

CON-6652

**ORACLE
OPEN
WORLD**

October 1–5, 2017
SAN FRANCISCO, CA

Introducing Oracle Data Integration Platform Cloud Governance Edition

**Deliver Trusted Data for Applications and Analytics using the power of
Oracle Public Cloud**

Mike Matthews
Senior Director, Product Management

October 4th 2017

ORACLE

Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Oracle Cloud Platform



Develop & Deploy



Integrate & Extend



Publish & Engage



Analyze & Predict



Secure & Manage

Innovate with a
**Comprehensive, Open,
Integrated and Hybrid**
Cloud Platform
that is
**Highly Scalable, Secure
and Globally Available**

Oracle Cloud Platform

Comprehensive

Open




Integrated



Hybrid

Oracle
Public Cloud



Oracle
Data
Center

-  Data Management
-  Application Development
-  Enterprise Integration
-  Data Integration

-  Analytics and Big Data
-  Content & Experience
-  Identity & Security
-  Systems Management

Oracle Cloud
at Customer



Your
Data
Center

Built on High Performant Oracle Cloud Infrastructure

Oracle Cloud Platform Momentum

14,000+

Oracle
Cloud Platform
Customers



3,000+

Apps in the
Oracle Cloud
Marketplace



\$1.4 Billion

FY17 Oracle Cloud
Platform
Revenue
(60% YoY Growth)



10 PaaS

Categories where
Oracle is a **Leader**
According to
Industry
Analysts



Oracle Cloud Platform for Integration

Application and Data Integration

Complete

Simplified

Open





Oracle Data Integration Platform Cloud Governance Edition (DIPC-GE)

Business Drivers

Data is the **Single Biggest Asset** for Most Companies

“Computing hardware used to be a capital asset, while data wasn’t thought of as an asset in the same way. Now, hardware is becoming a service people buy in real time, and **the lasting asset is the data.**”

– Erik Brynjolfsson, Director, MIT Initiative on the Digital Economy

But **Data is Hard**

- ❑ average cost of an unplanned data outage at \$7,900 a minute, a 41% increase from 2010, when the cost per minute at \$5,600.
- ❑ Typical data outages last 86 minutes, totaling an average of \$690,200 of costs.
- ❑ 72% of big data projects have issues with data integration reliability
- ❑ 17 month average implementation time for a BI project; 5 months before there is any usable BI artifacts
- ❑ <32% of BI projects are declared successful
- ❑ 36% confidence rate that the right data is available to the right people at the right time
- ❑ 42% firms managing over \$200B in assets do not have regular processes and procedures to control data usage

NEW: Oracle Data Integration Platform Cloud

Key Capabilities:

1. Data High Availability
2. Data Migrations
3. Data Warehouse Automation
4. Databus & Stream Integration
5. Data Governance

DIPC Solution Use Cases





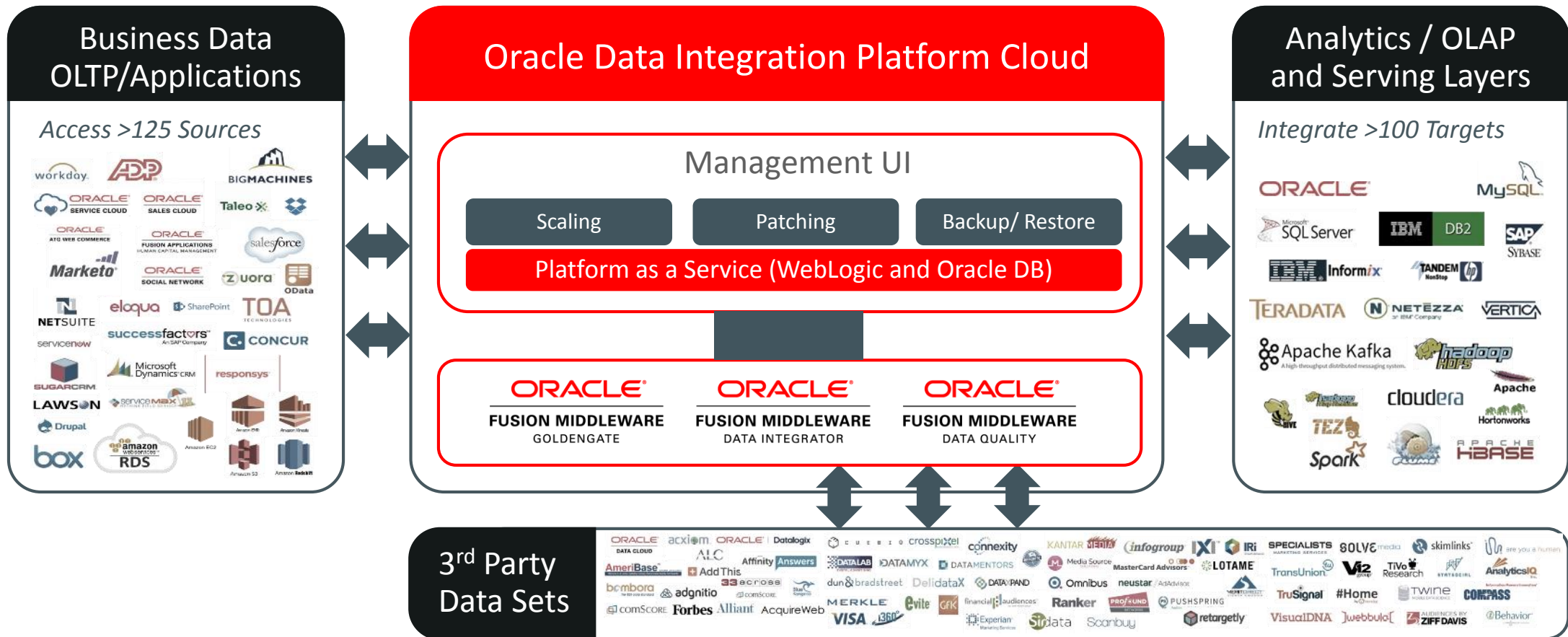


Oracle Data Integration Platform Cloud Governance Edition (DIPC-GE)

Introduction and Key Features

DIPC-GE | What is it?

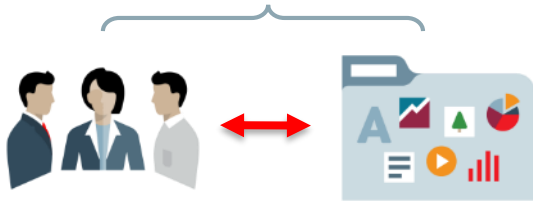
Oracle Data Integration Platform Cloud Governance Edition is a fully provisioned environment running on Oracle Platform as a Service catering to all your enterprise data replication, integration, quality and governance requirements.



DIPC-GE | What can I do with it?

Understand Data

Profile and analyze data to find quality issues



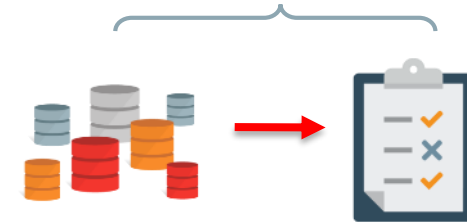
Standardize and Enhance Data

Fix the data problems you find



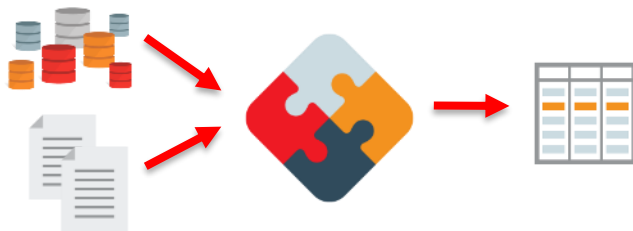
Implement Business Rules

Apply governed business rules across many sources



Match and Merge

Deduplicate, Enhance, Link and Consolidate Enterprise Data



Track Data Quality

Measure data quality and track it over time to prove trust



Deploy DQ Services

Embed Data Quality processes in real-time data interchange



DIPC-GE | Usability

For Data Governance use cases, DIPC provides a fully provisioned instance of **Oracle Enterprise Data Quality**, the market-leading Data Quality Management tool for non-technical users.

- **Accelerate** your data insights with fast data ingestion, manipulation and profiling
- **Engage** Business Users directly with the data they know
- **Collaborate** to understand and improve your data
- **Design** reusable Data Services that integrate easily in any application or data flow

The screenshot displays the Oracle Enterprise Data Quality Director interface. The main window shows a data flow process for '2 - Prepare Master Data'. The process starts with 'US Customers' data, which is processed through several steps: 'Normalize Whitespace (all attributes)', 'Parse Name', 'Add Gender', and 'Convert DoB to formal date'. The output of 'Convert DoB to formal date' is split into 'Successful' and 'Unsuccessful' paths, with the 'Unsuccessful' path leading to 'Country Standardization'. Below this, a 'Make Identifiers for Matching' block contains 'Make WholeAddress' and 'Make GivenName' steps. A yellow callout box indicates 'Match Rules approved by CDO, 5th Oct 2015 - see Project Note re. Approva...'. The 'Make Identifiers for Matching' block outputs 'Working Data', 'Reference Data', and 'Decisions', which are further processed into 'Groups', 'Relationships', and 'Merged' data. A 'Results Browser' window is open, showing a table of match results:

Review status	Relationships
Automatic match	18
Manual match	5
Pending	0
Possible Match	2
Manual No match	1

A 'Graph' window is also open, showing a pie chart for '2 - Prepare Master Data' with the following data:

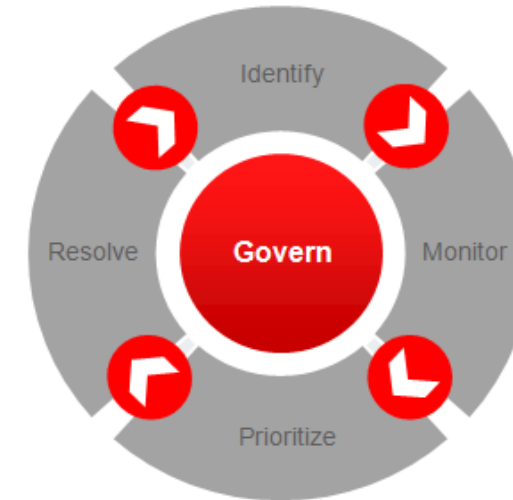
Match Type	Percentage
Automatic match	69%
Manual match	19%
Pending	0%
Possible Match	8%
Manual No match	4%

The 'Results Browser' and 'Graph' windows are part of the '2 - Prepare Master Data' job. The 'Data Stream' is 'US Customers'. The 'Process' is '2 - Prepare Master Data'. The 'Results Browser' and 'Graph' windows are both titled '2 - Prepare Master Data'. The 'Graph' window shows a pie chart with a legend for 'Automatic match', 'Manual match', 'Pending', 'Possible Match', and 'Manual No match'. The 'Results Browser' window shows a table with columns for 'Review status' and 'Relationships'. The 'Graph' window shows a pie chart with a legend for 'Automatic match', 'Manual match', 'Pending', 'Possible Match', and 'Manual No match'. The 'Results Browser' window shows a table with columns for 'Review status' and 'Relationships'. The 'Graph' window shows a pie chart with a legend for 'Automatic match', 'Manual match', 'Pending', 'Possible Match', and 'Manual No match'.

DIPC-GE | Governance Features

Use EDQ's extensive Collaboration and Governance features to engage Stakeholders, and to give flexible rule management, review and remediation options to Data Stewards when they need to change rules or make manual decisions

- Simple management of **Business Rules** for data quality without the need to modify DQ services
- Zero training user-tailorable **Data Quality Dashboard** for KPI monitoring and trend analysis
- Highly flexible **Case Management** application for data stewards to review data quality issues and make decisions



Data Analysts



Director



Data Stakeholders



Dashboard



Data Stewards

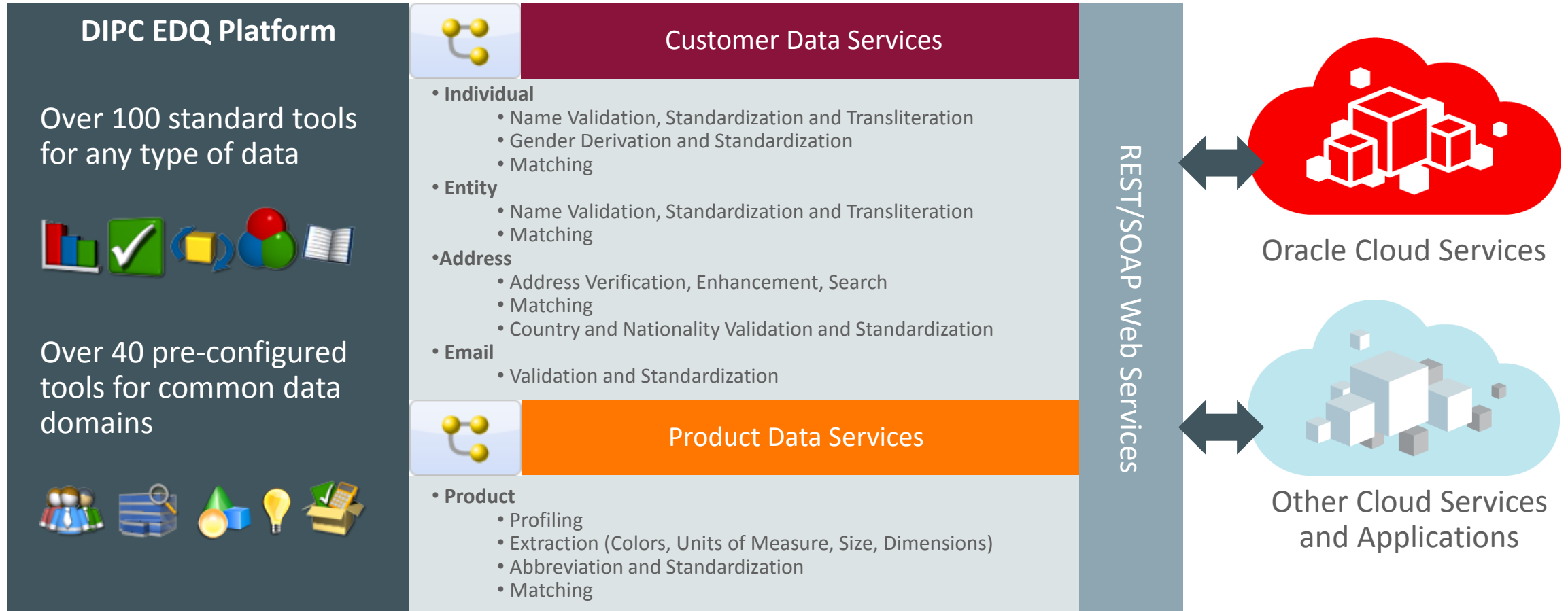


Case Management



DIPC-GE | Data Diversity and Specialism

EDQ is data agnostic and offers unparalleled time-to-value with **any** type of data in **any** language. It also provides out-of-the-box, pre-integrated, tailorable, services and tools for common data domains such as Customer and Product Data.



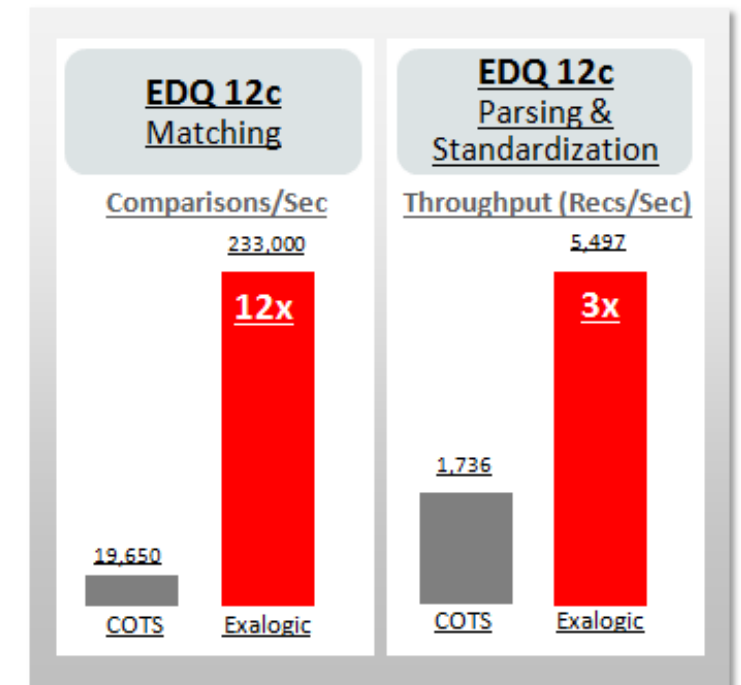
DIPC-GE | Performance and Scalability

DIPC and EDQ fully support Oracle Cloud's elastic scalability features, enabling seamless scaling up, down, in and out, controlled from simple options on the PaaS UI.

EDQ is also fully compatible with Oracle Exalogic, Exadata and Oracle Cloud Machines, providing optimized ultra-fast performance for both on-premise and cloud customers where needed.

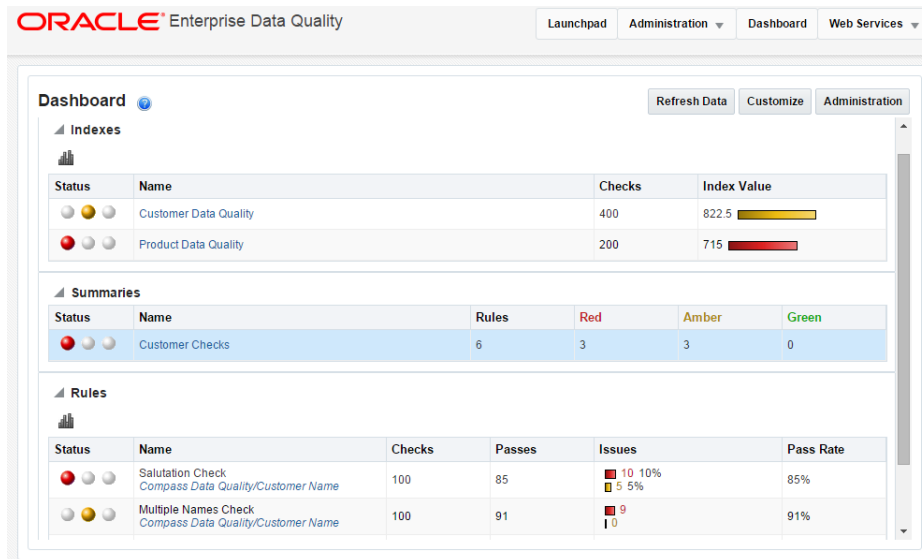
- Recent EDQ benchmarks:
 - **1 billion** records parsed, standardized and matched against **2 million** in an **8 hour** window (6 node cluster)
 - **1000s** of concurrent data stewards working alerts
 - **660** transactions per second per OCPU when matching 60 records per message in real-time with 100% sub-second response time
- Efficiency improvements in recent releases to ensure good operation when scaled down as well as up
- Fully co-ordinated multi-machine processing in a cluster
- Support for dynamic clustering, automatic spin-up of real-time services on new nodes, immediate availability for batch processing

EDQ Benchmark on Engineered Systems (100m records, 1 node)

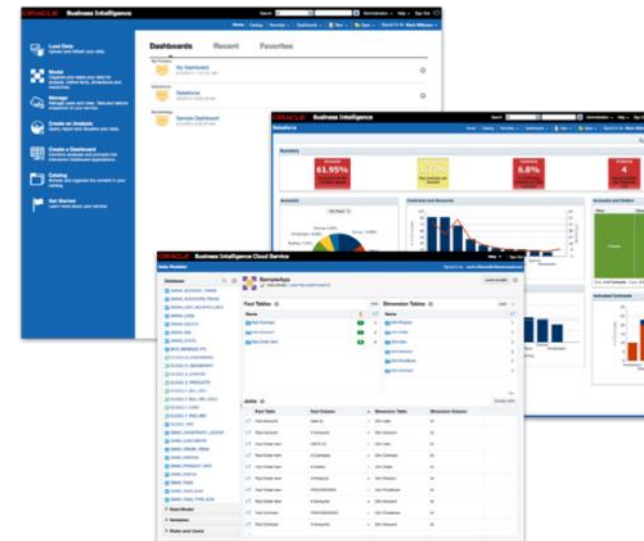


DIPC-GE | Reporting and Analytics

Built-in reporting provides stakeholder view of Data Quality metrics at any aggregation level from both batch and real-time processing, including trend analysis. Integrated with Oracle BI Cloud Service for deep analytics and visualizations where required. EDQ also provides immediate operational reporting in Case Management to check on Data Stewardship progress.



Built-in KPIs for Governance Stakeholders



Advanced Analytics using DQ metadata with BI Cloud Service

DIPC-GE | **Advanced Capabilities**

EDQ provides proven flexibility and functional breadth for richer data processing requirements such as multi-language processing, item classification, unstructured data parsing, and advanced data matching. Decision interfaces facilitate the use of machine learning and automation for data stewardship.

- **Effective cross-script matching and name variant recognition powered by over 4 million rows of reference data**
- **Match Engine can combine Deterministic, Probabilistic, Semantic and Learning approaches**
- **Rich unstructured text parsing supports item classification, name and address parsing, context-specific standardization etc.**
- **EDQ processor creation allows encapsulation and reuse of any configured logic**
- **All manual decisions can be made externally – and automated if required**

DIPC-GE | Migration from On Premise

For most EDQ customers, migration of configuration to DIPC should be straight-forward.

1. Package all projects from on-premise EDQ (to DXI file)
 2. Provision cloud DIPC instance
 3. Use SFTP interface to load in any extensions and landing area files
 4. Import the projects from DXI file
- **Note 1:** The [Oracle Network Cloud Service](#) may be needed for ongoing connectivity to on-premise data stores.
 - **Note 2:** Customers using Case Management or Dashboard will need to migrate their EDQCONFIG schemas into the provisioned DBaaS instance in order to preserve Case and Dashboard data
 - **Note 3:** Customers using LDAP for user access will need to integrate the Oracle Identity Cloud Service to their LDAP store using a bridge

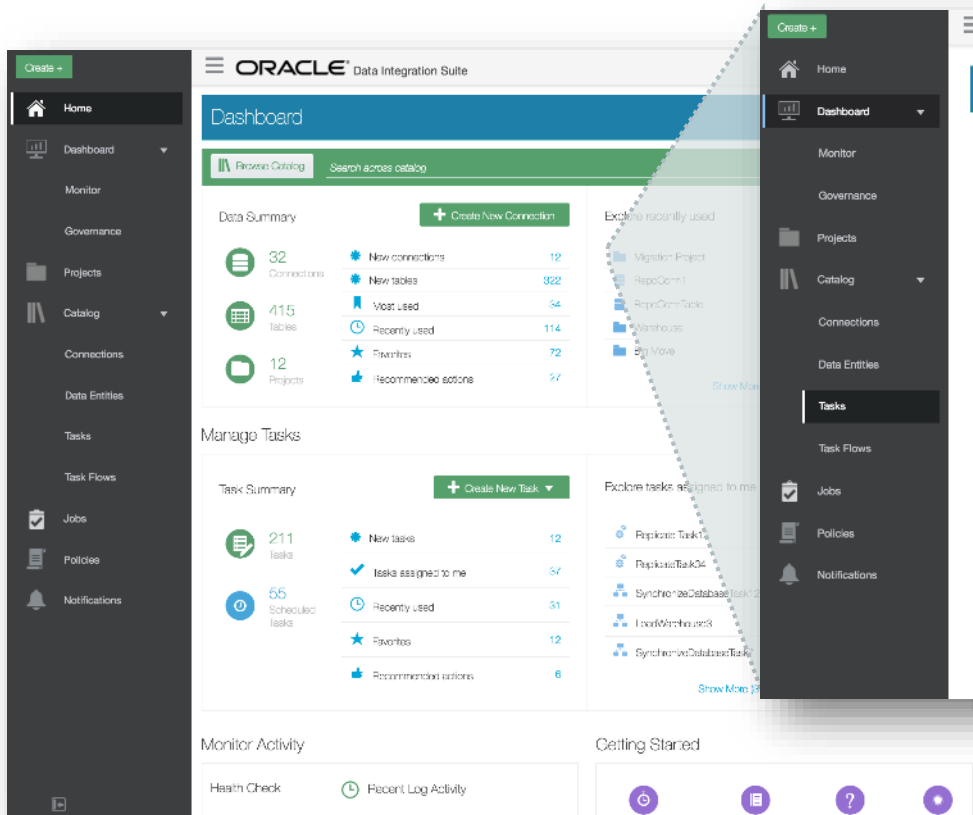
A man with a beard and mustache, wearing a dark suit, light blue shirt, and dark tie, is looking down at a smartphone he is holding in his hands. The background is a blurred city night scene with bokeh lights. The entire image is overlaid with a semi-transparent teal color.

Oracle Data Integration Platform Cloud Governance Edition (DIPC-GE)

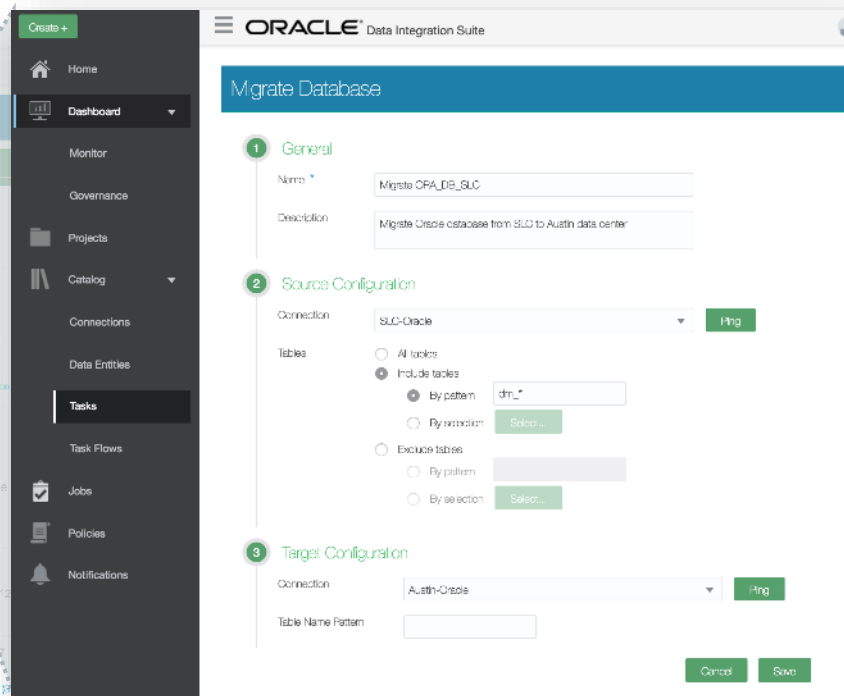
Product Vision and Roadmap

Data Integration Platform Cloud

...a **Unified** solution



...that's **Easy** to use



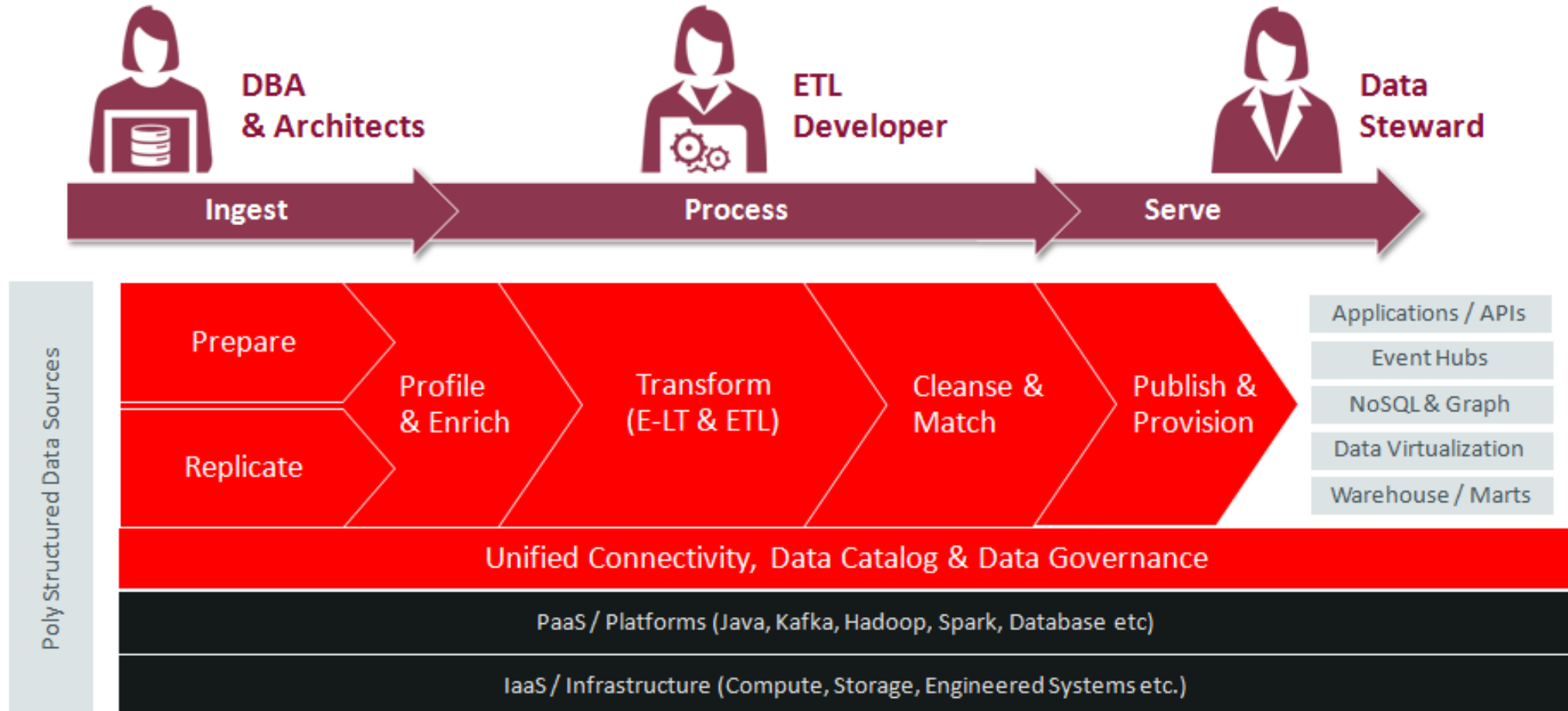
...for **Powerful** data-driven solutions

Key Capabilities

1. Data High Availability
2. Data Migrations
3. Data Warehouse Automation
4. Databus & Stream Integration
5. Data Governance

DIPC-GE | Product Vision

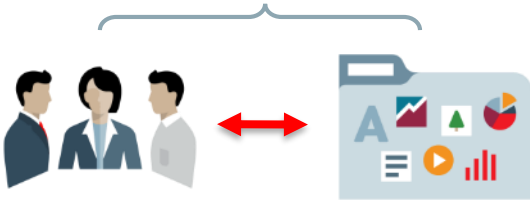
Over time, Oracle Data Integration Platform Cloud will grow to provide a new integrated user experience and new functionality for all data integration, replication, quality and governance purposes. Upgrades will be seamless.



DIPC-GE | What will I be able to do?

Capture Data Knowledge

Enhance the value of data assets with crowd-sourced knowledge



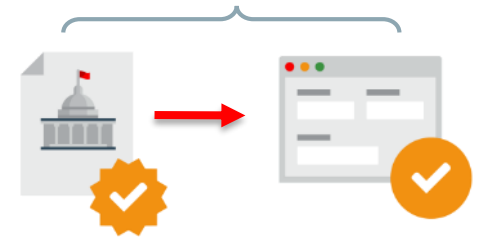
Define Business Terms

Define business vocabulary once, use everywhere



Understand Data Quicker

Quicker, machine-assisted profiling and rule linkage



Manage Policies

Manage data standards with approved policies



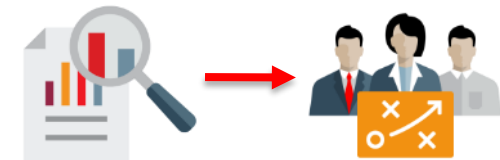
Track Data Lineage

Show where data is created and how it flows to targets

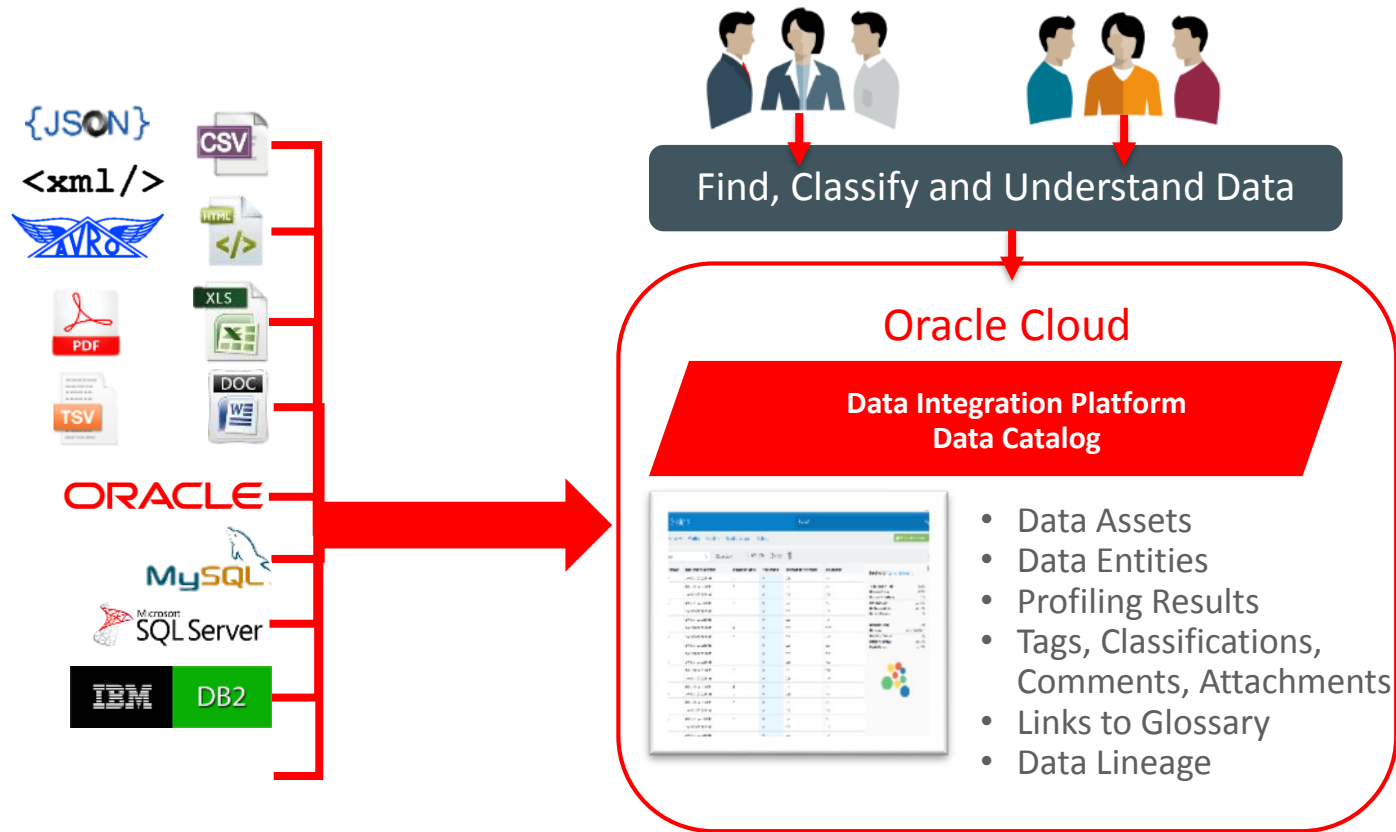


Review and Improve

Keep on top of frequent data changes



Data Catalog



- Easily searchable inventory of enterprise data assets
- Instant data profiling and previews to understand data
- Capture knowledge about data for all users
- Define data asset ownership and access control for Data Governance
- Intelligent and automated linkage to Business Glossary (Phase 2)
- Track usage and lineage of data assets (Phase 2)

Business Glossary and Rules



Data Stewards define Business Glossary Terms and Business Rules, approved by Governance Council

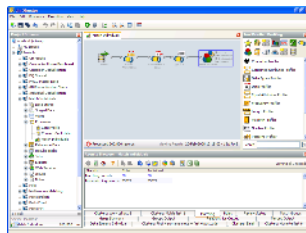
Identifier	Name	Metrics	Status	Domains
R1011	Account Name Populated	3	Approved	Customer
R3005	Address 1 Populated	2	Approved	Customer,Site
R3004	Address can be Verified by AV	4	Pending	Customer,Site
R3001	Address Populated	2	Pending	Customer,Site
R1015	At Least 1 Phone Supplied	3	Approved	Customer
R3002	City Populated	5	Pending	Customer,Site
R1013	Date of Birth in the Past	3	Approved	Customer
R1012	Date of Birth Populated	3	Approved	Customer
R1014	Email Populated	3	Approved	Customer
R1008	Email Valid Format	3	Approved	Customer
R1003	First Name Length Check	2	Approved	Customer

Attributes	Thresholds	Metrics	Workflow	History	Links
Identifier		R1008			
Name		Email Valid Format			
Domains		Customer			
Priority		1 - High			
Description		Email address has a valid format			
Created		14-03-2016 04:41			
Created by		Peter Spencer			
Last updated		28-03-2016 05:07			
Last updated by		Jack Davis			

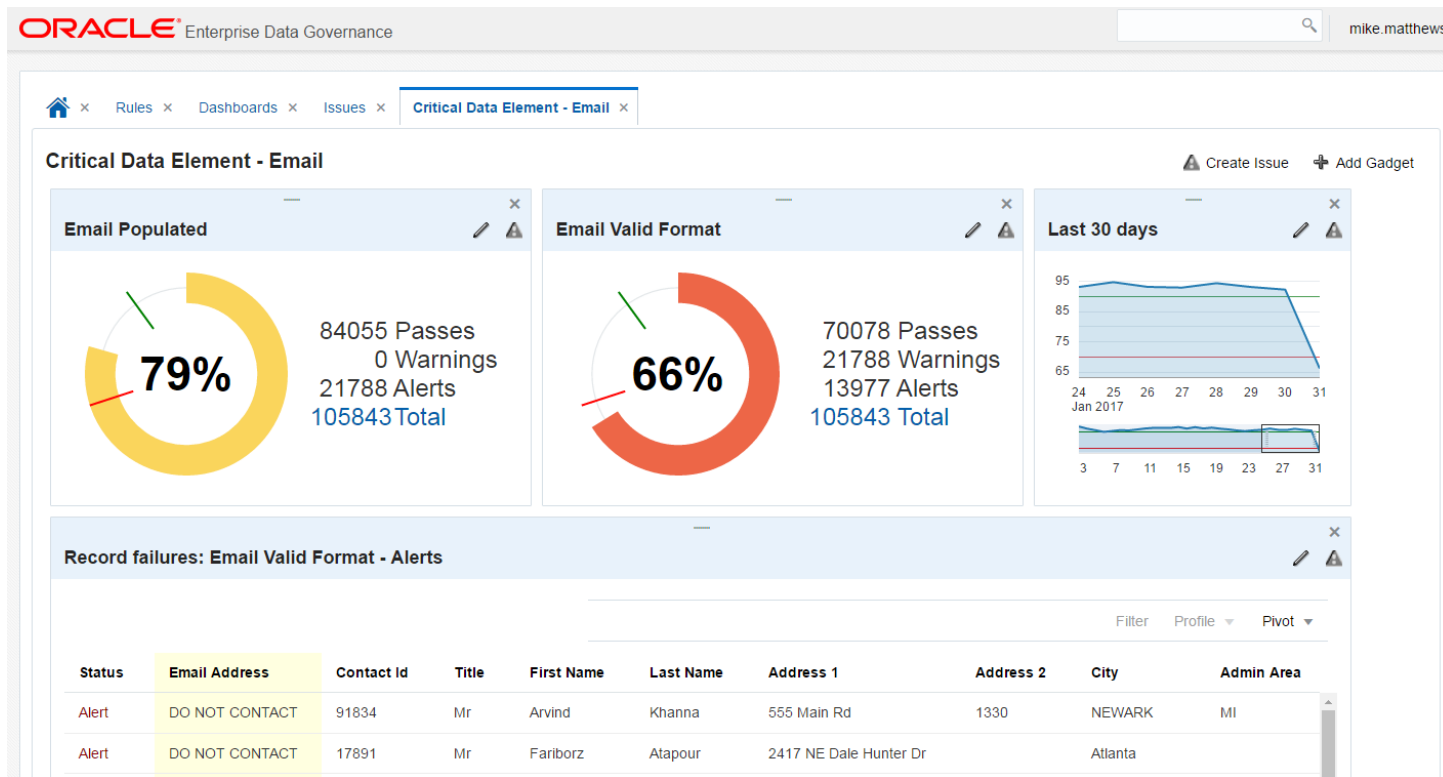
- Define common business terms and definitions
- Attach business rules to measure the quality of data at logical level
- Link physical data assets (data entities) to the glossary
- Implement business rules using DQ engine (EDQ, eventually natively in Suite) and publish results
- Publish DQ metrics for data stewards
- API to access Glossary externally



Data Analysts implement business rules (using EDQ for now) and publish metrics/results for stewardship

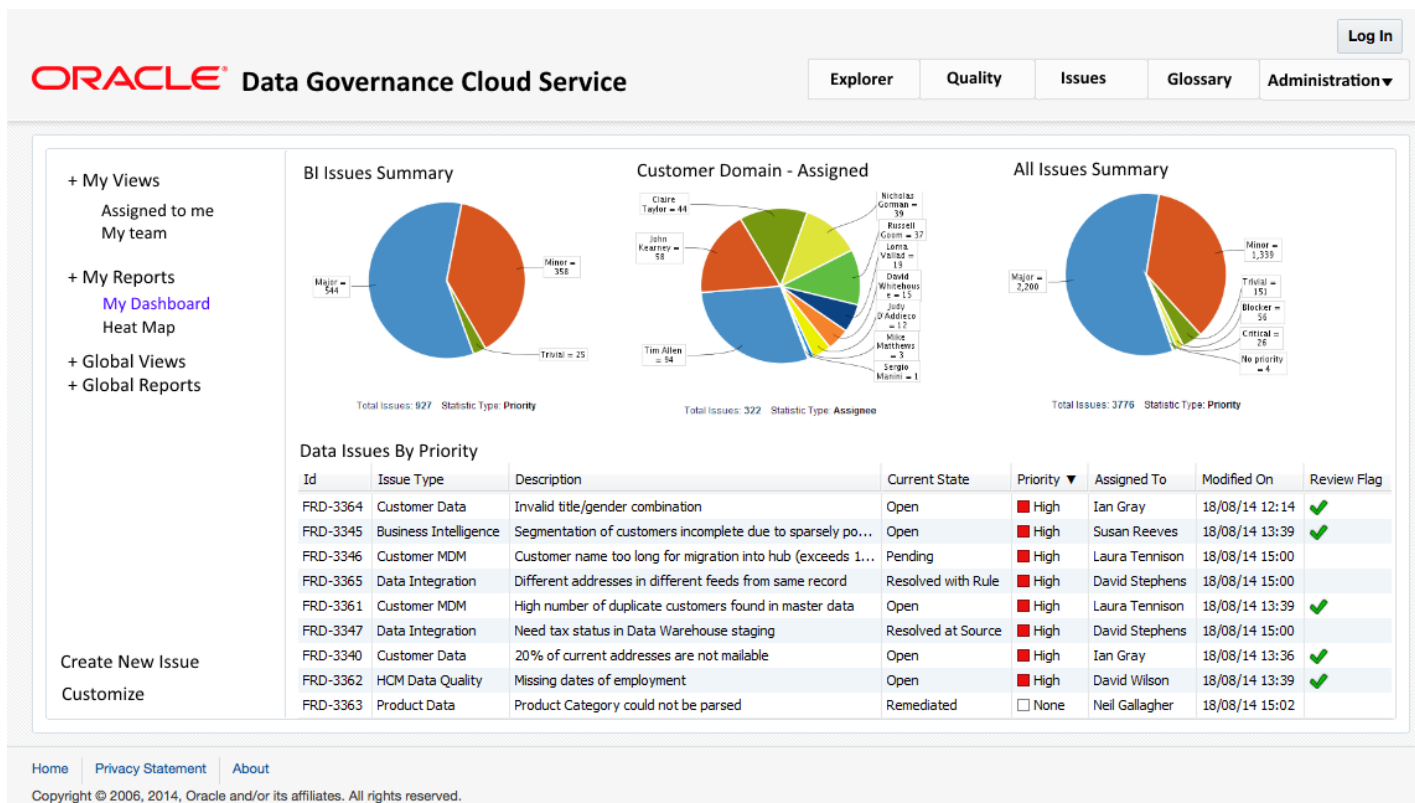


Data Quality Assurance



- Use EDQ, the market-leading DQ Platform for Business Users, to measure and improve data quality
- EDQ to implement business rules, and publish metrics and results to the Suite UI
- Provides instant visibility of data quality, and data issues to data stewards
- Implement services to standardize, enhance and match data
- Deliver trusted, high quality data to data consumers

Data Stewardship



- Business Glossary and Business Rules to define enterprise data standards
- Highly configurable Data Quality Dashboards with drilldowns to data
- Automated notifications on KPI threshold breaches
- Issue Management to collaborate and resolve data issues
- Data Lineage diagrams to understand and trace data flows



Customer Spotlight · Cummins

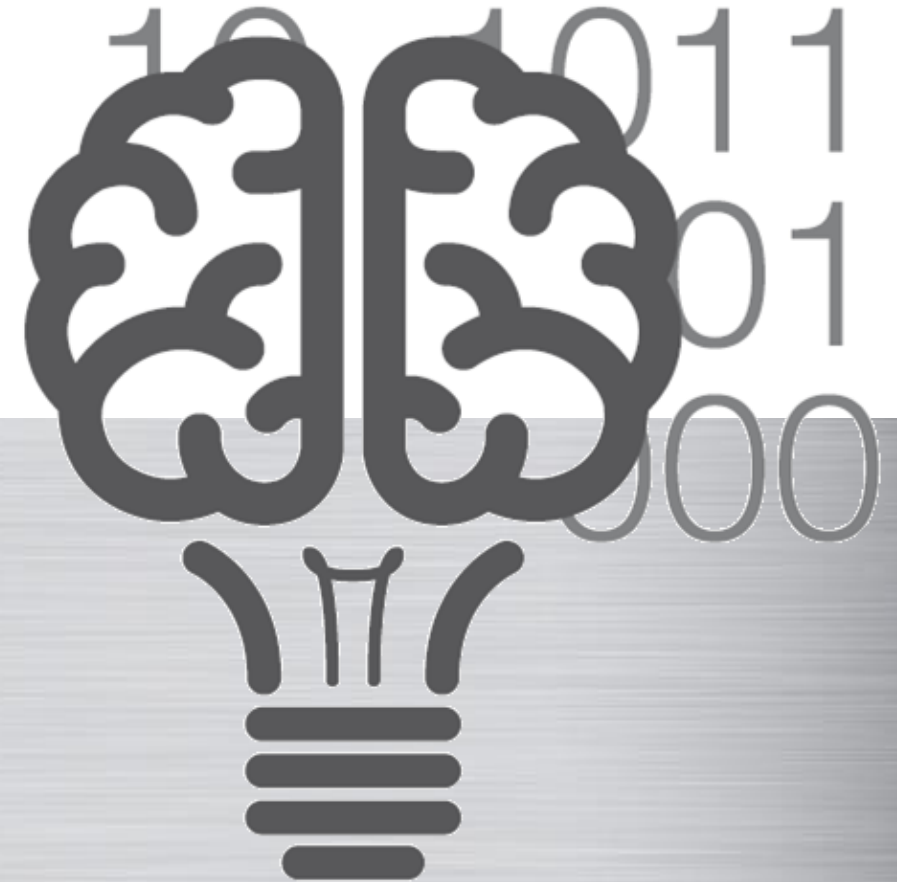
Neha Kaptan, Information Quality and Governance Leader

Data Governance at Cummins

Neha Kaptan

October 2017

Internal Use Only



CONNECTING CUMMINS
an Envision Cummins capability, Partnering to Transform

Data Governance at Cummins



Data
Quality
Improvement



Data
Quality
Control
Services



Business
Alignment
for
Analytics

Data Quality Improvement

Supplier Data Quality Dashboard Summary - Jun'17

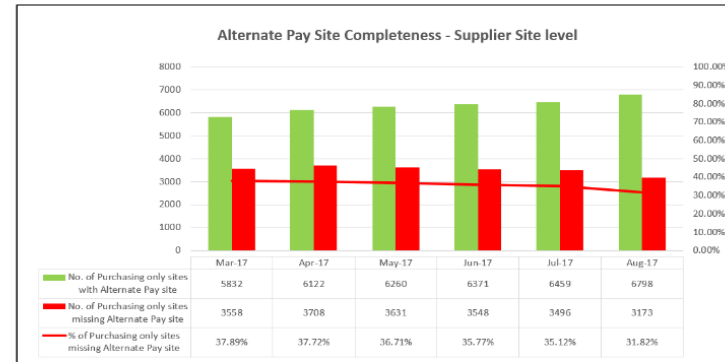
Status	System Name	DQ Dimension	Rules	Red	Yellow	Green	Comments
●	Supplier Data Hub	Supplier - Uniqueness	2			2	Detective controls: 2; Preventative Controls: 0
●	Supplier Data Hub	Supplier - Validity	35	1	2	32	Detective controls: 9; Preventative Controls: 26 Refer Supplier DQ dashboards for DQ improvement status of detective controls.
●	Supplier Data Hub	Supplier - Completeness	20	6			Detective controls: 10; Preventative Controls: 10
●	Supplier Data Hub	Supplier - Accuracy	5				

Green: Error <= 5% or <=Agreed target
Yellow: Error = 5- 15% or 5 - 15% of Agreed target
Red: Error > 15% or >(Agreed target -15%)

For Detailed Supplier data quality dashboards - [click here](#)

Alternate Pay Site Completeness – Aug'17

DQ Rule: Alternate Pay Site must be assigned to all supplier sites which are selected as Purchasing only Sites.



2017 Q3/Q4 DQ Target: Reduce error count by 1000~ Intercompany and non US operating unit site errors. Target for US OU sites - TBD

DQ Resolution status: mass upload completed for Intercompany supplier sites in Aug'17. Remaining Supplier sites with spend in 2016/17 shared with regional team for review and correction.

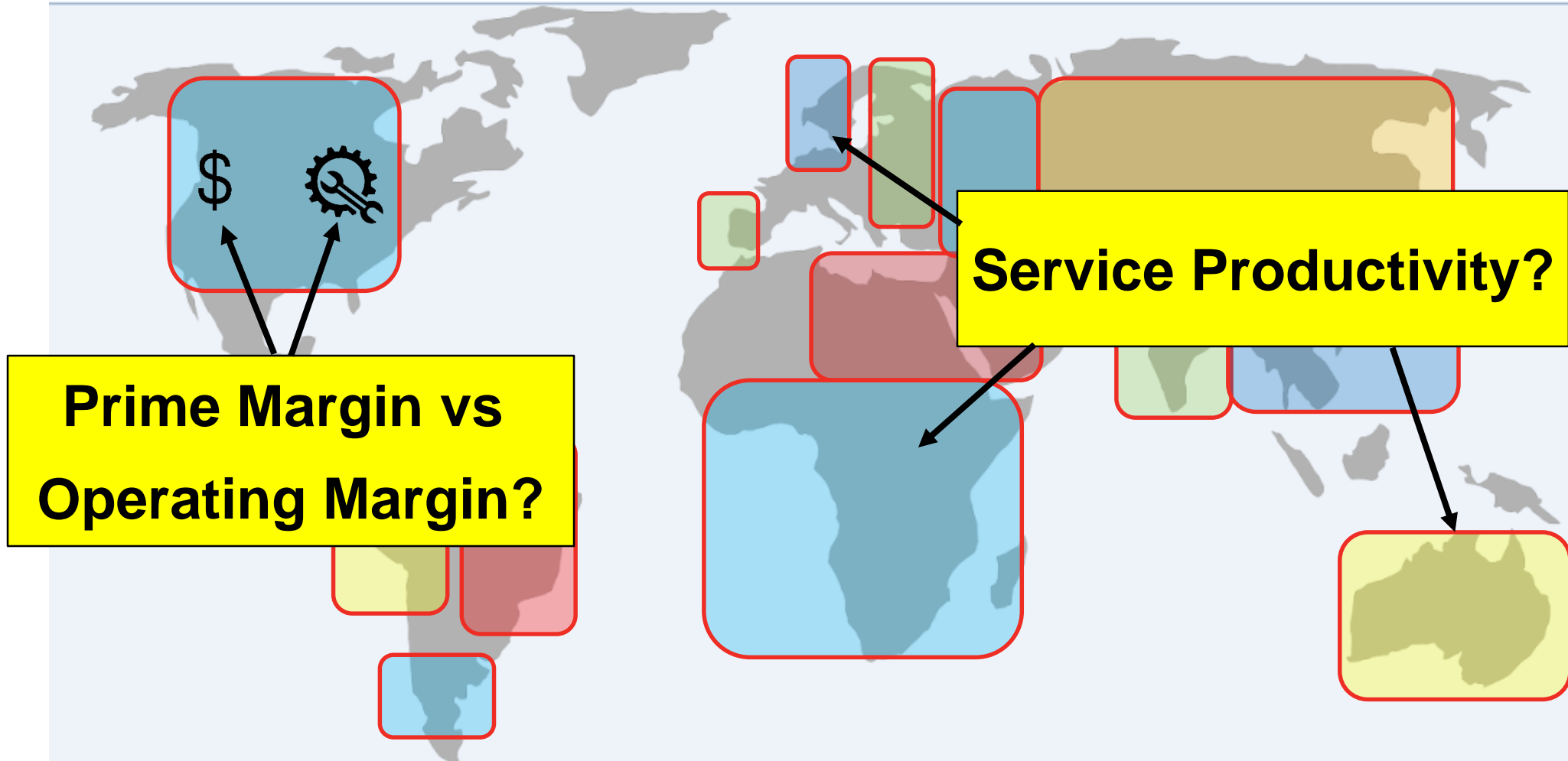


Data Quality Control Services

The screenshot shows a web application interface for account creation. The top navigation bar includes 'File', 'Edit', 'View', 'Navigate', 'Query', 'Tools', and 'Help'. Below the navigation bar, there are several tabs for account management, including 'Home', 'Accounts', 'Account Plan', 'Activities', 'Sales Order', 'Vehicles', 'Contacts', 'Service Requests', 'Products', 'Leads', 'Opportunities', 'Call Reports', 'Quotes', 'Assets', and 'Support'. The main content area is titled 'New Account Creation' and contains a form for 'Account Details'. The form includes fields for Name, Main Phone, Main Fax, Address Line 1-4, City, County, State, Country, and Zip Code. Below the form, there are two sections: 'Standardized Address' and 'Recommended Address'. The 'Recommended Address' section is highlighted with a red box, and an arrow points from the 'Standardized Address' to it. The 'Recommended Address' section includes fields for Address Line 1-4, City, State, Country, and Zip Code, along with a 'Verification Messages' field and a 'DD Error' checkbox. The 'Verification Messages' field contains the text 'Partially Verified: 2 - Locality, Identified and parsed, Prima'. The 'DD Error' checkbox is unchecked. The 'DD Error Message' field is empty. At the bottom of the form, there are two buttons: 'Use Standardized' and 'Use Recommended'.

Real-time address validation and search/match services

Business Alignment for Analytics

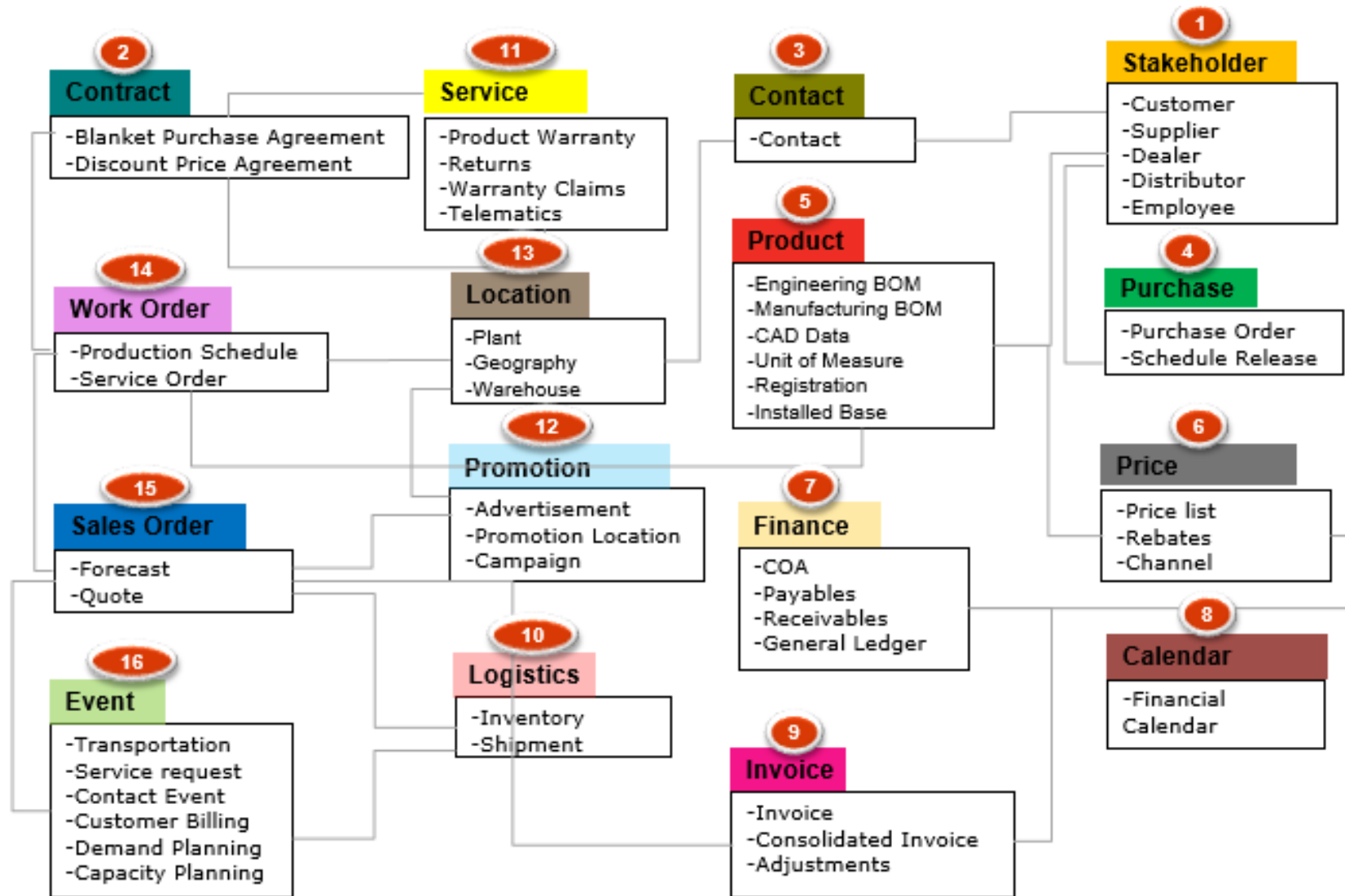




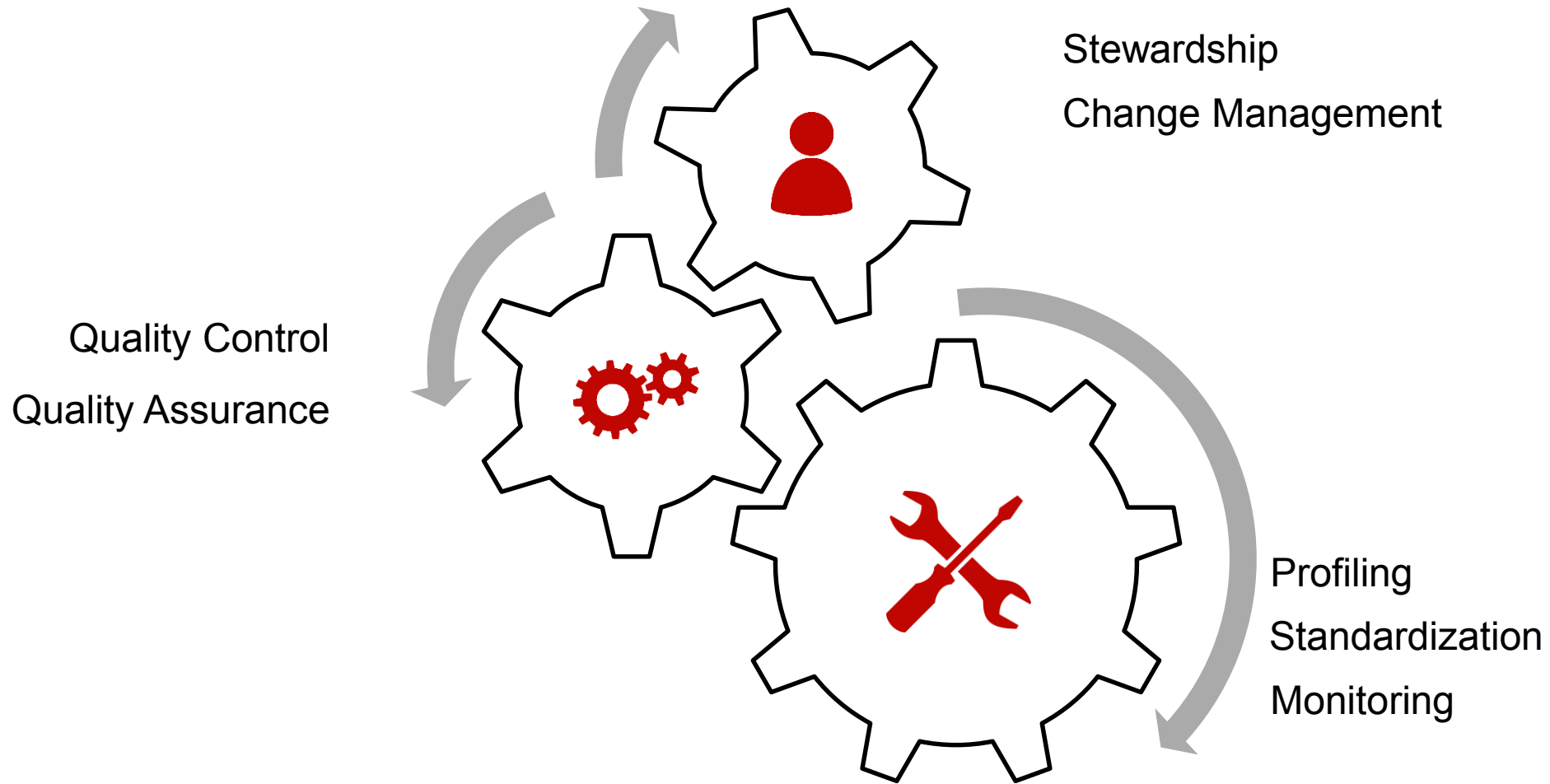
What's Next?



Not Just Master Data



Data Quality As a Service



Governing the Lake

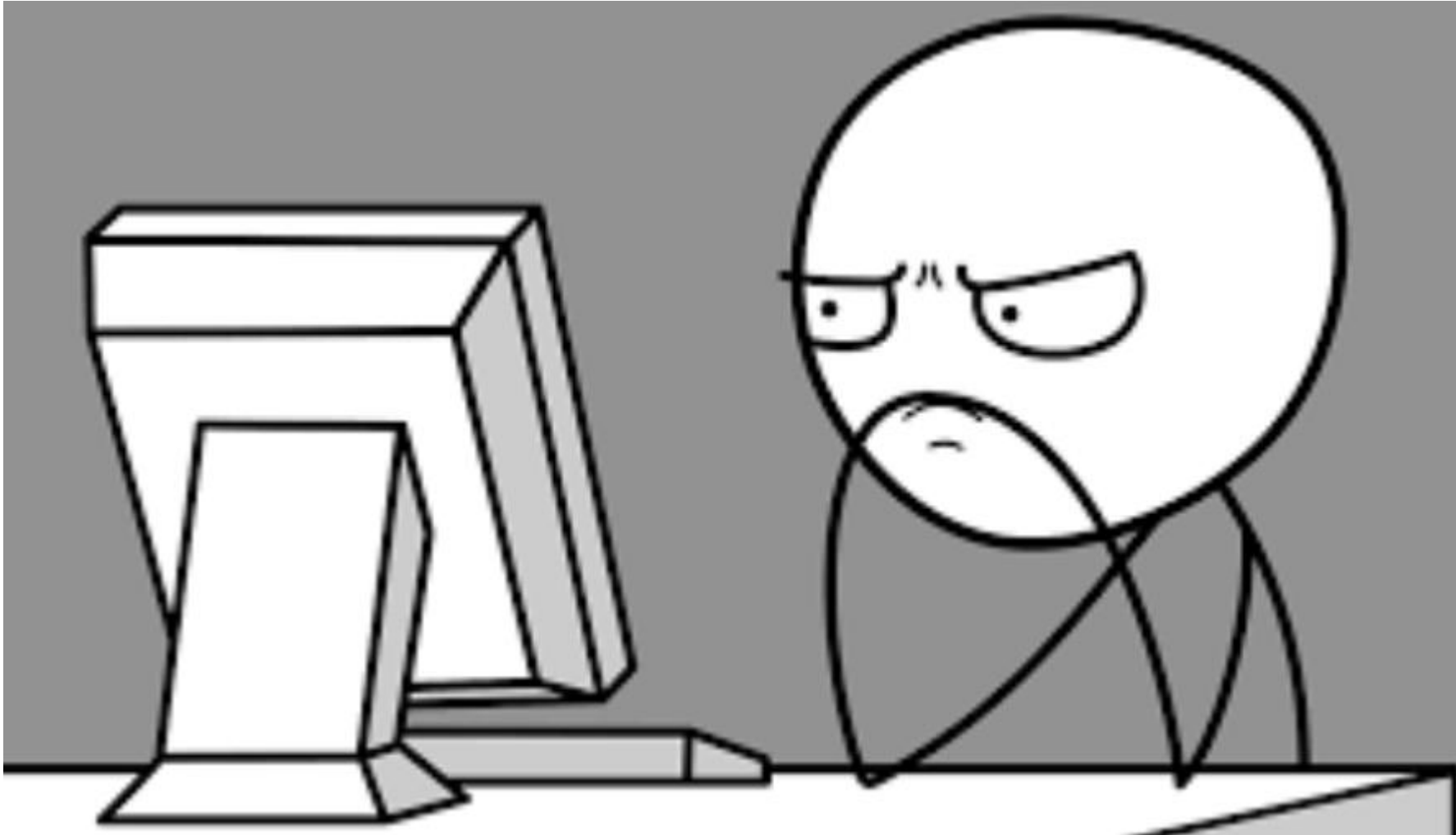




What's The Value?



More Work For Me?



How Quickly Can We Do This?



Get a sneak peek at cutting-edge data integration designs and receive a free gift!

- Oracle is constantly developing new software and features that will make your work easier, and Oracle's User Experience team would love to get your feedback on new data integration designs.
- Feedback sessions will take place at a date and time of your own choice.
- You can take part via webconference, from the comfort and convenience of your own office.
- If you're interested, please fill out the 1-page form at <http://bit.ly/2vIH1Sg>
uppercase I lowercase l
- To show our appreciation, we will post all participants their choice from a wide selection of thank-you gifts.

Data Integration Programme – [FOCUS ON DOC LINK](#)

Presentations on:

Oracle
Data Integration
Platform Cloud

Oracle
Data Integrator

Oracle
GoldenGate

Oracle
Enterprise
Data Quality

Oracle
Enterprise
Metadata
Management

Hands-on Labs:

Oracle GoldenGate
Real-Time Data Replication
in the Cloud
HOL7715

Oracle
Enterprise Data Quality
HOL7653

ODI and OGG
for Big Data
HOL7708

Oracle
Data Integration
Platform Cloud
HOL7673

Demo Stations:

The EXchange
Integration Area
- Moscone West

The EXchange
Data Management Area
- Moscone West

The EXchange
Analytics & Big Data Area
- Moscone West

Data Integration Programme – FOCUS ON DOC LINK

Sunday, October 1

- Lift and Shift Workloads to Cloud with Oracle Data Integration Platform Cloud [SUN6653]
- Data Movement between On-Prem, Fusion ERP Cloud, Fusion HCM Cloud and Salesforce [SUN7286]
- Accelerate Migration to Cloud Infrastructure with Data Integration Platform [SUN6896]

Monday, October 2

- Oracle Data Integration Platform Strategy and Roadmap [CON6646]
- Filling Your Data Lake with Potable Data, Using Data Integration [CON5465]
- GoldenGate : Deep Dive into Automating OGG using the new Microservices [CON6569]
- Oracle Data Integration Platform: Foundation for Cloud Integration [CON6650]
- Oracle Data Integration Platform Empowers Enterprise Grade Big Data Solutions [CON6893]
- Oracle Data Integration Platform Cloud Deep Dive [CON6651]
- Oracle GoldenGate Cloud Service: Real-Time Data Replication in the Cloud [HOL7715]

Tuesday, October 3

- Oracle Data Integrator Product Update and Strategy [CON6654]
- Oracle Enterprise Data Quality: Product Overview and Roadmap [CON6656]
- Accelerate Cloud On-Boarding Using Oracle GoldenGate Cloud Service [CON6894]
- Oracle Enterprise Data Quality for All Types of Data [HOL7653]
- Oracle Data Integration Platform: a Cornerstone for Big Data [CON6655]
- GoldenGate: MAA and Best Practices for Oracle GoldenGate Microservices [CON6570]
- Oracle GoldenGate Product Update and Strategy [CON6897]

Wednesday, October 4

- A Practical Path to Enterprise Data Governance with Oracle Enterprise Data Quality [CON6657]
- Oracle Data Integrator and Oracle GoldenGate for Big Data [HOL7708]
- Introduction to Oracle Data Integration Platform Cloud [HOL7673]
- An Enterprise Databus: GoldenGate in the Cloud Working with Kafka and Spark (CON6895)
- GoldenGate: Best Practices & Deep Dive on OGG 12.3 Microservices at Cloud [CON6568]
- Oracle GoldenGate for Big Data [CON6898]
- **Oracle Data Integration Platform Cloud Service Governance Edition [CON6652]**

Connect with Oracle Integration



Oracle Data Integration



@OracleDI



[Blogs.oracle.com/DataIntegration/](https://blogs.oracle.com/DataIntegration/)



Oracle Data Integration



Oracle FMW



@OracleIntegrate



[Blogs.oracle.com/Integration/](https://blogs.oracle.com/Integration/)



Oracle SOA

ORACLE®

Integrated Cloud

Applications & Platform Services