

12월 9일

클라우드 시대 데이터 관리환경의 현대화

김상철

인포매티카 파트너 기술담당



Informatica®

발표자 소개



김상철 Partner Technical Consultant

- 현 인포매티카 KOREA POD 기술담당
- Data Integration , Data Management Solution 전문가
- APAC내 금융권, 전자통신 업체 Data Warehouse와 Data Analytics Solution 구축

인포매티카 소개

Informatica는 데이터 중심 디지털 변혁을 앞서 추진하고 있는 유일한 엔터프라이즈 클라우드 데이터 관리 리더입니다. Informatica는 지난 25년 동안 9500개 이상의 고객사가 데이터의 힘을 활용하도록 지원했습니다.

혁신 기술

100개 특허 보유 | 8곳 전 세계 R&D 센터 | 196억 8천만 달러 연간 R&D 투자

Informatica Intelligent Data Platform™은 마이크로서비스 아키텍처에 구축된 업계에서 가장 완벽한 모듈형 솔루션으로, 기업이 하이브리드 엔터프라이즈 전반에서 모든 데이터의 힘과 가치를 활용할 수 있도록 도움을 줍니다.

AI 중심 플랫폼은 신뢰할 수 있고 안전하며, 관리되고 액세스할 수 있으며, 관련성 높고 실행 가능한 적시 데이터를 보장함으로써 어디서든 온 프레미스, 다중 클라우드 및 빅 데이터를 아우릅니다. 이를 통해 세계에서 가장 진보적인 회사들은 데이터 중심 디지털 변혁의 성과를 달성할 수 있습니다.

CLAIRE®

머신 러닝 및 기타 AI 기술을 사용하는 CLAIRE® 엔진은 Informatica Intelligent Data Platform의 뛰어난 메타데이터 기능을 활용하여 핵심 데이터 관리 및 거버넌스 프로세스를 가속화하고 자동화합니다.



Intelligent
Cloud Services



Intelligent
Data Engineering



Intelligent
Data Integration



Intelligent
Data Quality



Intelligent
Data Privacy



Intelligent
MDM

최근의 혁신

- 차세대 iPaaS 재정의
- 데이터 거버넌스, GDPR 및 데이터 프라이버시 솔루션 출시
- Enterprise Data Catalog 출시
- MDM Customer 360 솔루션 출시
- CLAIRE 도입
- Enterprise Data Preparation 출시
- Data Engineering 제품군 출시

Quick Facts

- 1993년 설립
- 본사 소재지: 캘리포니아 레드우드 시티
- CEO, Amit Walia
- 2015년, Permira 및 Canada Pension Plan Investment Board에 약 53억 달러에 인수됨
- 25년간 82개 국가에서 9,500개 이상의 고객사 지원
- 고객사 목록에 포춘 100대 기업 중 85개 기업이 포함됨
- 연간 수익 13억 달러 이상
- 4,800명 이상의 직원
- 모든 에코시스템 전반에 걸쳐 매월 10조 이상의 클라우드 거래 처리 중

Intelligent Data Platform



MULTI-CLOUD



BIG DATA



REAL TIME/
STREAMING



TRADITIONAL

Solutions



ENTERPRISE
DATA PREPARATION



ENTERPRISE DATA
GOVERNANCE



CUSTOMER
360



CUSTOMER
360 INSIGHTS



PRODUCT
360



SUPPLIER
360



REFERENCE
360



DATA PRIVACY
MANAGEMENT

Products



DATA
INTEGRATION



iPaaS



DATA
ENGINEERING



DATA QUALITY &
GOVERNANCE



MASTER DATA
MANAGEMENT



ENTERPRISE
DATA CATALOG



DATA
PRIVACY

CLAIRE™



MONITOR AND MANAGE



DATA ENGINE



CONNECTIVITY

A Leader in Five Gartner Magic Quadrants

Informatica is placed furthest in completeness of vision and highest in ability to execute in ALL of these Magic Quadrants

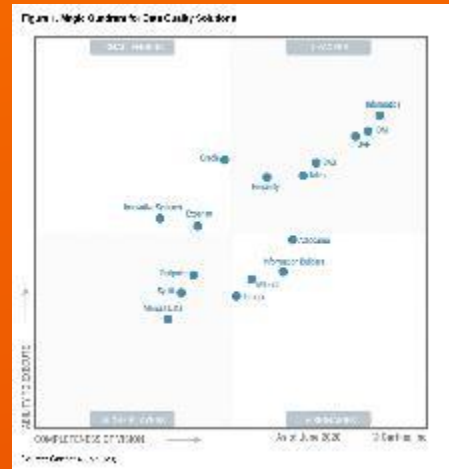
Enterprise Integration Platform as a Service



Sep 2020

Eric Thoo, et al.,
21 Sep 2020

Data Quality Solutions



July 2020

Melody Chien, et al.,
27 July 2020

Master Data Management Solutions



Jan 2020

Simon Walker, et al.,
13 Jan 2020

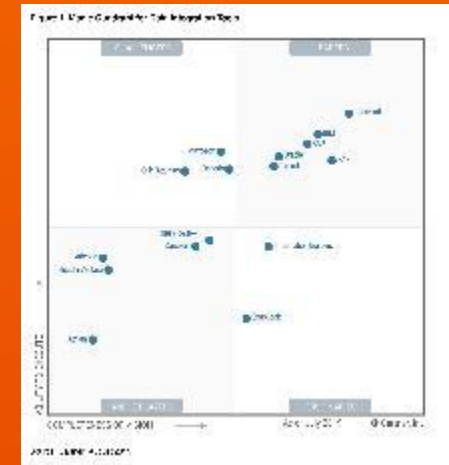
Metadata Management Solutions



Nov 2020

Guido De Simoni, et al.,
16 Oct 2019

Data Integration Tools



Aug 2020

Ehtisham Zaidi, et al.,
18 Aug 2020

These graphics were published by Gartner, Inc. as part of larger research documents and should be evaluated in the context of the entire document. The Gartner documents are available upon request from Informatica. Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

Today's Agenda

현대 데이터 분석환경 트렌드

인포매티카 클라우드 서비스와 아키텍처

인포매티카 데이터 카탈로그

고객 사례

데모

현대 데이터 분석환경 트렌드



데이터 3.0 시대



Explosion in Data Volume

20.6 zettabytes per year in global data center traffic



New Users

500 million business data users and growing



New Data Types

(mobile, social, IoT)

20 billion connected devices



Data in the Cloud

Over 94% of data center traffic will come from the Cloud



Machine Learning/AI

1 billion workers will be assisted by machine learning or AI



Informatica

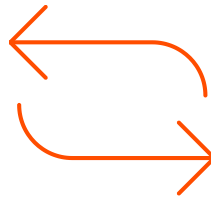
클라우드 : 디지털 전환의 중심



83%

of enterprise
workloads will be in
the **cloud** by 2020

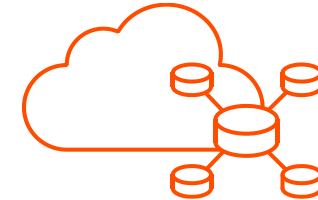
Source: Forbes, Logic Monitor Cloud Adoption



83%

of organizations expect
workloads to **move freely**
between **clouds**

Source: Turbonomic's 2019 State of Multicloud

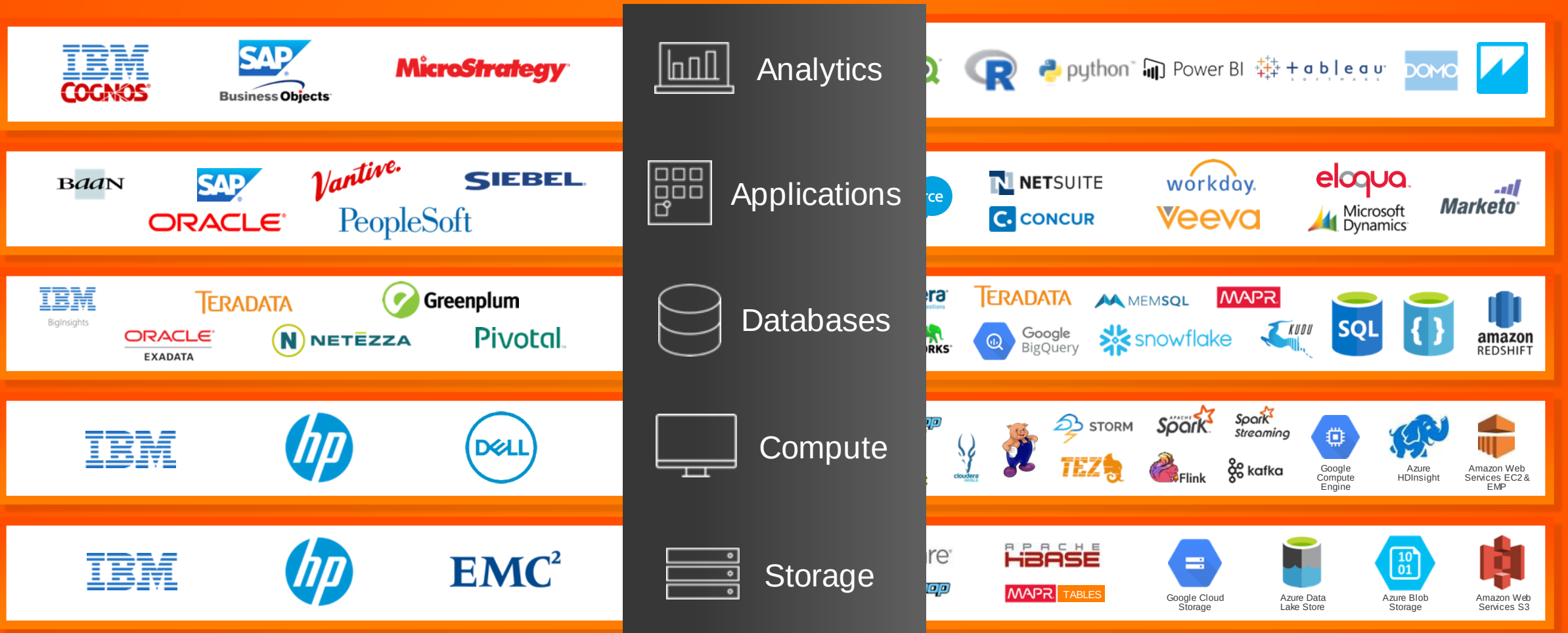


75%

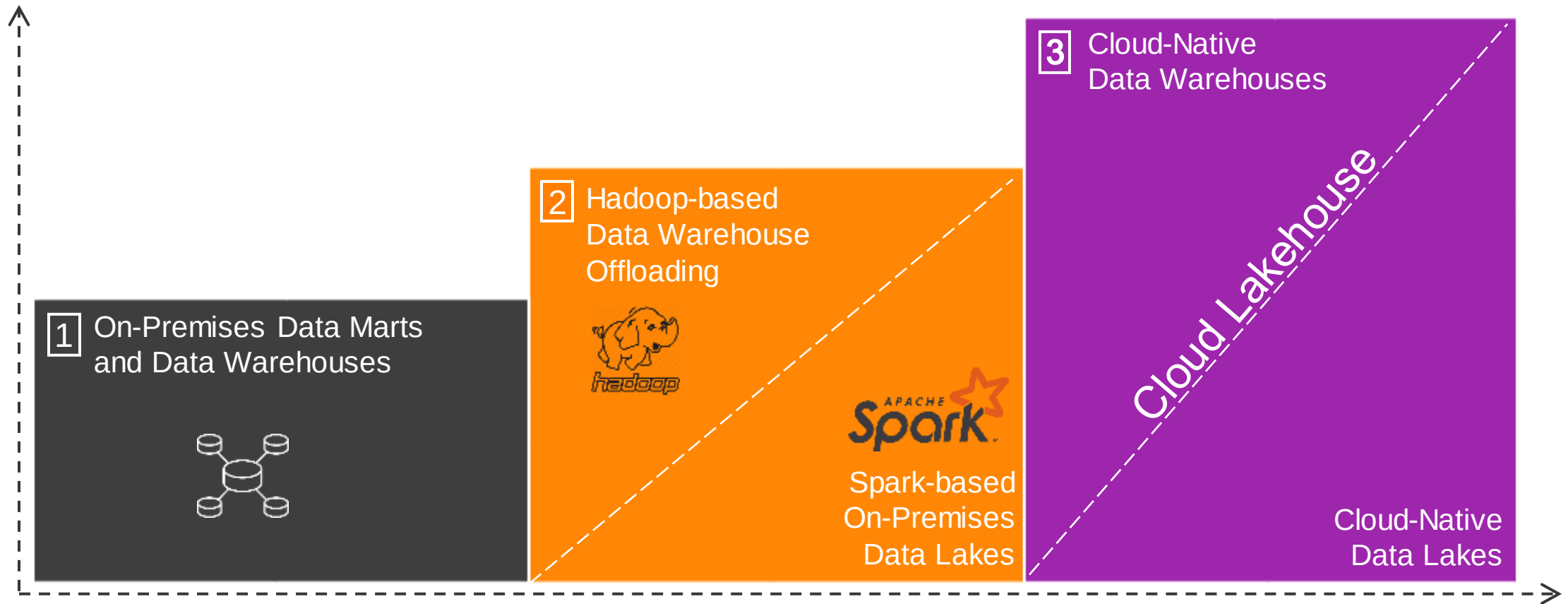
of all **databases** will be
deployed or migrated to a
cloud platform by 2022

Source: Gartner, The Future of the DBMS Market Is Cloud by
Adam Ronthal, Donald Feinberg, Merv Adrian | June 23, 2019

새로운 데이터 분석환경



데이터 분석환경의 변화



Déjà Vu – 같은 문제점의 반복 : Hand Coding

SQL Query

```
select l_orderkey, sum(l_extendedprice * (1 -  
l_discount)) as revenue, o_orderdate, o_shippriority  
from CUSTOMER, ORDERS, LINEITEM where  
c_mktsegment = 'AUTOMOBILE' and c_custkey =  
o_custkey and l_orderkey = o_orderkey and o_orderdate  
< date '1995-03-13' and l_shipdate > date '1995-03-13'  
group by l_orderkey, o_orderdate, o_shippriority order  
by revenue desc, o_orderdate limit 10;
```

Spark Code

```
package main.scala  
import org.apache.spark.sql.DataFrame  
import org.apache.spark.SparkContext  
import org.apache.spark.sql.functions.sum  
import org.apache.spark.sql.functions.udf  
/**  
 * Query 3  
 */  
class Q03 extends TpchQuery {  
  override def execute(sc: SparkContext, schemaProvider:  
    TpchSchemaProvider): DataFrame = {  
    // this is used to implicitly convert an RDD to a DataFrame.  
    val sqlContext = new org.apache.spark.sql.SQLContext(sc)  
    import sqlContext.implicits._  
    import schemaProvider._  
    val decrease = udf { (x: Double, y: Double) => x * (1 - y) }
```



Proprietary



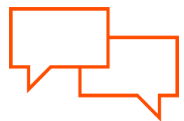
Informatica

Déjà Vu – 같은 문제점의 반복 : 데이터 품질관리

60%

of organizations
challenged by
data quality

Source: IDC



While some data quality issues are unfathomably complex, many yield quickly and produce outsize gains. Eliminating a single root cause can prevent thousands of future errors, save millions, and make things better for all involved.

Redman, Salmon and Nagle – Harvard Business Review

클라우드 데이터 레이크 하우스를 위한 메타데이터 관리

DATA ARCHITECT

How can I get complete view of our data landscape?



DATA CONSUMER

How can I find, discover, understand data required for my analysis?



GOVERNANCE OFFICE

How can we validate and enforce our data governance policies and definitions?



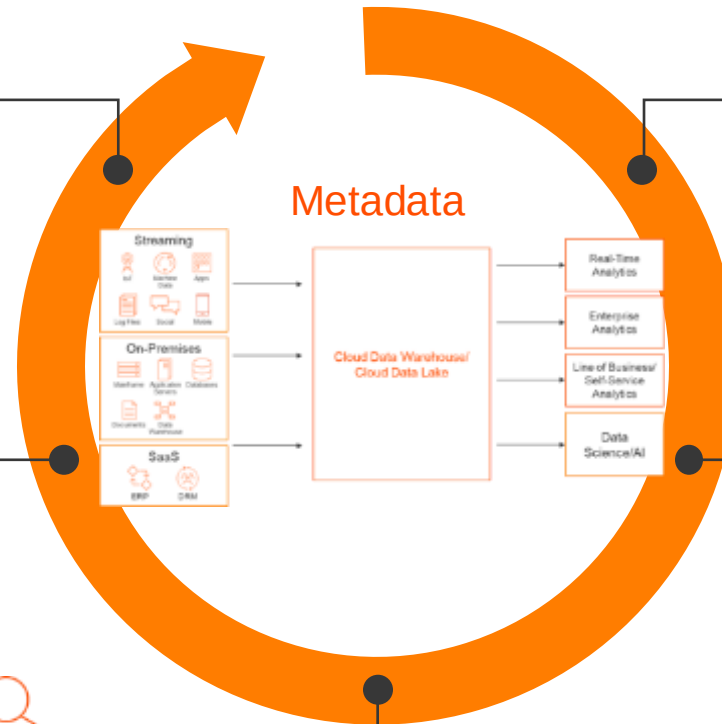
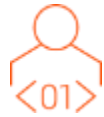
DATA STEWARD

How can I manage metadata for key enterprise data assets?



DATA WAREHOUSE DEVELOPER

How can I find technical and business metadata about source systems and mappings?



If you don't understand your assets, how do you use them? Need visibility to the metadata for those data assets...provide visibility into data lineage.

클라우드 데이터 관리

DATA INTEGRATION



- Codeless integration
- Mass ingestion for files, DB and streaming
- Push-down optimization
- Serverless and elastic scaling
- Spark-based processing in the cloud
- Broad connectivity
- Stream processing
- MLOps

DATA QUALITY



- Data profiling
- Data quality rules
- Dictionaries to manage values lists
- Cleansing, standardization, parsing, verification and deduplication/consolidation processes
- Integrated into data integration
- Data quality analytics

METADATA MANAGEMENT



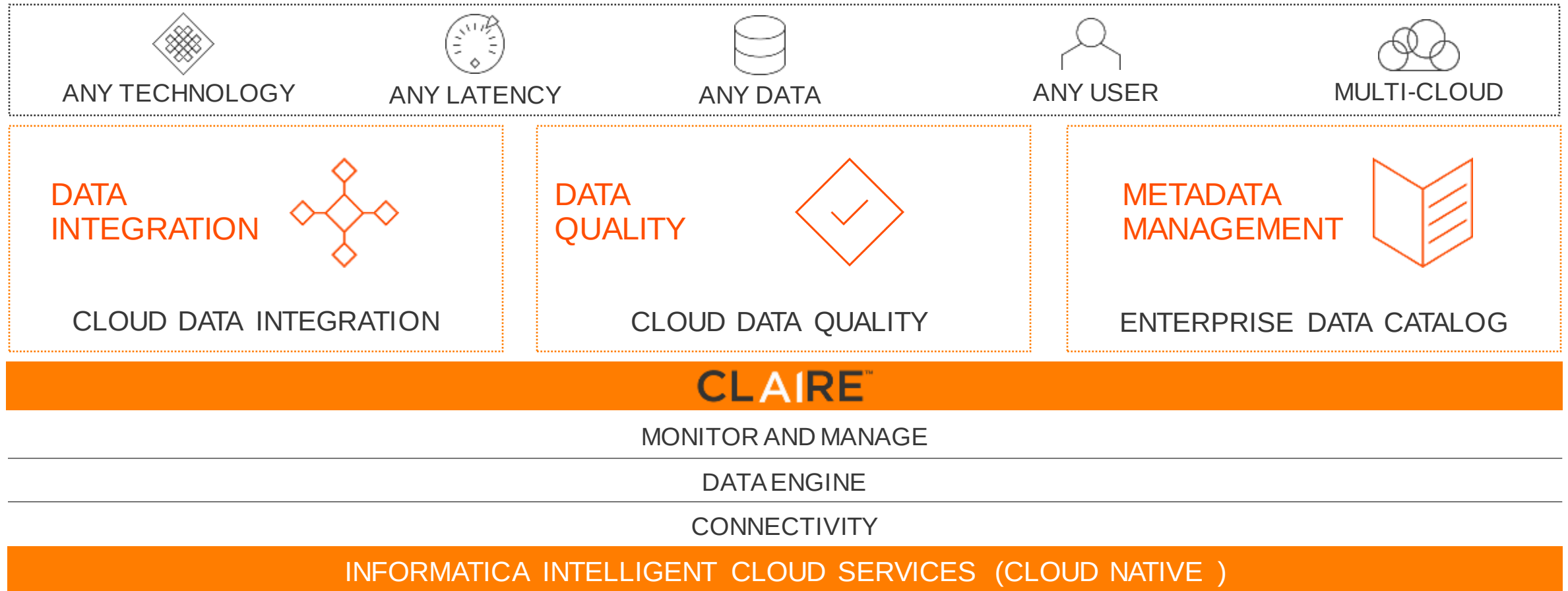
- Data discovery
- End-to-end lineage
- Metadata – technical, business, operational, usage
- Connect and scan metadata – databases (DW, DL), apps, ETL, BI tools and others
- Common metadata foundation

AI-DRIVEN AUTOMATION

CLOUD NATIVE

Multi-cloud • API driven • Microservices • Containerization • Serverless architecture
Minimal install and setup • Auto-upgrades • Usage-based pricing • Trust certifications
Integrated full-stack high availability and advanced security (i.e., platform, network, infrastructure)

인포매티카 클라우드 서비스 (IICS)



데이터 카타로그: 필요한 데이터 검색, 분류, 정리

The image displays two screenshots of the Informatica Data Catalog interface, overlaid on a cityscape background.

Left Screenshot (Main Page):

- Header: Informatica Data Integration
- Left Navigation Menu: New..., Home, Explore, Data Catalog, Bundles, My Jobs, My Import/Export Logs
- Main Content: Enterprise Data Catalog for Cloud Data Integration
- Search Bar: Search
- Text: Search over 1248 assets in the data catalog

Right Screenshot (Search Results):

- Header: Informatica Data Integration
- Left Navigation Menu: Same as the left screenshot
- Search Bar: Search Results: crm
- Filters:

 - Asset Type: All (11), Table (7), DelimitedFile (7)
 - Resource Type: All (10), Oracle (7), HDFS (7), Microsoft SQL Server (1)
 - Resource Name: All (4), ACME_HDFS (4), ACME_Oracle (4), ORACLE_ACME_CRM (4), HDFS_ELEVATE (2), Acme_Retail_ORC_Master,... (1)
 - Rows: N/A (5), Under 50K (1), 100K to 150K (1)
 - Data Domains

- Search Results (1-18 of 18):

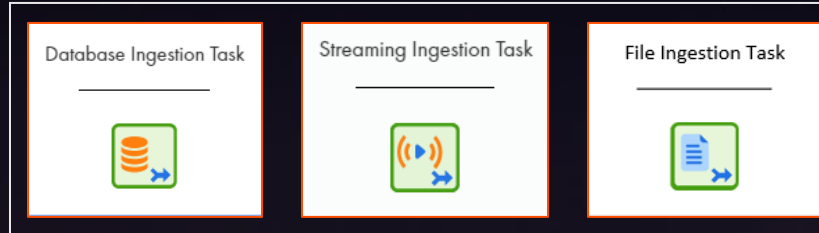
 - CRM_CUSTOMER_MAIN**
Type: Table, Source: Oracle, Last Updated: 2019-05-03T10:05:01.917Z
 - CRM_CUSTOMER_COMMUNICATION**
Type: Table, Source: Oracle, Last Updated: 2019-04-22T19:38:18.914Z
 - CRM_CUSTOMER_DEMOGRAPHICS**
Type: Table, Source: Oracle, Last Updated: 2019-04-22T19:38:18.914Z
 - CRM_CUSTOMER_COMMUNICATION**
Type: Table, Source: Oracle, Last Updated: 2019-05-03T10:06:22.874Z
 - CRM_CUSTOMER_DEMOGRAPHICS**
Type: Table, Source: Oracle, Last Updated: 2019-05-03T10:05:01.917Z
 - DIM_RETAIL_CUSTOMER_ADDRESS**
Type: Table, Source: Oracle, Last Updated: 2019-04-28T23:42:11.092Z
 - CRM_CLIENT**
Type: Table, Source: Microsoft SQL Server, Last Updated: 2019-03-04T15:04:22.549Z

Cloud Data Integration



Informatica

클라우드로의 실시간 데이터 적재: DB, 파일, 스트림



DATABASE INGESTION

The screenshot shows the Informatica Data Integration console with a task named "demo_dbm_mszm_mstload_1". The task is a "Database Ingestion Task" and is currently in a "Completed" state. The configuration details are as follows:

- Job Name:** demo_dbm_mszm_mstload_1
- Task Name:** demo_dbm_mszm_mstload
- Task Type:** Database Ingestion Task
- Load Type:** Initial Load
- Task Location:** QA Demo
- Start Time:** Dec 6, 2019, 3:52:32 AM
- Records Read:** 10000
- Records Written:** 10000

The "Source" tab is selected, showing the following configuration:

- Connection:** Oracle Staging DB - Patient Data
- Schema:** PATIENT_STAGING
- Table Selection Rules:** Operator: Include, Condition: *
- Advanced:** Restart Point for Incremental Load: Default

STREAMING INGESTION

The screenshot shows the Informatica Data Integration console with a task named "Ingest_Kafka_To_Amazon_S3". The task is a "Streaming Ingestion Task" and is currently in a "Up and Running" state. The configuration details are as follows:

- Job Name:** Ingest_Kafka_To_Amazon_S3
- Task Name:** Ingest_Kafka_To_Amazon_S3
- Task Type:** Streaming Ingestion Task
- Load Type:** Initial Load
- Task Location:** QA Demo
- Start Time:** Dec 6, 2019, 3:52:32 AM
- Records Read:** 10000
- Records Written:** 10000

The "Target" tab is selected, showing the following configuration:

- Connection:** Kafka
- Topic:** DemoTopic1
- Advanced Properties:** Producer Configuration Properties: key1=value1, key2=value2; Metadata Fetch Timeout in milliseconds: 5000; Batch Flush Size in bytes: 1048576

FILE INGESTION

The screenshot shows the Informatica Data Integration console with a task named "ADLSGH2_FutemFile_mdl_Unm-S101". The task is a "File Ingestion Task" and is currently in a "Completed" state. The configuration details are as follows:

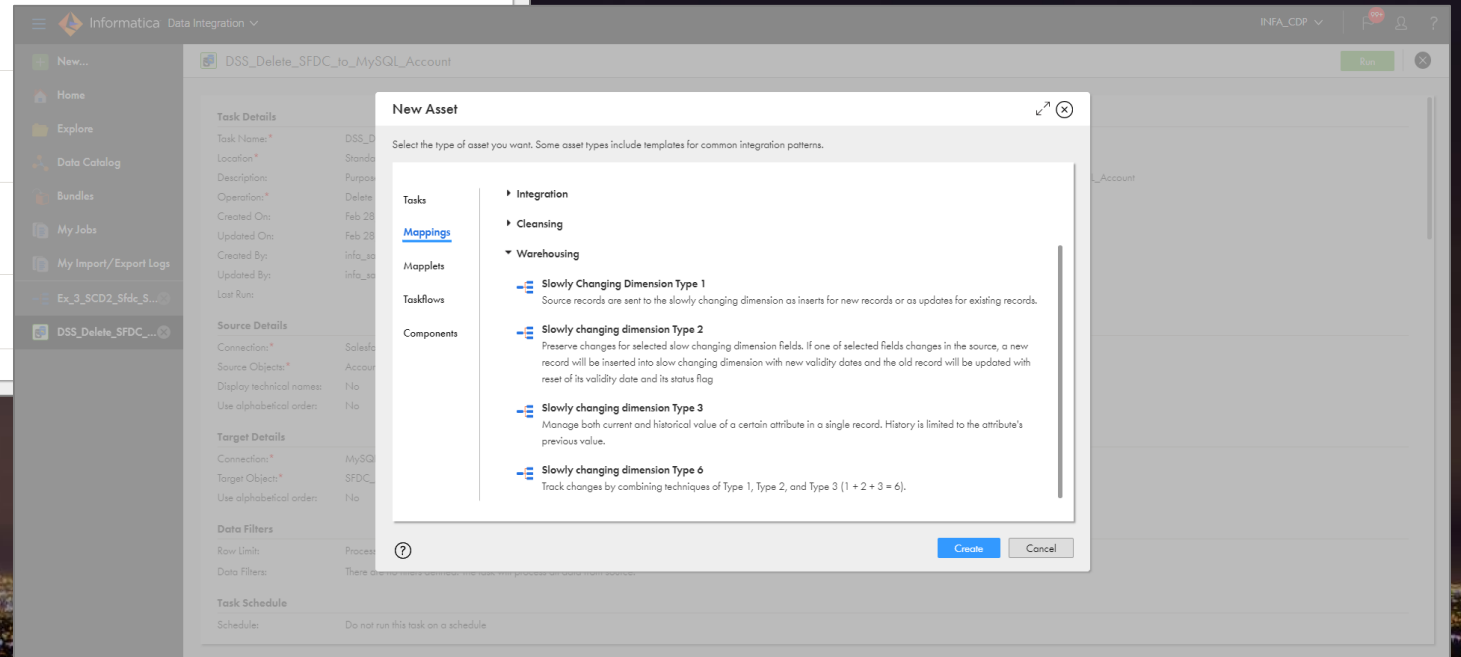
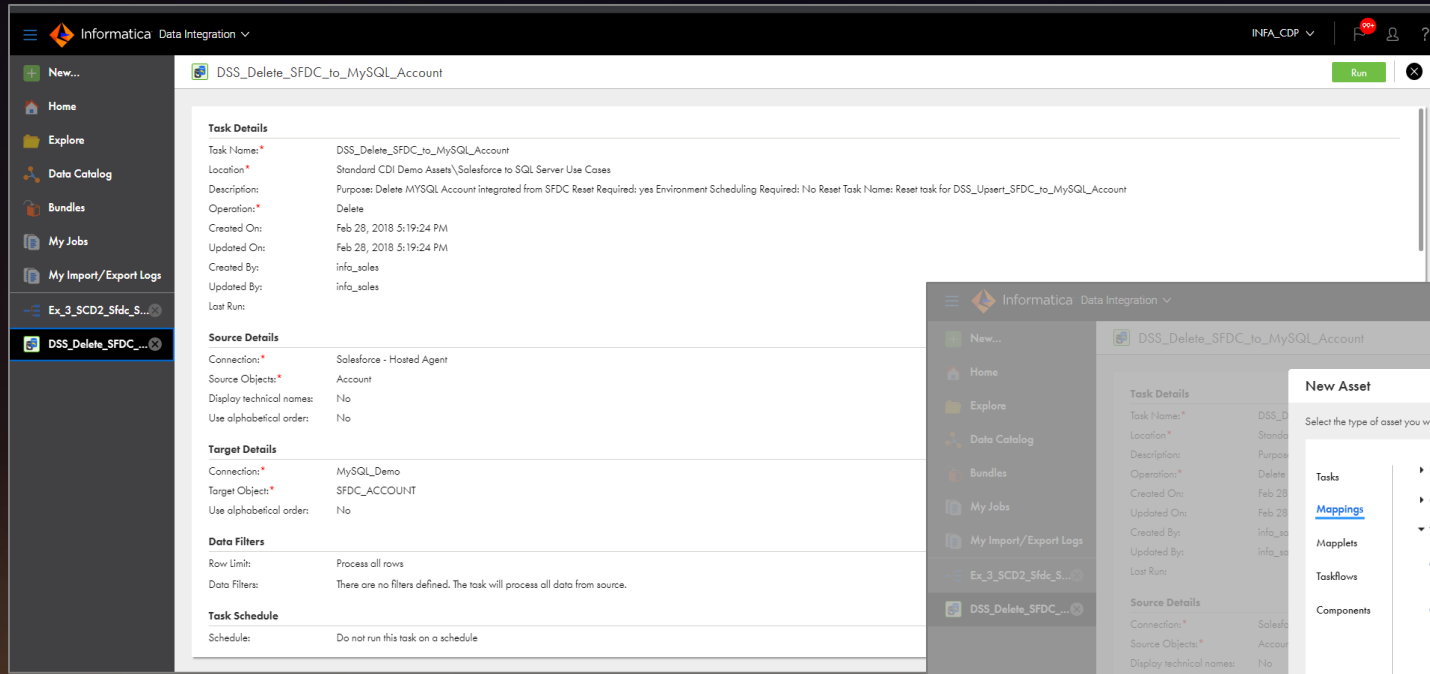
- Job Name:** ADLSGH2_FutemFile_mdl_Unm-S101
- Task Name:** ADLSGH2_FutemFile_mdl_Unm-S101
- Task Type:** File Ingestion Task
- Load Type:** Initial Load
- Task Location:** QA Demo
- Start Time:** Dec 6, 2019, 3:52:32 AM
- Records Read:** 10000
- Records Written:** 10000

The "Source" tab is selected, showing the following configuration:

- Connection Type:** File System
- Connection:** Local File
- Source Options:** Source Directory: /mnt/gh2/101
- Transfer Mode:** File Pattern: *.*

Cloud Data Integration

템플릿과 위저드 기반의 개발 환경

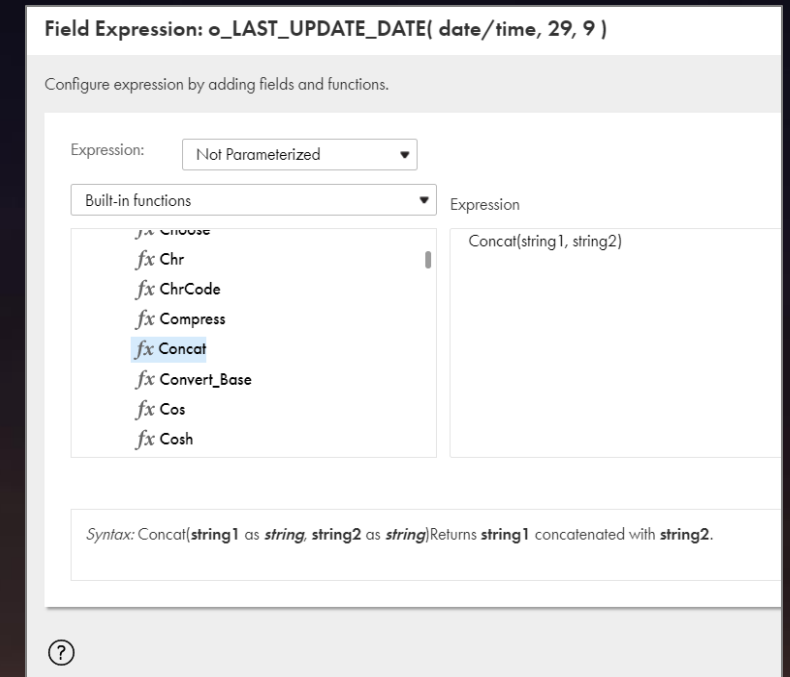
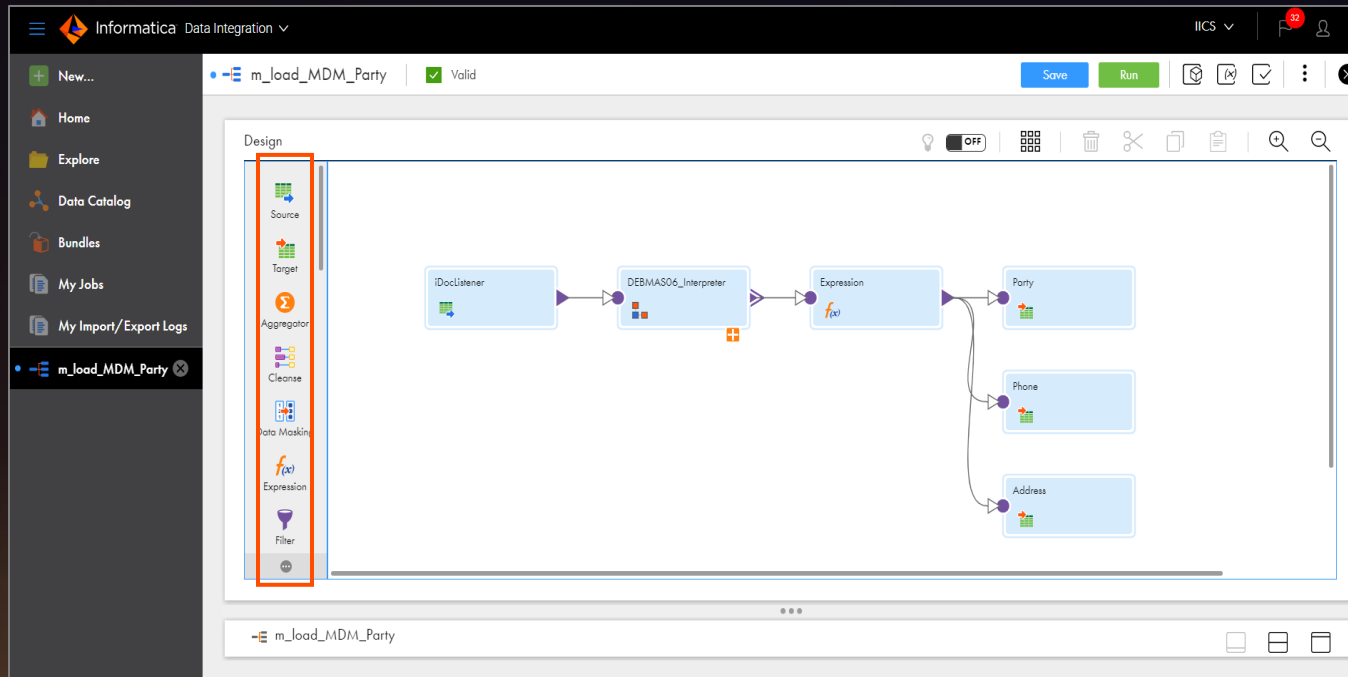


Cloud Data Integration



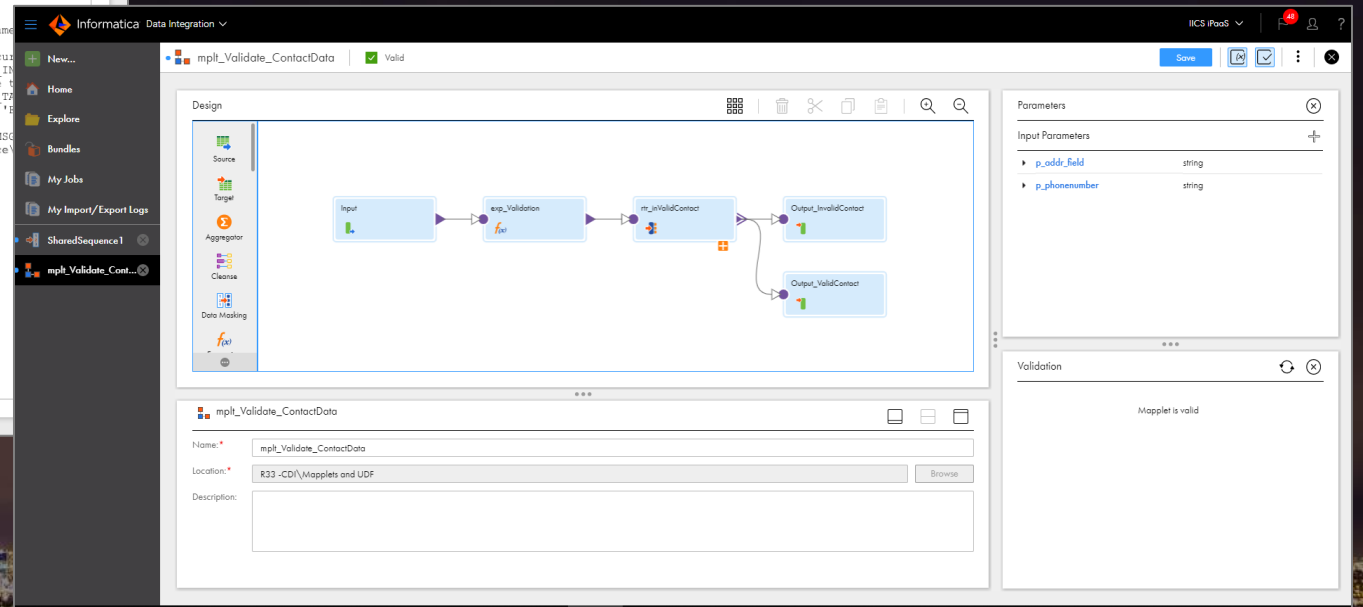
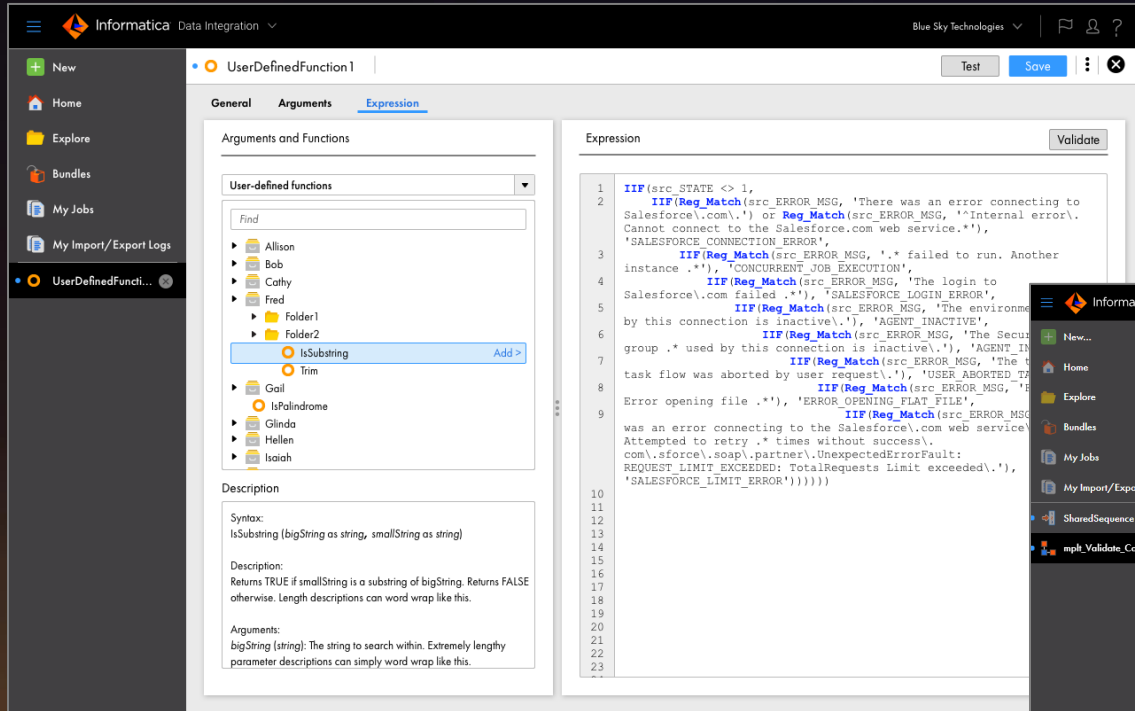
Informatica

수동 코드개발이 필요없는 GUI기반의 개발환경



Cloud Data Integration

맵핑의 재활용과 개발 효율성 증대



Cloud Data Integration



데이터 품질관리: 데이터 정제, 표준화 기능

The Informatica Data Integration interface displays a workflow for data cleansing. The design canvas shows a sequence of steps: Source (m_Customer) → Cleanse_Status → AddressValidation → RuleSpecification → Expression → Target. The Properties panel for the Cleanse_Status step is visible, showing 'Cleanse Options' set to 'Manual' and 'Incoming Fields' mapped to 'Cleanse'.

The Informatica Data Quality interface displays a data quality profile for the 'Superstore' dataset. The profile shows 21 columns, 2 rules, and 9994 rows. The 'Columns and Rules' tab is active, showing a table of data quality metrics for various columns. The 'Alerts' tab is also visible, showing a list of alerts and rules.

Columns	Values Distribution	% Null	% Distinct	% Non-Distinct	# Patterns	% of Top Pattern
Category		0.00%	0.03%	99.96%	3	60.30%
City		0.00%	5.31%	94.68%	9	20.63%
Country		0.00%	0.01%	99.98%	1	100.00%
Customer_ID		0.00%	7.93%	92.06%	1	100.00%
Customer_Name		0.02%	7.93%	92.04%	4	6.84%
Discount		37.45%	4.86%	57.68%	4	39.95%
ID		0.00%	100.00%	0.00%	3	90.00%
Order_Date		0.00%	12.37%	87.62%	1	100.00%
Order_ID		0.02%	50.12%	49.86%	3	96.10%
Postal_Code		0.00%	6.31%	93.68%	2	95.51%
Product_ID		0.00%	18.63%	81.36%	1	100.00%
Product_Name		0.00%	18.11%	81.88%	2	5.72%
Profit		7.51%	57.30%	35.18%	8	19.22%
Quantity		0.00%	13.12%	86.87%	2	76.50%

Data Preview (Random sample)

ID	Order_ID	Order_Date	Ship_Date	Ship_Mode	Customer_ID	Customer_Name	Segment	Country	State	Postal_Code	
1	1	CA-2016-152156	08/11/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Kentucky	42420
2	2	CA-2016-152156	08/11/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Kentucky	42420
3	3	CA-2016-138688	12/06/2016	16/06/2016	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	California	90036
4	4	US-2015-108966	11/10/2015	18/10/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Florida	33311
5	5	US-2015-108966	11/10/2015	18/10/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Florida	33311
6	6	CA-2014-115812	09/06/2014	14/06/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	California	90032
7	7	CA-2014-115812	09/06/2014	14/06/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	California	90032

Cloud Data Quality & Governance



서버 셋업이 필요없는 Spark Processing with Elasticity

The screenshot displays the Informatica Data Integration interface, split into two main panels. The left panel, titled 'Demo_Cluster', shows the 'Basic Configuration' for a new cluster. The right panel, titled 'MyMappingTask', shows the 'Task Mapping' and 'Advanced Session Properties' for a specific task.

Basic Configuration (Left Panel):

- Name: Demo_Cluster
- Description: (empty)
- Runtime Environment: sa_group
- Cloud Platform: Amazon Web Services (AWS)
- Region: US East (Ohio)
- Master Instance Type: m4.2xlarge
- Master Instance Profile: (empty)
- Worker Instance Type: m4.2xlarge
- Minimum Worker Nodes: 2
- Maximum Worker Nodes: 5
- Worker Instance Profile: (empty)
- Availability Zones: Select Availability Zones
- EBS Volume Type: SSD (gp2)
- EBS Volume Size: 150 GB
- Idle Timeout: 5 minutes
- Staging Location: integrationalscaledemo/staging
- Log Location: integrationalscaledemo/logs
- ☒ Encrypt Data

Task Mapping (Right Panel):

- Created By: sabgupta@informatica.com
- Updated By: sabgupta@informatica.com
- Last Run: Nov 15, 2019 8:50:45 PM
- Task Mapping Name: MappingWithAnExpression - Copy 1
- Location: sabira
- Updated Date: Oct 15, 2019 3:55:07 PM

Advanced Session Properties (Right Panel):

Session Property Name	Session Property Value
Session Log File Directory	asdasd
Write Backward Compatible Session Log File	No

☒ Enable cross-schema pushdown optimization

Tuning Recommendation (Right Panel):

Property	Value
spark.driver.extraJavaOptions	some-text
spark.driver.maxResultSize	some-text
spark.driver.memory	some-text
spark.dynamicAllocation.maxExecutors	some-text
spark.executor.cores	some-text
spark.executor.extraJavaOptions	some-text

Cloud Data Integration



Informatica

인공지능을 이용한 맵핑 추천기능

The screenshot displays the Informatica Data Integration (IICS) interface. The main workspace shows a workflow design with a 'WD' (Source) connector, a 'Filter' connector, and a 'SalesforceContacts' (Target) connector. A context menu is open over the 'Filter' connector, listing options: Expression, Router, Lookup, Aggregator, Cleanse, and Data Masking. The 'Properties' panel at the bottom shows the 'Filter' connector's details. A 'Configure Tuning' dialog box is overlaid on the right, providing recommendations for Spark properties to optimize performance. The dialog includes the text: 'CLAIRE Tuning runs the mapping task multiple times in succession to optimize performance. To set resource limits, configure the appropriate Spark properties before you tune the mapping task.' It features a text input field for 'Number of Mapping Task Runs' with the value '25' entered. The dialog is powered by CLAIRE™ and has 'Tune' and 'Cancel' buttons.

CLAIRE™ Powered



인포매티카 클라우드 서비스 (IICS)

DATA
INTEGRATION



CLOUD DATA INTEGRATION

DATA
QUALITY



CLOUD DATA QUALITY

METADATA
MANAGEMENT

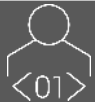


ENTERPRISE DATA CATALOG

Informatica Cloud Data Management Platform



Citizen Integrator
and Business User



Ad Hoc Integrator
LOT IT Developer

User Personas



IT Integration
Expert



Integration
Admin Ops



Common Administration
& Monitoring

Templated and Prepackaged Solutions

Uniform Connectivity

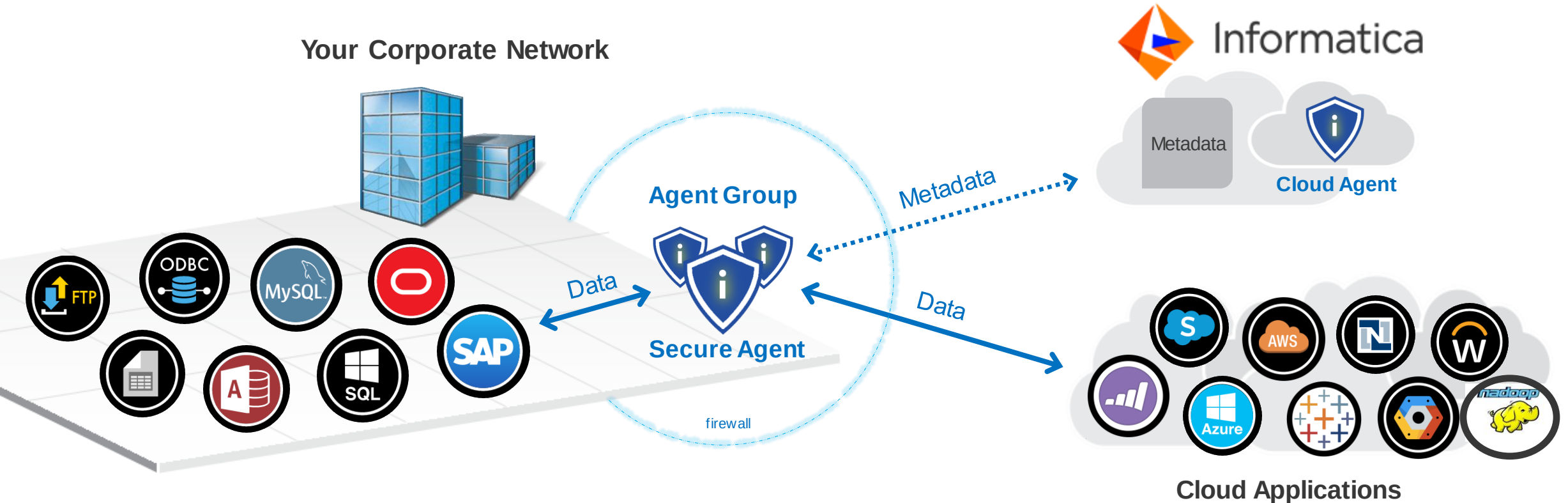
Pluggable Runtime Engines

CLAIRE™



Common Security

Informatica iPaaS: Secure Agent Architecture



- ✓ No staging required
- ✓ Data transmission is secure
- ✓ Multiple security certifications
- ✓ Agent Groups for High Availability
- ✓ Cloud Hosted Agent

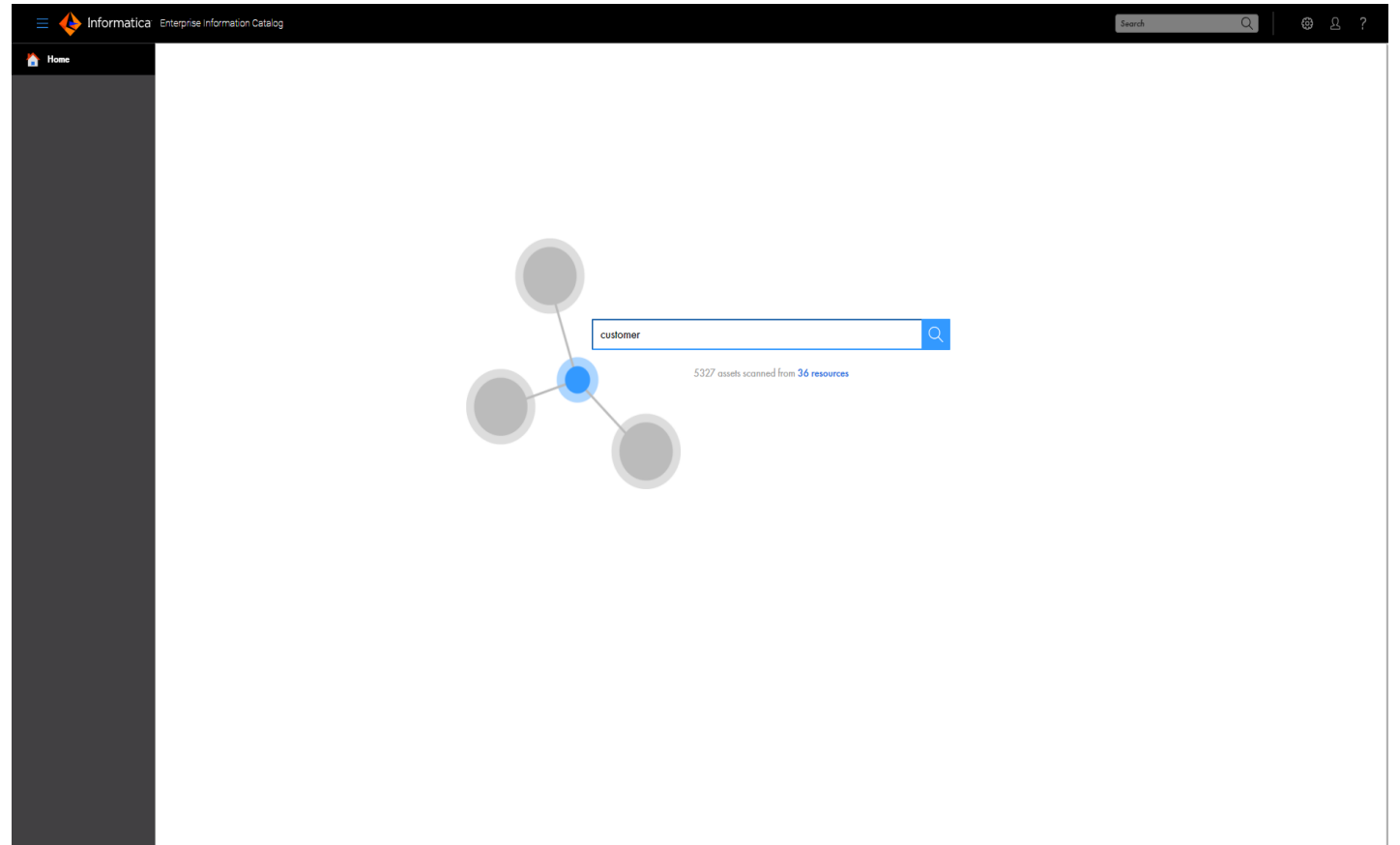
인포매티카 Enterprise Data Catalog



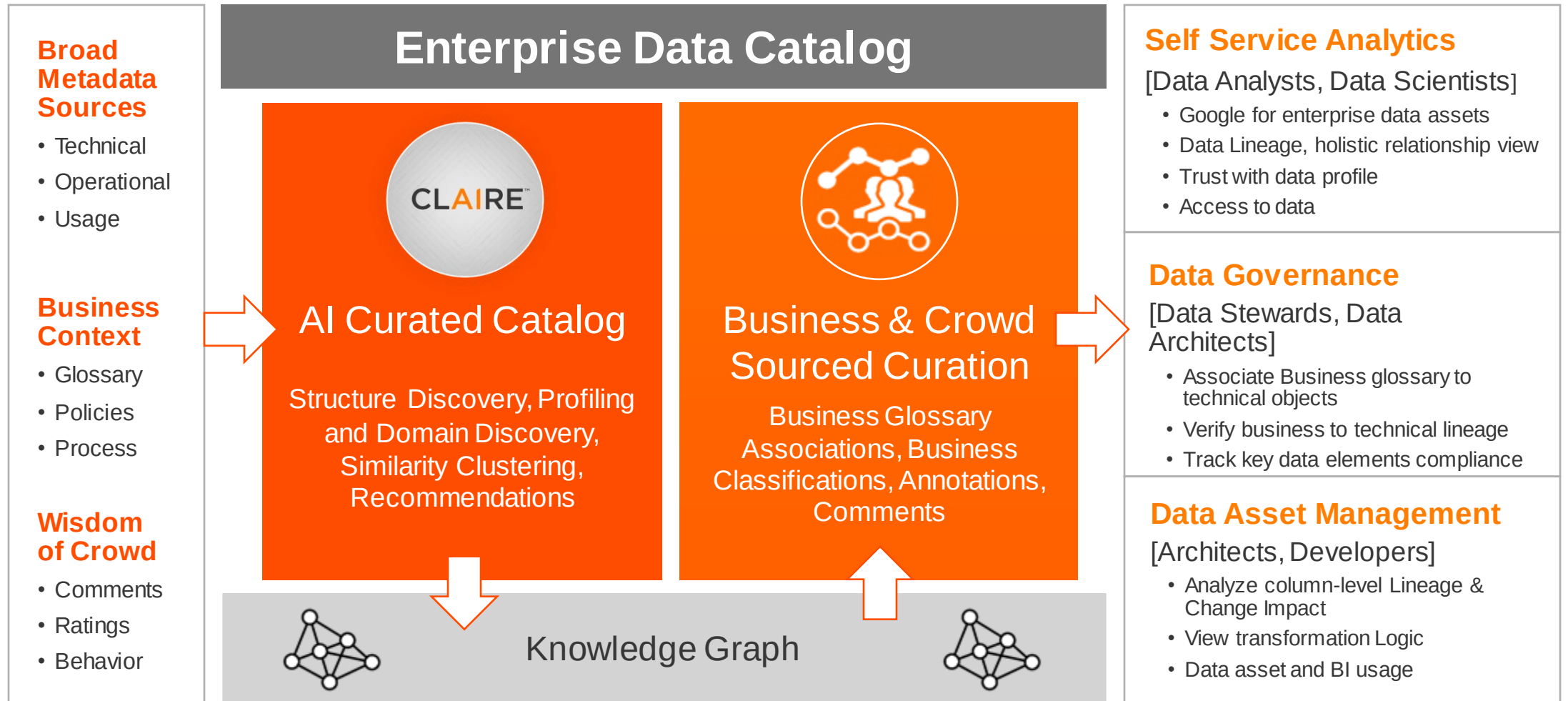
Enterprise Data Catalog

Powered by **CLAIRE™**

Google
like search to
find and **discover**
relevant data in the
enterprise



Enterprise Data Catalog Powered by CLAIRE™



Enterprise Data Catalog

Powered by **CLAIRE™**

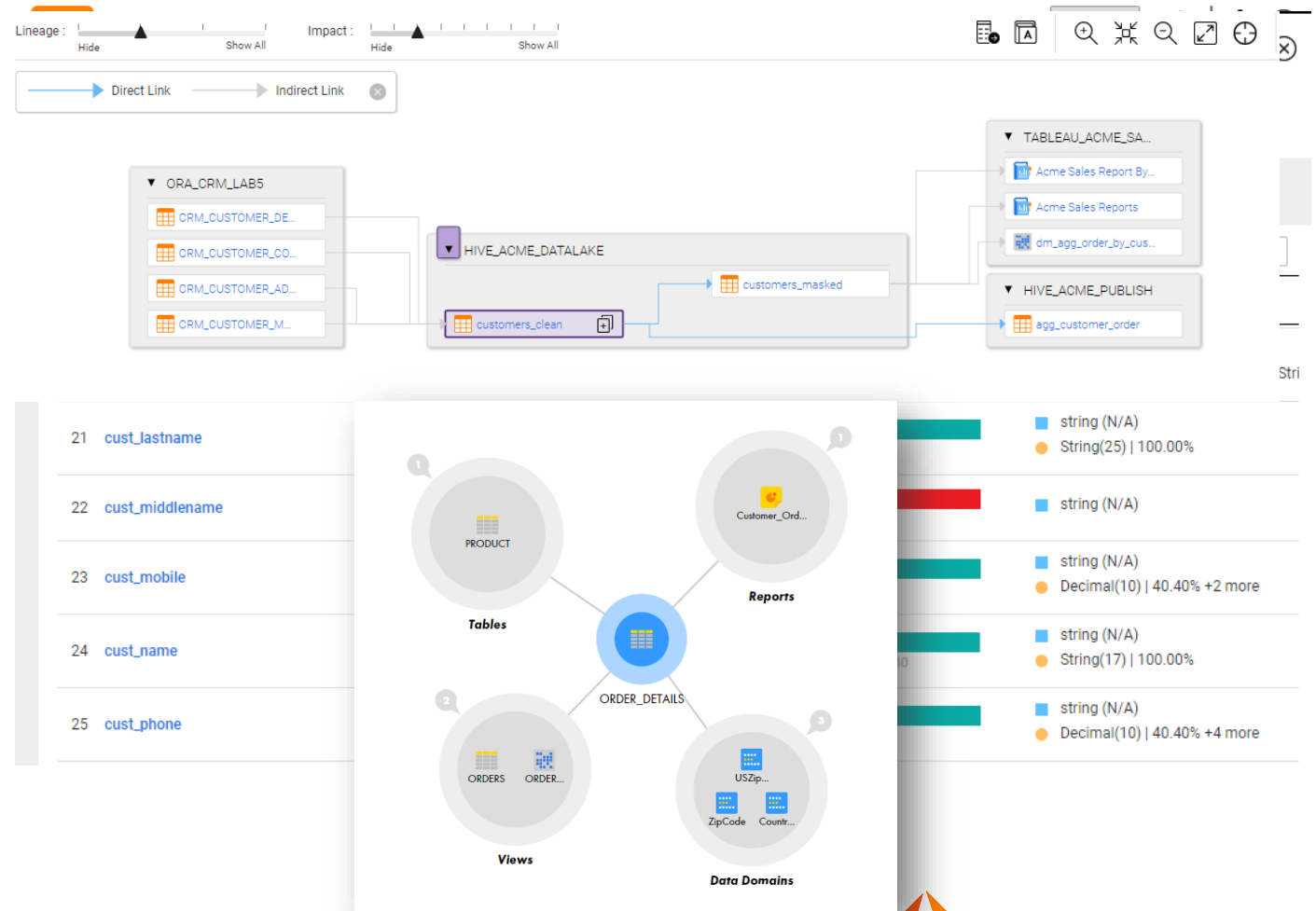
- 직관적인 **검색화면**과 Semantic Search 나 비즈니스 용어를 이용한 검색
- **Data Asset Overview** 와 **데이터 프로파일링** 기능. 그리고 **End-to-End Lineage**를 통한 데이터 영향도 및 관계 파악
- **AI와 머신러닝**을 이용해 자동화한 데이터 분류와 비즈니스 글로서리 태깅
- 엔터프라이즈 시스템을 위한 방대한 **메타데이터 추출 커넥터**
- 사용자들의 **참여와 협업**을 통한 데이터 정보 확장과 큐레이션
- **Open Framework**을 통해 다른 Application과 쉬운 통합과 고객 개발 App과의 쉬운 통합을 위한 제공되는 **방대한 API**

The screenshot displays the Enterprise Data Catalog interface. On the left, a 'Filter by' sidebar contains several sections: 'Certified Assets' with a dropdown set to 'All Assets'; 'Rating' with radio buttons for 'All', '★★★★☆ and above (3)', '★★★★☆ and above (4)', '★★★☆☆ and above (4)', and '★☆☆☆☆ and above (4)'; 'Resource Name' with checkboxes for 'All', 'datalake_cdr (37)', 'HIVE_ACME_DATA LAKE (17)', 'nga_crm (17)', 'nga_dw_redshift (17)', and 'nga_data_lake (10)', plus 'Add' and 'Show All' buttons; and 'Asset Type' with checkboxes for 'All', 'Column (71)', 'Table (42)', 'Mapping (13)', 'Delimited File (10)', and 'Folder (8)', plus 'Add' and 'Show All' buttons. The main area on the right is titled 'Search Results (1 - 20 of 193)' and lists three items: 1. 'customers_clean' (★★★★★), Data Owner: Shaun Kim, Data Steward: Kaiyu Lee, with tags 'Address_Information', 'Person', 'Customer', Type: Table. 2. 'customer_master' (★★★★★), Data Owner: Shaun Kim, Data Steward: Edwyn Lee, with tags 'Person', 'Address_Information', 'Customer', Type: Table. 3. 'Customer' (Composite Data Domain), Type: Composite Data Domain. Below these are two 'Customer' terms: 'The party that receives, buys, or consumes an item or service.' (Type: Term) and 'a party, either person or company, who purchases a product/service from one of our business units' (Type: Term).

Enterprise Data Catalog

Powered by **CLAIRE™**

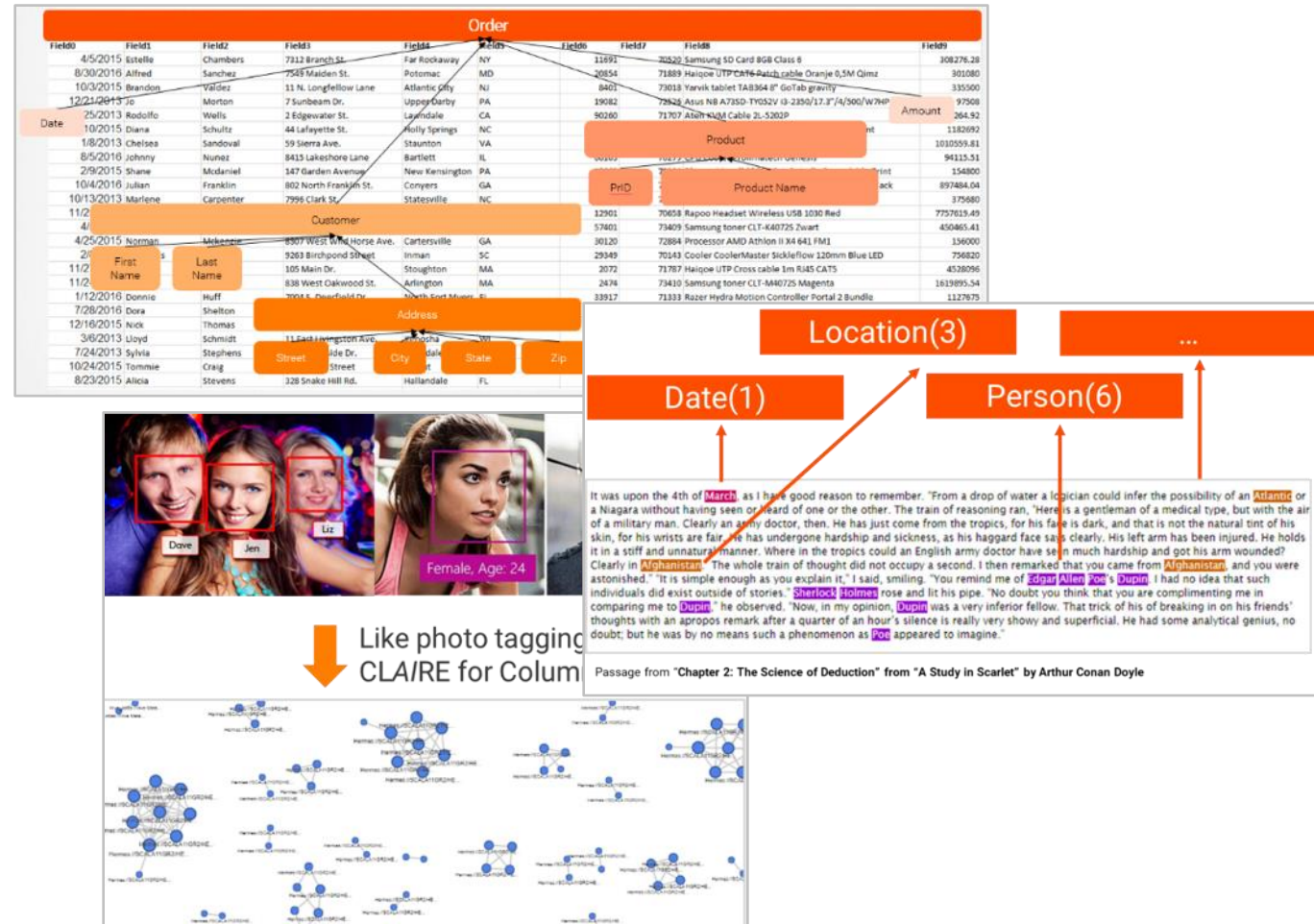
- 직관적인 **검색화면**과 Semantic Search 나 비즈니스 용어를 이용한 검색
- Data Asset Overview** 와 **데이터 프로파일링** 기능. 그리고 **End-to-End Lineage**를 통한 데이터 영향도 및 관계 파악
- AI와 머신러닝**을 이용해 자동화한 데이터 분류와 비즈니스 글로서리 태깅
- 엔터프라이즈 시스템을 위한 방대한 **메타데이터 추출 커넥터**
- 사용자들의 **참여과 협업**을 통한 데이터 정보 확장과 큐레이션
- Open Framework**을 통해 다른 Application과 쉬운 통합과 고객 개발 App과의 쉬운 통합을 위한 제공되는 **방대한 API**



Enterprise Data Catalog

Powered by **CLAIRE™**

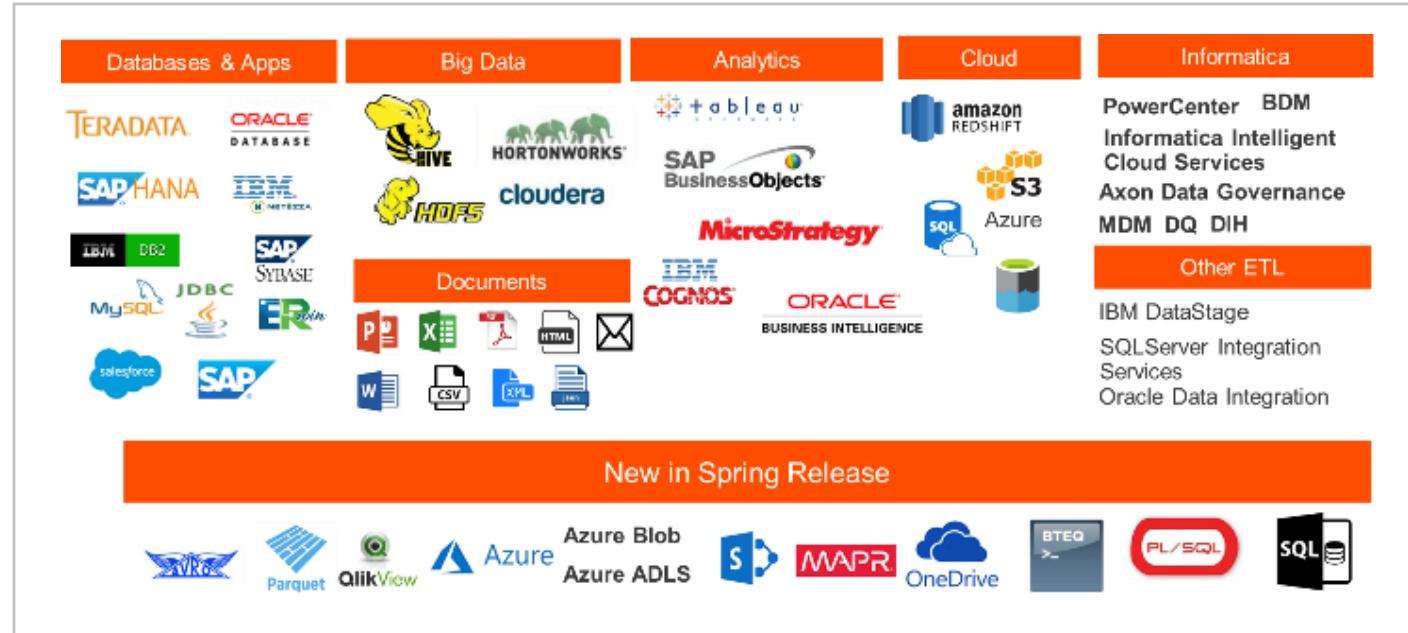
- 직관적인 **검색** 화면과 Semantic Search 나 비즈니스 용어를 이용한 검색
- Data Asset Overview** 와 **데이터 프로파일링** 기능. 그리고 **End-to-End Lineage**를 통한 데이터 영향도 및 관계 파악
- AI와 머신러닝**을 이용해 자동화한 데이터 분류와 비즈니스 글로서리 태깅
- 엔터프라이즈 시스템을 위한 방대한 **메타데이터 추출 커넥터**
- 사용자들의 **참여와 협업**을 통한 데이터 정보 확장과 큐레이션
- Open Framework**을 통해 다른 Application과 쉬운 통합과 고객 개발 App과의 쉬운 통합을 위한 제공되는 **방대한 API**



Enterprise Data Catalog

Powered by **CLAIRE™**

- 직관적인 **검색** 화면과 Semantic Search 나 비지니스 용어를 이용한 검색
- **Data Asset Overview** 와 **데이터 프로파일링** 기능. 그리고 **End-to-End Lineage**를 통한 데이터 영향도 및 관계 파악
- **AI와 머신러닝**을 이용해 자동화한 데이터 분류와 비지니스 글로서리 태깅
- 엔터프라이즈 시스템을 위한 방대한 **메타데이터 추출 커넥터**
- 사용자들의 **참여와 협업**을 통한 데이터 정보 확장과 큐레이션
- **Open Framework**을 통해 다른 Application과 쉬운 통합과 고객 개발 App과의 쉬운 통합을 위한 제공되는 **방대한 API**



Enterprise Data Catalog

Powered by **CLAIRE™**

- 직관적인 **검색** 화면과 Semantic Search 나 비즈니스 용어를 이용한 검색
- Data Asset Overview** 와 **데이터 프로파일링** 기능. 그리고 **End-to-End Lineage**를 통한 데이터 영향도 및 관계 파악
- AI와 머신러닝**을 이용해 자동화한 데이터 분류와 비즈니스 글로서리 태깅
- 엔터프라이즈 시스템을 위한 방대한 **메타데이터 추출 커넥터**
- 사용자들의 **참여과 협업**을 통한 데이터 정보 확장과 큐레이션
- Open Framework**을 통해 다른 Application과 쉬운 통합과 고객 개발 App과의 쉬운 통합을 위한 제공되는 **방대한 API**

The screenshot displays the Informatica Enterprise Data Catalog interface. At the top, there's a navigation bar with tabs for Overview, Columns, Lineage and Impact, Relationships, Reviews, and Questions. The main area is titled 'Assets Followed' and shows a grid of data assets. Each asset card includes a title, a star rating, the Business Owner, the Data Steward, and a brief description. Assets shown include 'Flight Delay', 'Flight Departures', 'Counter Sales', 'Airlines sales', 'Ticket Sales', and 'Flight Traffic Data'. Below the asset grid, there's a 'Notifications' section. It features a table with columns for Asset Name, Path, Description, and Modified. The first notification is for the 'CONTACT' asset, detailing a system update of the profiled row count from 129880 to 129879. The second notification is for the 'CUSTOMER_ID' asset, detailing a system update of the inferred data type value in CSV format from Decimal(6),129880,100.00 to String(6),129879,100.00. At the bottom of the notification table, there are buttons for 'Helpful (0)' and 'Delete'.

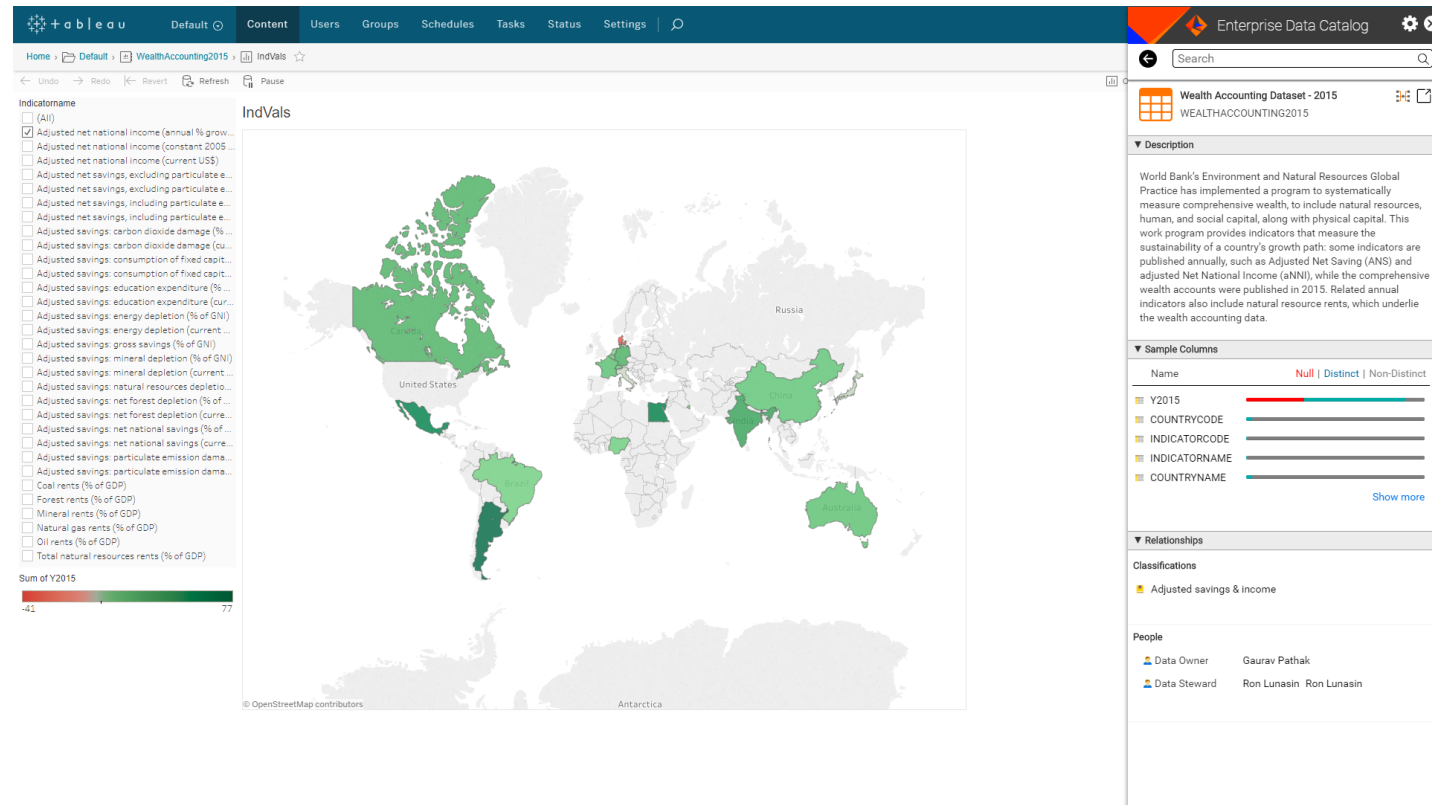
Asset Name	Path	Description	Modified
CONTACT	CRMS/ORCL/C##CRM	System updated value of Profiled Row Count from 129880 to 129879 in asset CONTACT. System updated value of Rows covered by profiling from 129880 sample rows. to 129879 sample rows. in asset CONTACT.	2 months ago
CUSTOMER_ID	CRMS/ORCL/C##CRM/CONTACT	System updated value of inferred DataType value in CSV format from Decimal(6),129880,100.00 to String(6),129879,100.00 in asset CUSTOMER_ID. System updated value of inferred DataType value in CSV format from String(6),129880,100.00 to Decimal(6),129879,100.00 in asset CUSTOMER_ID. System updated value of all pattern value in CSV format from Others,999,77 to Others,998,77 in asset CUSTOMER_ID. System updated value of inferred DataType value in CSV format from Integer(6),129880,100.00 to Integer(6),129879,100.00 in asset CUSTOMER_ID. System updated value of all pattern value in CSV format from 99999,90000,69.29 to 99999,90000,69.30 in asset CUSTOMER_ID. System updated value of Unique Count from 129880 to 129879 in asset CUSTOMER_ID.	2 months ago

Enterprise Data Catalog

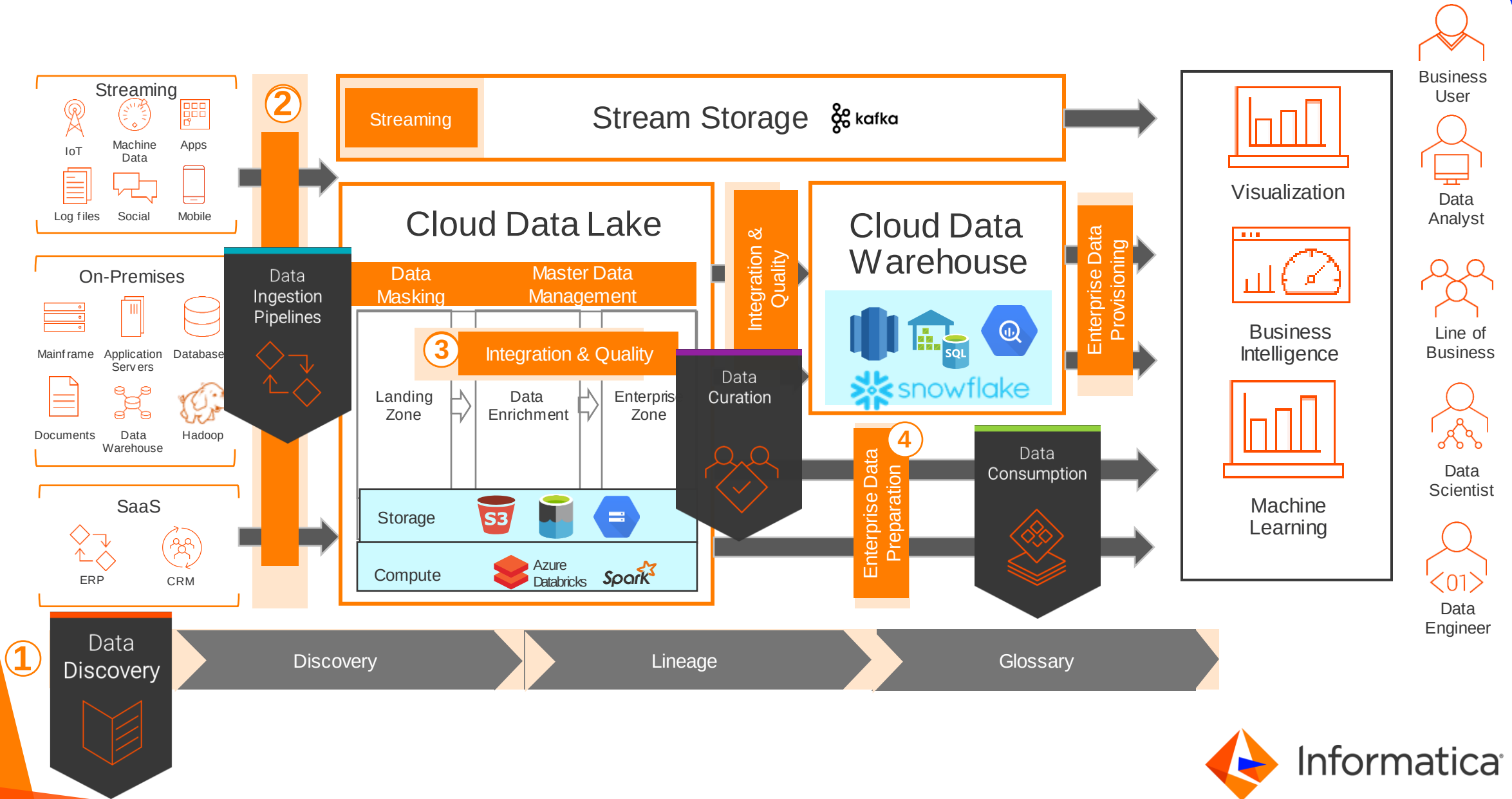
Powered by **CLAIRE™**

- 직관적인 **검색** 화면과 Semantic Search 나 비즈니스 용어를 이용한 검색
- Data Asset Overview** 와 **데이터 프로파일링** 기능. 그리고 **End-to-End Lineage**를 통한 데이터 영향도 및 관계 파악
- AI와 머신러닝**을 이용해 자동화한 데이터 분류와 비즈니스 글로서리 태깅
- 엔터프라이즈 시스템을 위한 방대한 **메타데이터 추출 커넥터**
- 사용자들의 **참여과 협업**을 통한 데이터 정보 확장과 큐레이션
- Open Framework**을 통해 다른 Application과 쉬운 통합과 고객 개발 App과의 쉬운 통합을 위한 제공되는 **방대한 API**

Tableau Plug-in



클라우드 데이터 분석 레퍼런스 아키텍처



고객 사례



Building a Complete Data Architecture

- Avis, Budget 및 Zipcar를 통해 업계에서 가장 인정받는 브랜드 3 개를 운영하는 모빌리티 솔루션의 선도적 인 글로벌 제공 업체
- 650,000 대의 대규모 차량을 실시간으로 완벽한 글로벌 뷰와 연결하여 효율성을 높이고 비용을 절감하고 수익을 극대화할 목적
- AWS에 Informatica 솔루션을 배포하여 데이터를 운영하고 차세대 플랫폼의 일부로 실시간 분석을 수행
- 엔드 투 엔드 데이터 파이프 라인으로 글로벌 차량 분석을 지원하여 차량 관리자가 실시간으로 차량을 추적 할 수있는 더 빠른 액세스를 제공



Connecting Great People to Great Companies

- 더 많은 사람들을 전 세계 직업기회를 찾아주는 리크루팅 에이전시
- 더 나은 고용 일치를 위해 고품질 데이터를 생성하고 레거시 데이터웨어 하우스를 Microsoft Azure SQL 데이터웨어 하우스로 마이그레이션
- Informatica Intelligent Cloud Services 및 Azure를 통해 현대적인 인재 매칭 플랫폼을 구성하고 Axon Data Governance 및 Enterprise Data Catalog를 통해 데이터 거버넌스 수행
- Informatica MDM으로 마스터 데이터 관리
- 통합된 조직, 고객, 직원 및 파트너 정보를 통해 보다 정확한 Job Matching 플랫폼 구축



Foundation Series

Cloud Lakehouse Data Management



Available in Korean! January 1, 2021

Earn a Cloud Lakehouse Data Management (Foundation Level) Certification badge when you complete the free online training.

<https://now.informatica.com/foundation-level-certification-series.html>

데모

