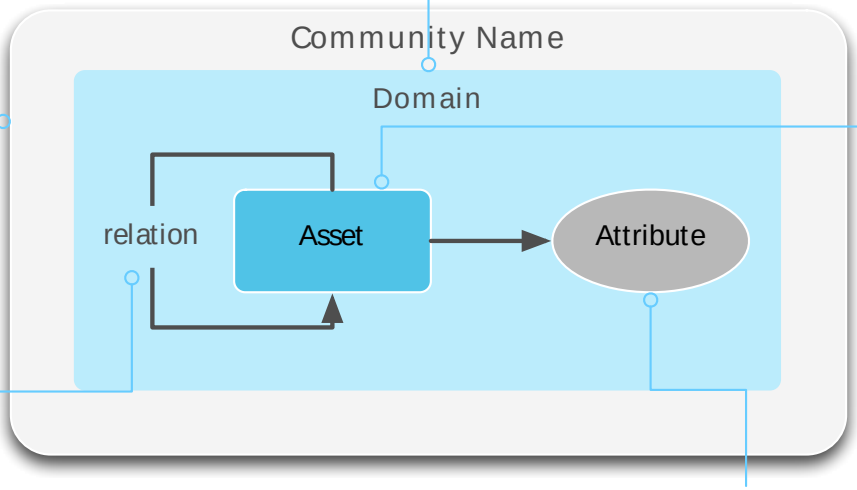
Other Data Management Vendor products

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.
E.g., Customer Domain groups all assets related to customer relationship management
E.g., 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.
E.g., 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*



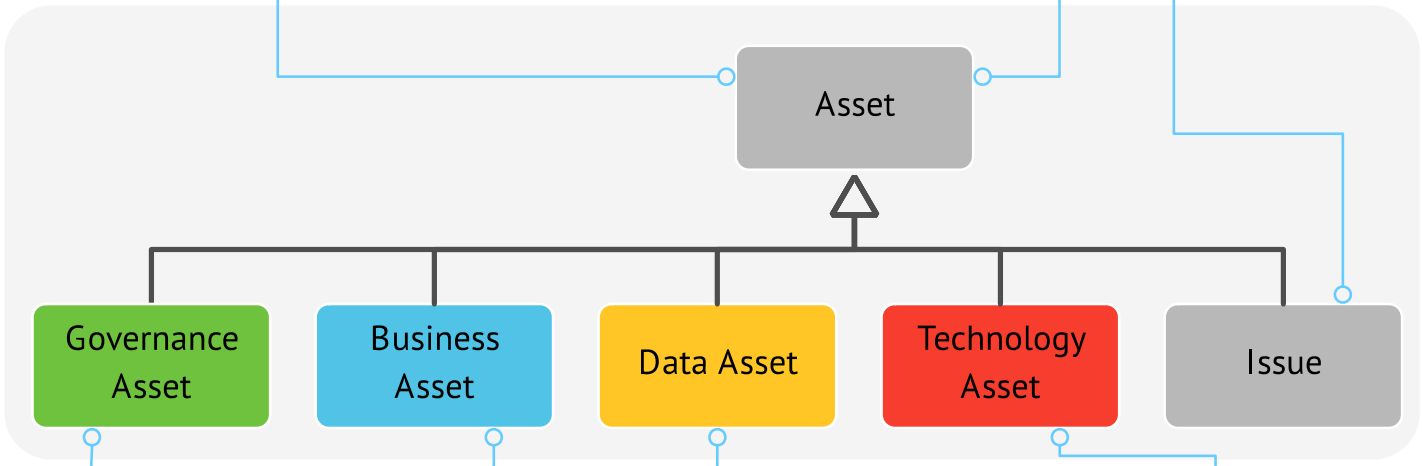
Relations semantically relate 2 assets
E.g., *between assets "Customer" and "CRM": "Customer has system of record / is system of record for CRM"*
E.g., between assets "Customer" and "Gender": "Customer *has gender / gender of* Gender"

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*



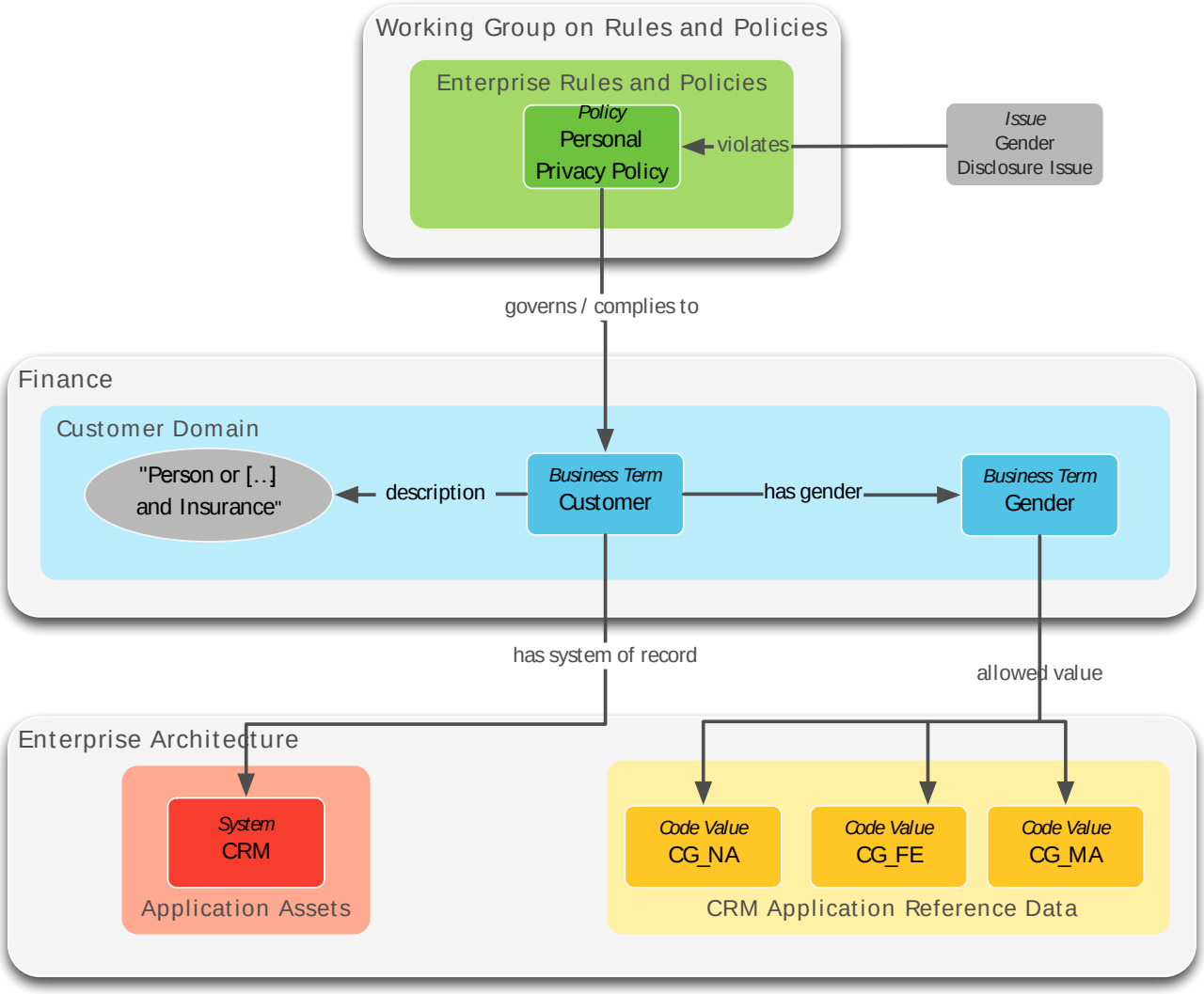
includes asset types such as *Policy* and *Rule*

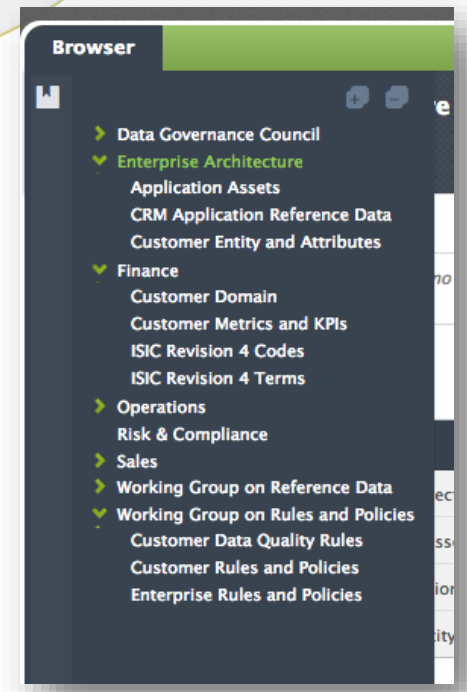
includes asset types such as *System* and *Database*

subsumes asset types such as *Code Value*

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

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

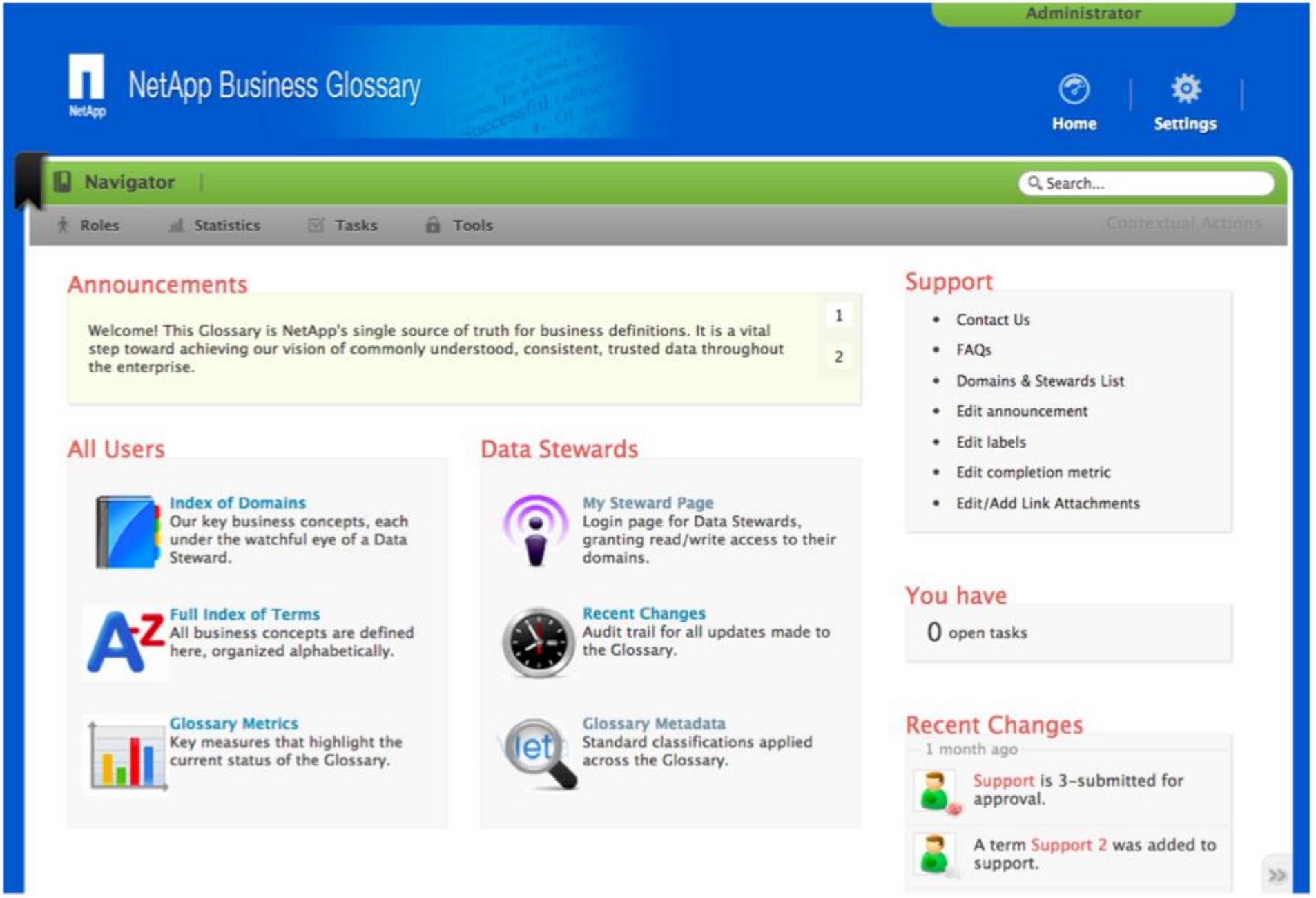
3

Publish

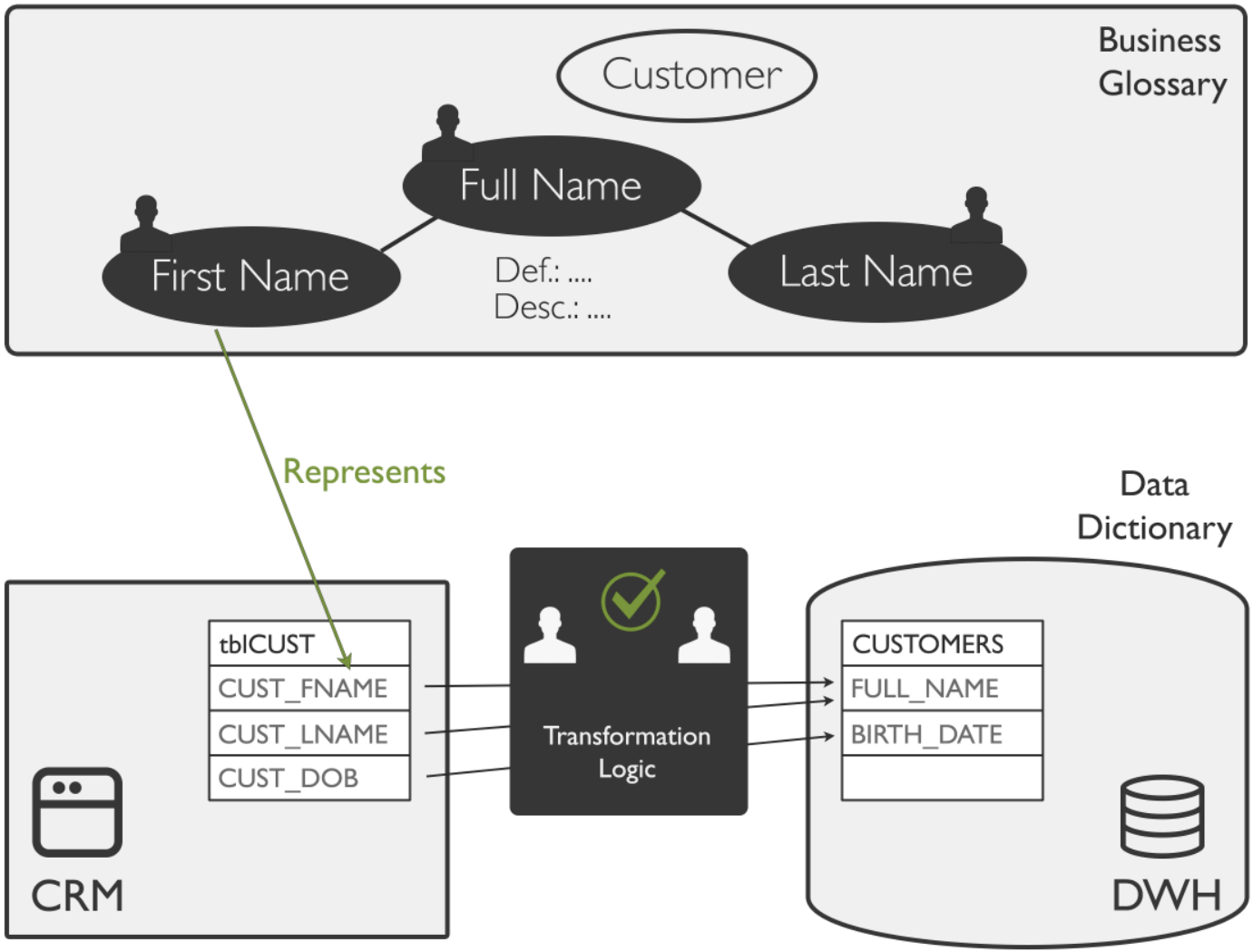
4

Use

DG in Cloud Provider



Data Dictionary

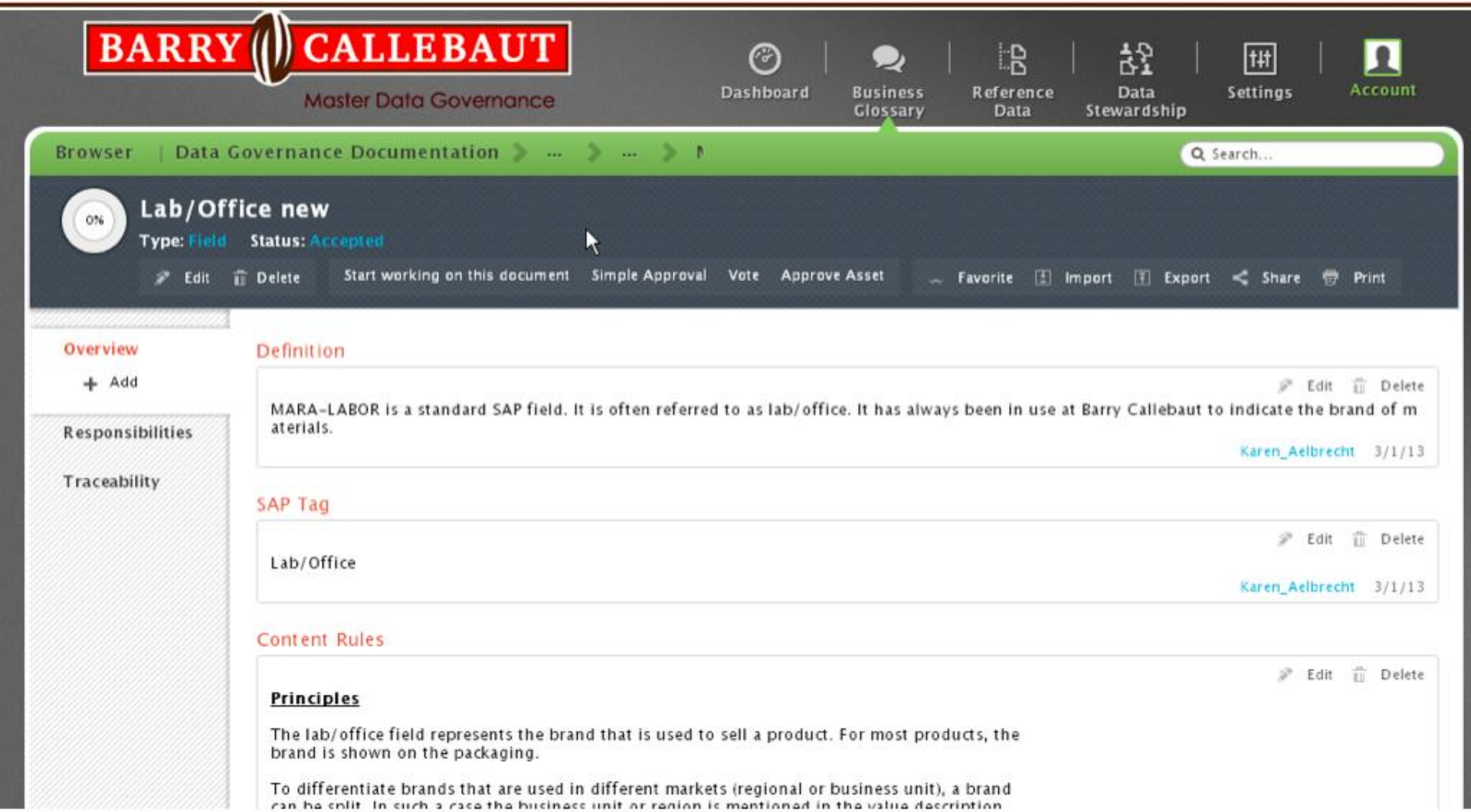


1 Define

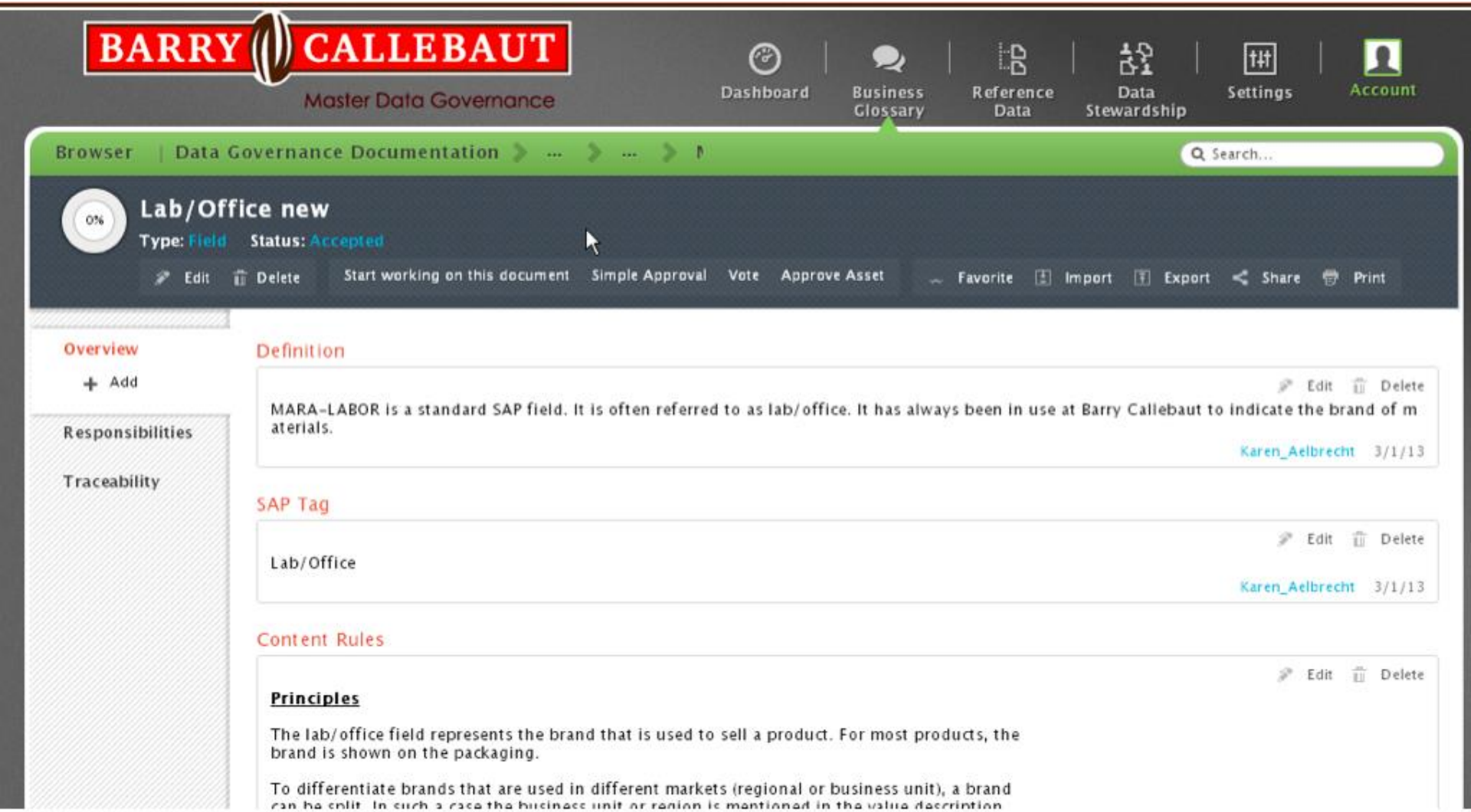
2 Map

3 Visualize

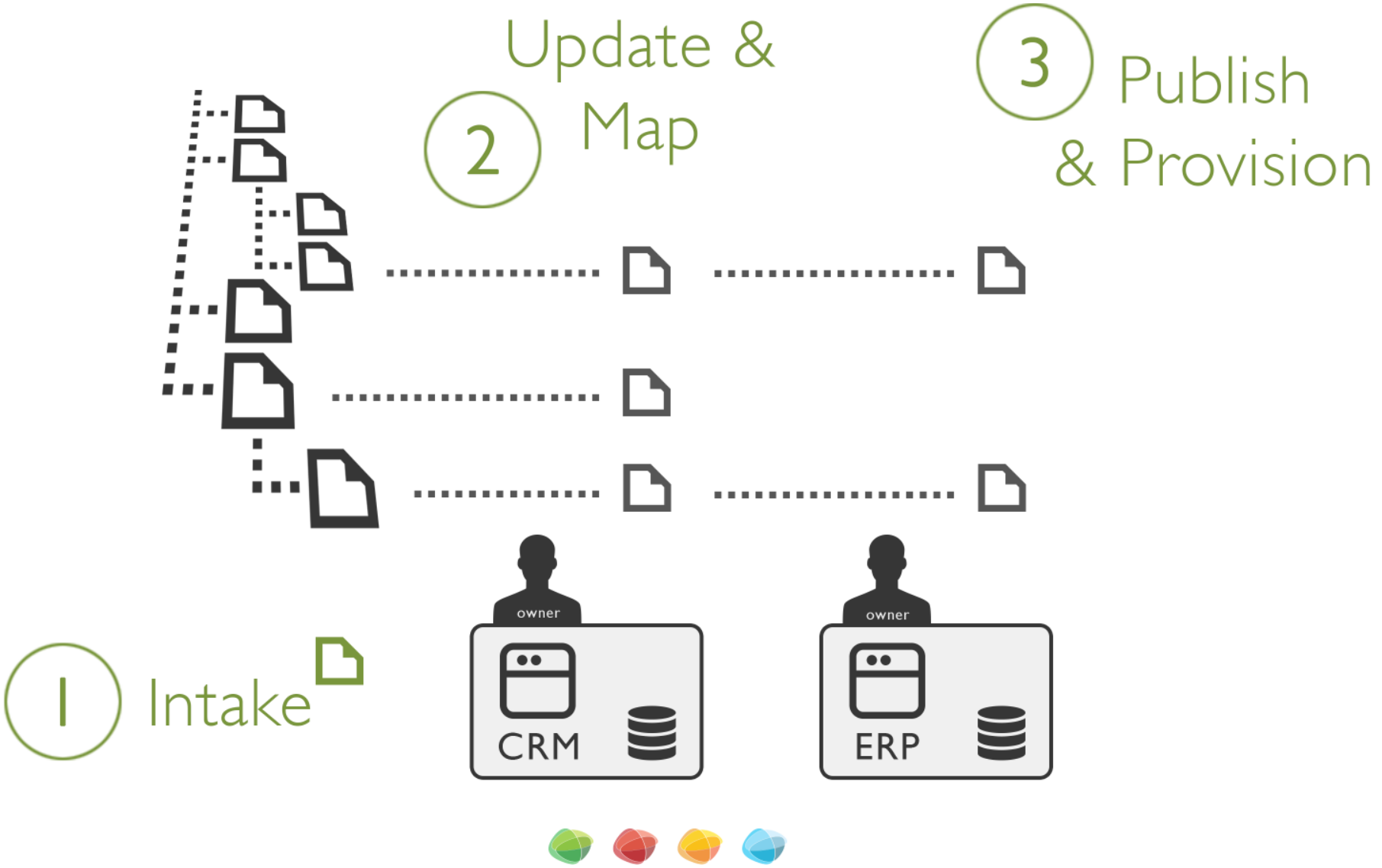
Business Glossary at the #1 Chocolate Factory



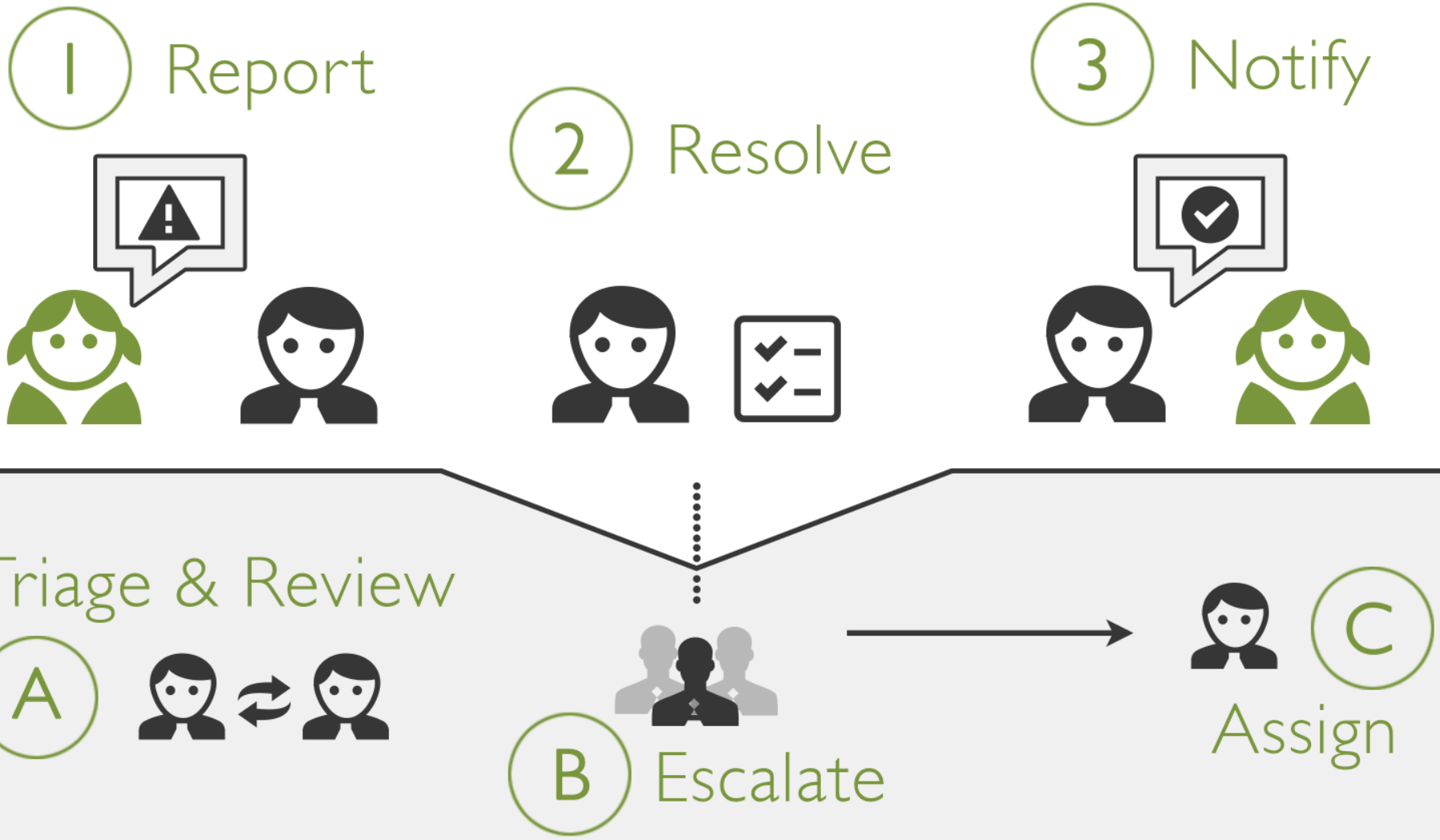
Business Glossary at the #1 Chocolate Factory



Reference Data

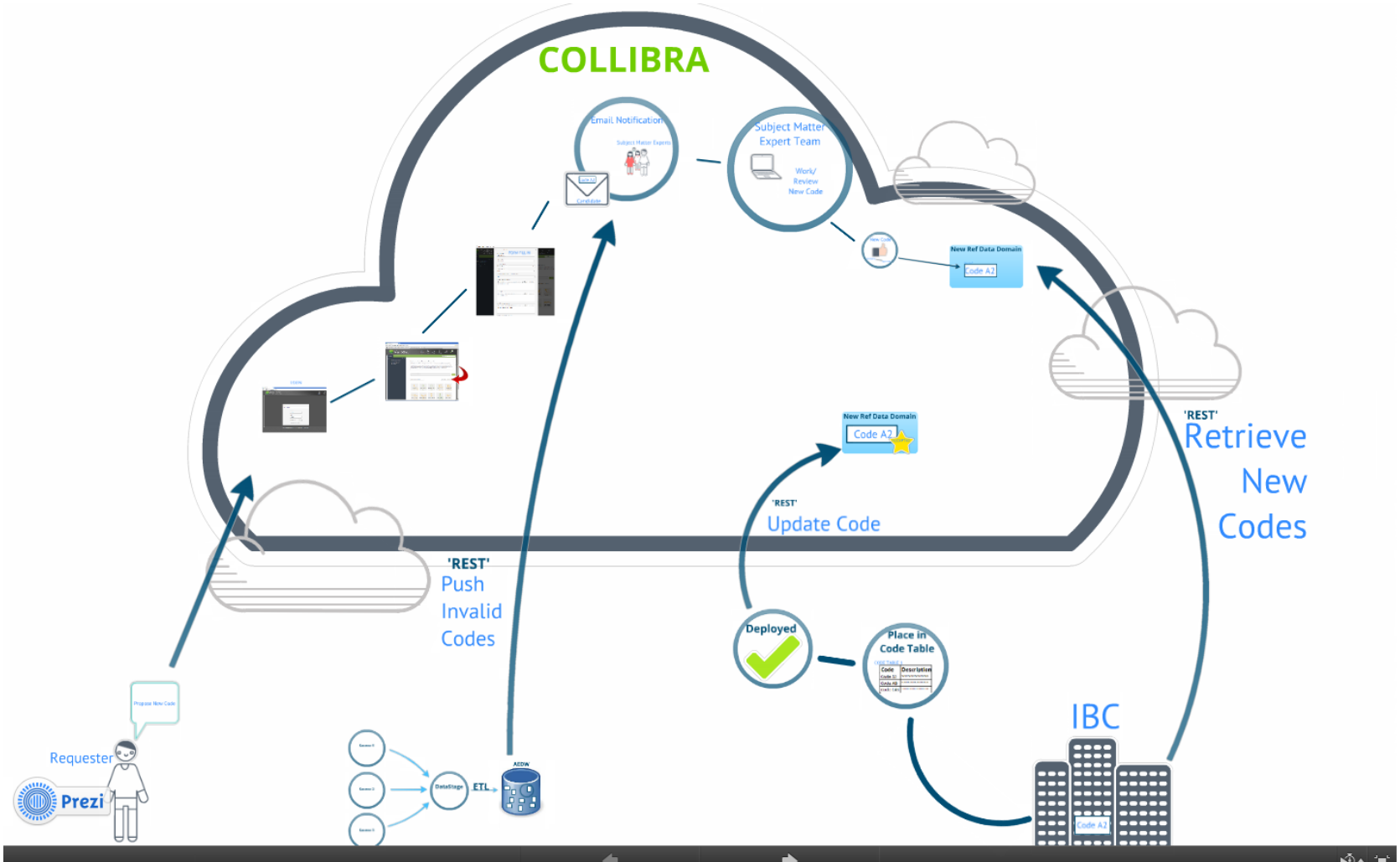


Issue Management



Reference Data & Issue Mgt at Health Insurance Co.

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



Policy Management

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