

# IBM Watson Knowledge Catalog

*An intelligent, integrated enterprise data catalog and data governance platform*



## Andres Avalos

**Offering Management**  
IBM Watson Data and AI

[aravalos@us.ibm.com](mailto:aravalos@us.ibm.com)



## Please Note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice and at IBM's sole discretion.

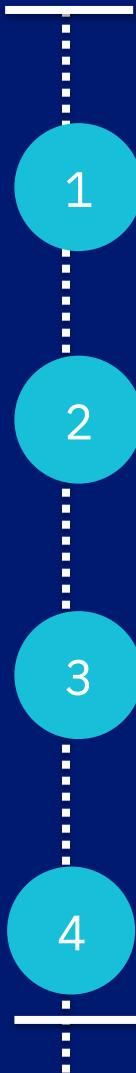
Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

# Agenda



**Todays Data challenges**

**Evolving Data Governance to drive value with  
Watson Knowledge Catalog**

**Snapshot**

**Q&A**

# It's Complicated

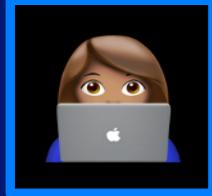
Messy & Unclassified; Hard to access & use



# Affected Personas

## Challenge

## Needs



### Knowledge Workers

Data Citizen  
Citizen Data Scientist  
Business Analyst  
Data Scientist

- Where can I find this data for my Customer Churn Analysis?
- What does this Product Data Attribute mean?
- I need to know what table(s) I need to query to write this extract of data.
- Can I trust this data? Where did it come from?



### Data Governance Teams

Head of Risk  
Head of Financial Reporting  
Chief Data Office

- How can I understand what impact this regulation will have on compliance?
- What does the 'Reference Salary' mean?
- How can I prove to auditors that these financial transactions are accurate?
- How can I enforce governance policies consistently across my enterprise data?
- We have huge potential within our data, how can I monetize it?



### Data Stewards

Data Quality Analysts  
Data Quality Stewards  
Data Custodian  
Data Curator

- How can I understand if this data is useful to the business?
- How can I tell if this data is accurate/trustworthy/useful?
- Who is the owner of this data, and how can I get them to update it?
- How can I improve the effectiveness of the Data Stewardship process?

- A fast way to find and understand what assets are valuable to them.
- Self service Data Prep and movement tools to gain fast access to discovered data.
- An enriched view of the information discovered and how it relates to the business terminology used.
- Trust and confidence that the data they use is accurate and of appropriate quality

- A mechanism to understand/manage corporate policies relating to industry regulations.
- Tools to easily find domain experts relating to data and policies
- Powerful tools to visualize the lineage of data as it flows across the enterprise.
- Automatic enforcement of governance policies as data is consumed
- Capabilities to leverage the governance program to drive better informed analytics and AI.

- An easy way to understand a technical and business view of the data and the relationships between them.
- Tools to manage the roll out of new business entities and their impact upon the data.
- Feedback mechanisms from the Knowledge Workers to improve the accuracy of my work.
- Tools to define, monitor and maintain the quality, trust and bias of data across the enterprise.

# The Changing Data Landscape

## Data Monetization Challenges

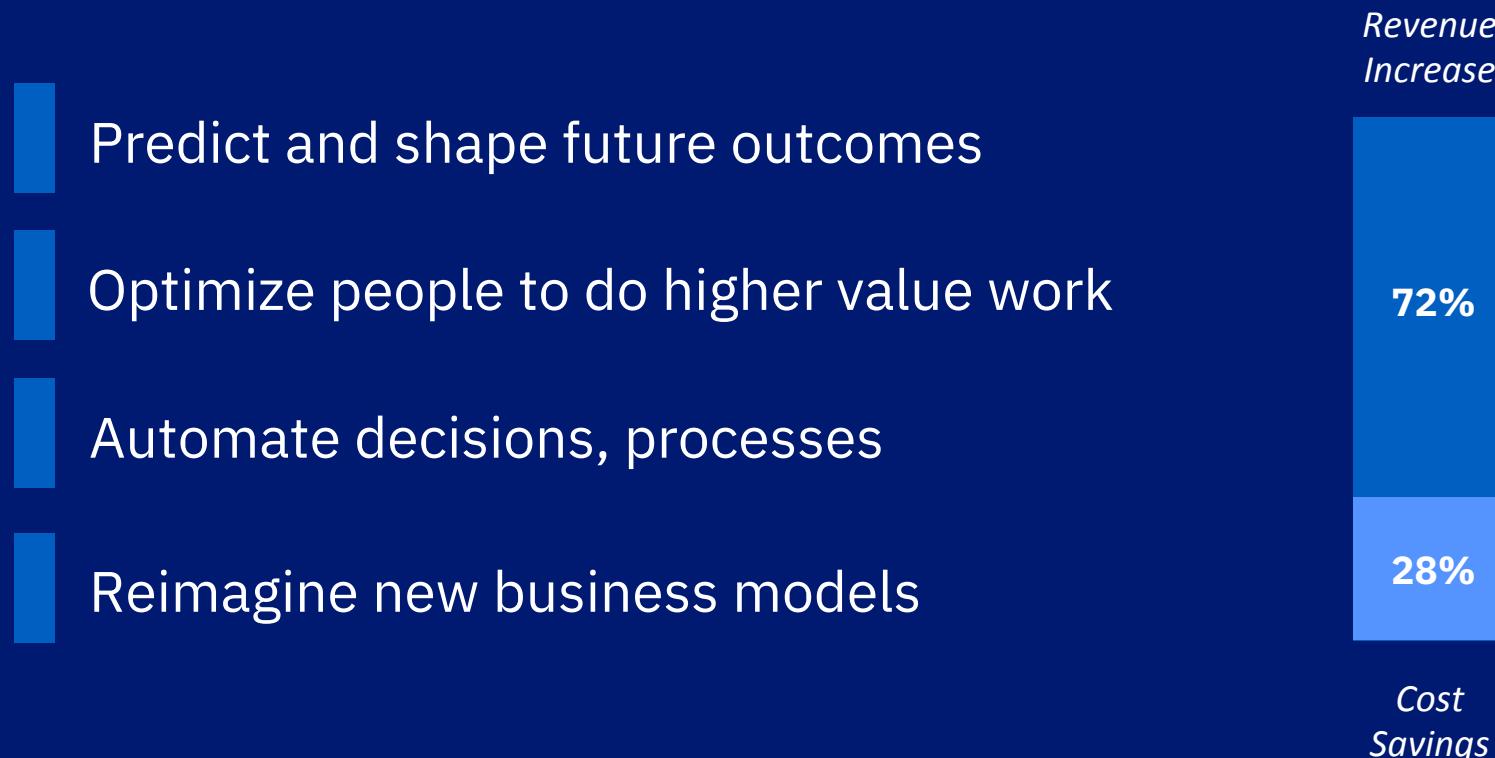
- **Speed** of data creation - **Structured & Unstructured & Semi-Structured Data**
- **Multiple** Data Location → Public Cloud, Private Cloud, On-Premises
- **Value Creation** from **Data** → High quality data for Analytics, ML & AI
- **Self Services access to all data** → Find and understand data
- **New Data Regulations:** GDPR, CCPA (CA), ARDR, etc.  
→ Auditability & Traceability & Data Provenance



# Put AI To Work

Surface hidden intelligence to transform how your business operates

**AI unlocks  
the value  
of data  
in totally  
new ways**



# There is no AI without an IA

(information architecture)

**“No amount of AI algorithmic sophistication will overcome a lack of data [architecture]**

**MITSloan**

# 81%

do not understand the data required for AI

# 80%

of data is either inaccessible, untrusted or unanalyzed

94%

are committed to multicloud

74%

using AI to modernize existing apps

87%

of AI developers use open source

# The AI Ladder

A prescriptive approach to accelerating the journey to AI



AI

**INFUSE** – Operationalize AI with trust and transparency

**ANALYZE** - Scale insights with AI everywhere

**ORGANIZE** - Create a trusted analytics foundation

**COLLECT** - Make data simple and accessible

Data of every type,  
regardless of where it lives



**MODERNIZE**  
your data estate for an  
AI and multicloud world

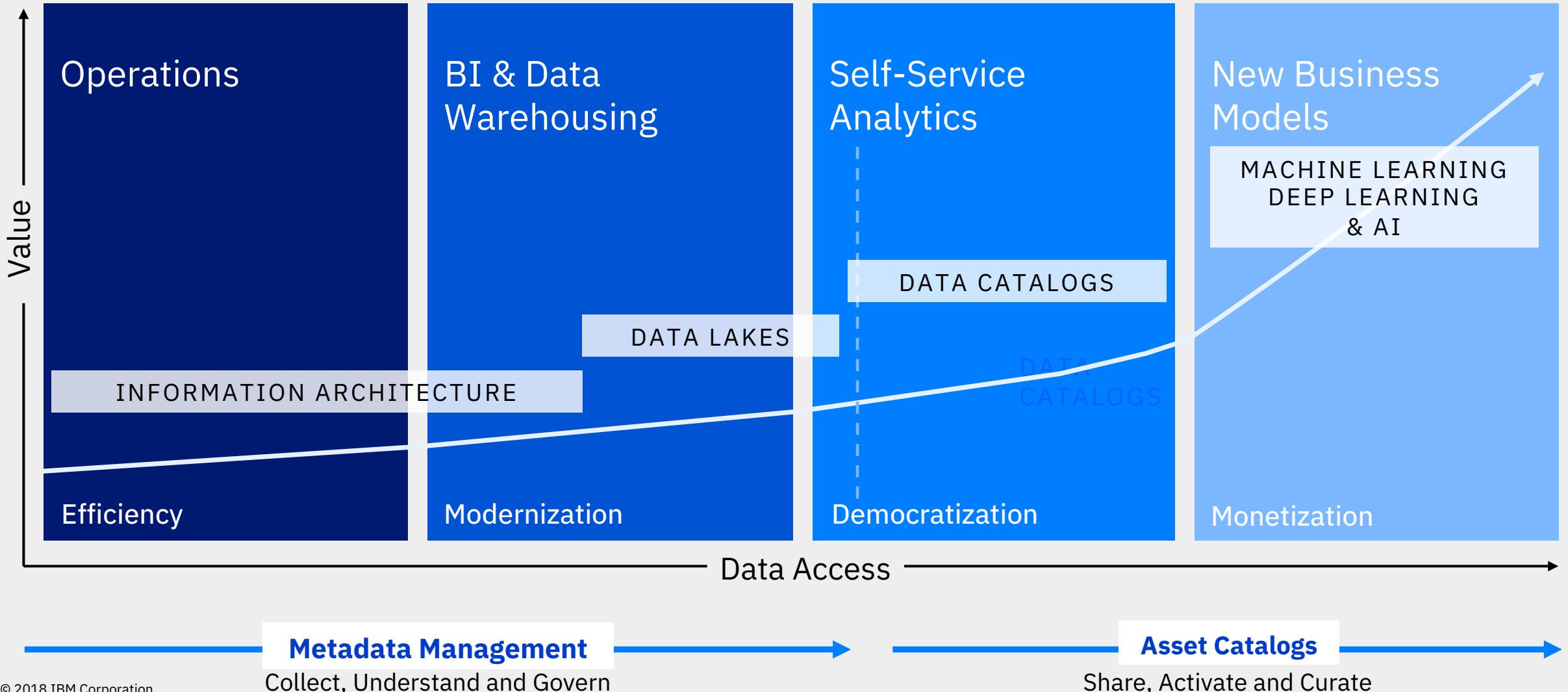
**But** organizations are hindered by data challenges.

80% of time is spent on collecting, cleaning, organizing data versus exploring and operationalizing insights to drive innovations everywhere.



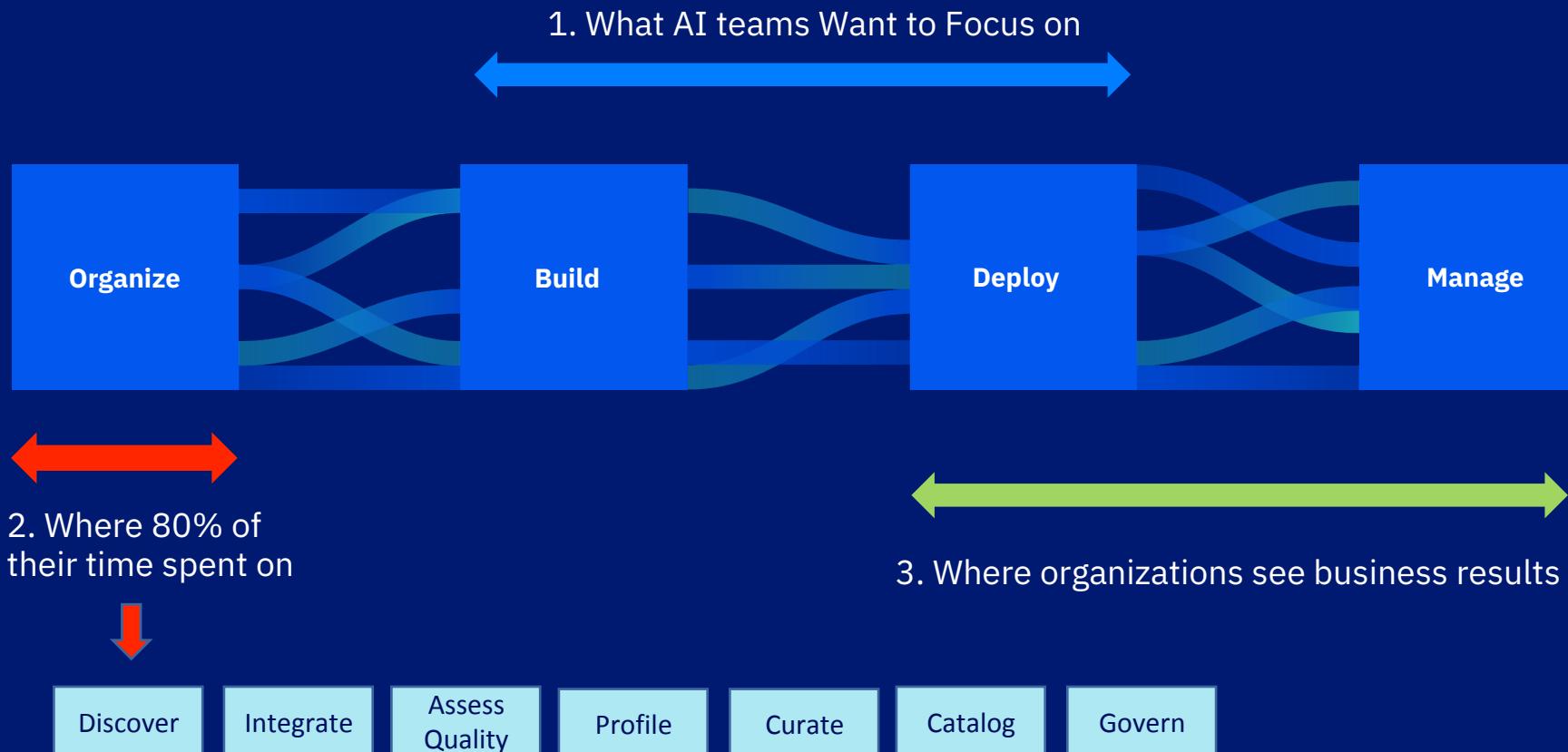
# Organizations are racing to unleash the power of data and apply AI & Analytics

IBM Data & AI provides the path to monetize your data



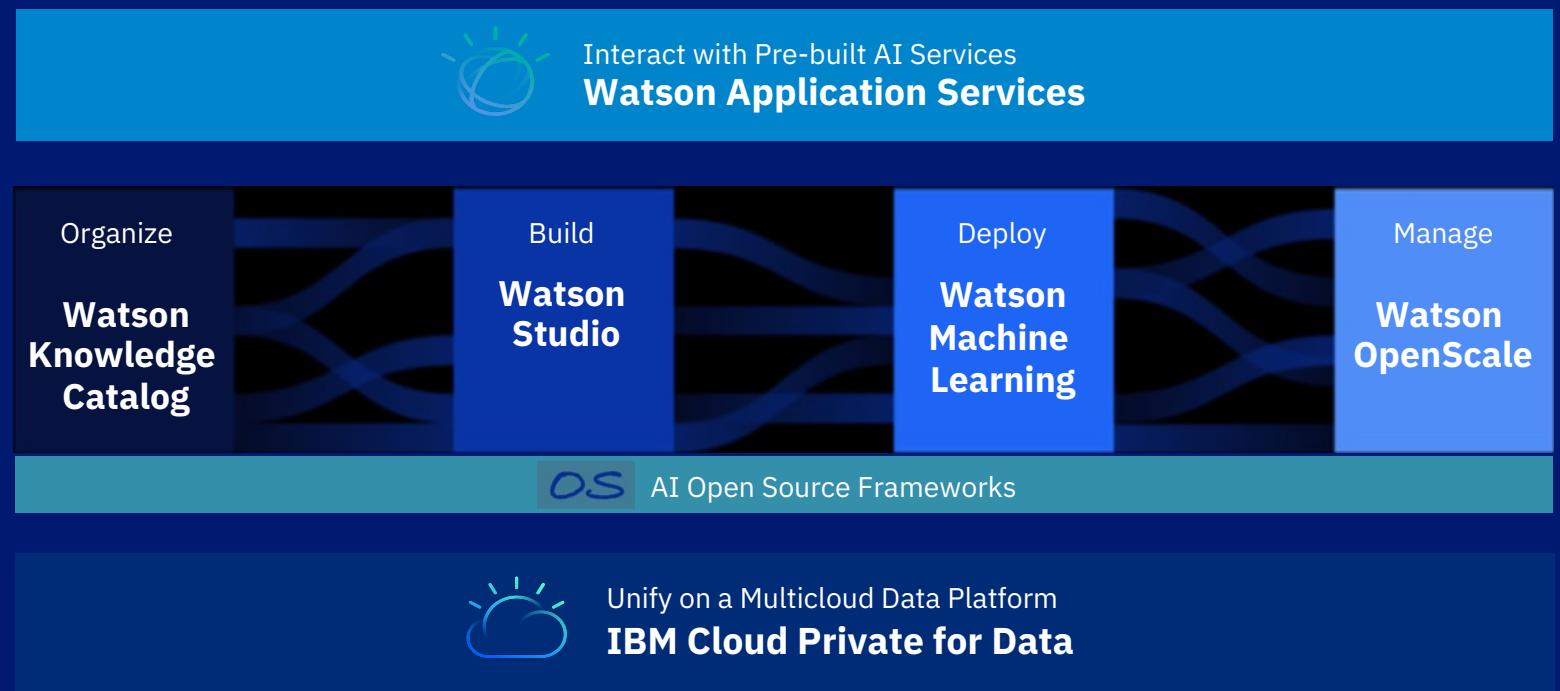
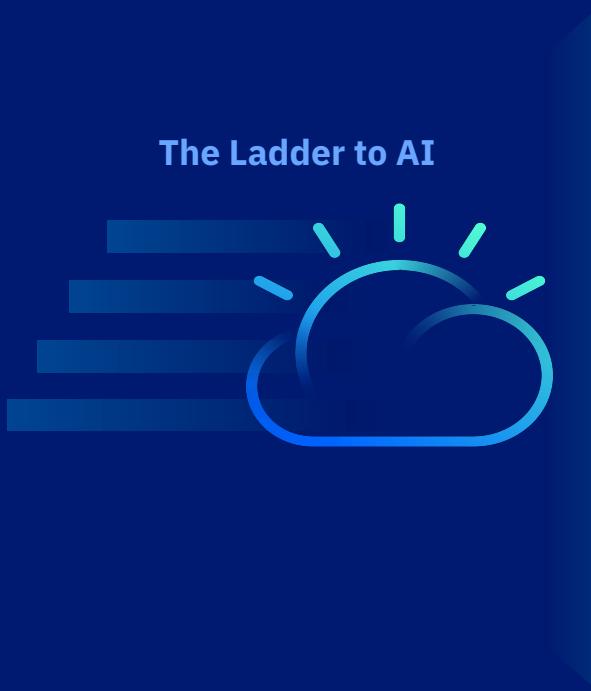
# 4 critical steps to operationalizing AI

## Business-Ready data



# IBM's AI Portfolio

Everything you need for Enterprise AI, on any cloud

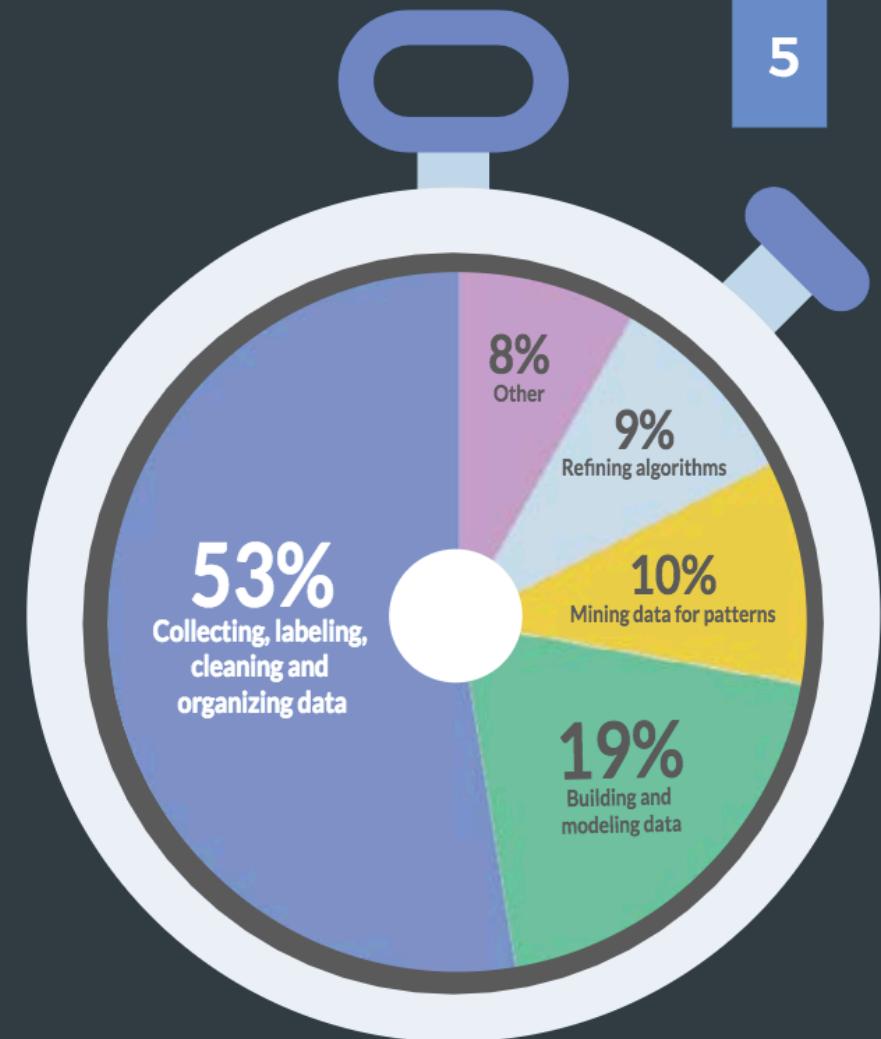


# The need for productivity in data science

## WHAT KEEPS DATA SCIENTISTS HAPPY? (and why aren't they doing more of it?)

**What activity takes up most of your time?**

5



# Introducing IBM Watson Knowledge Catalog



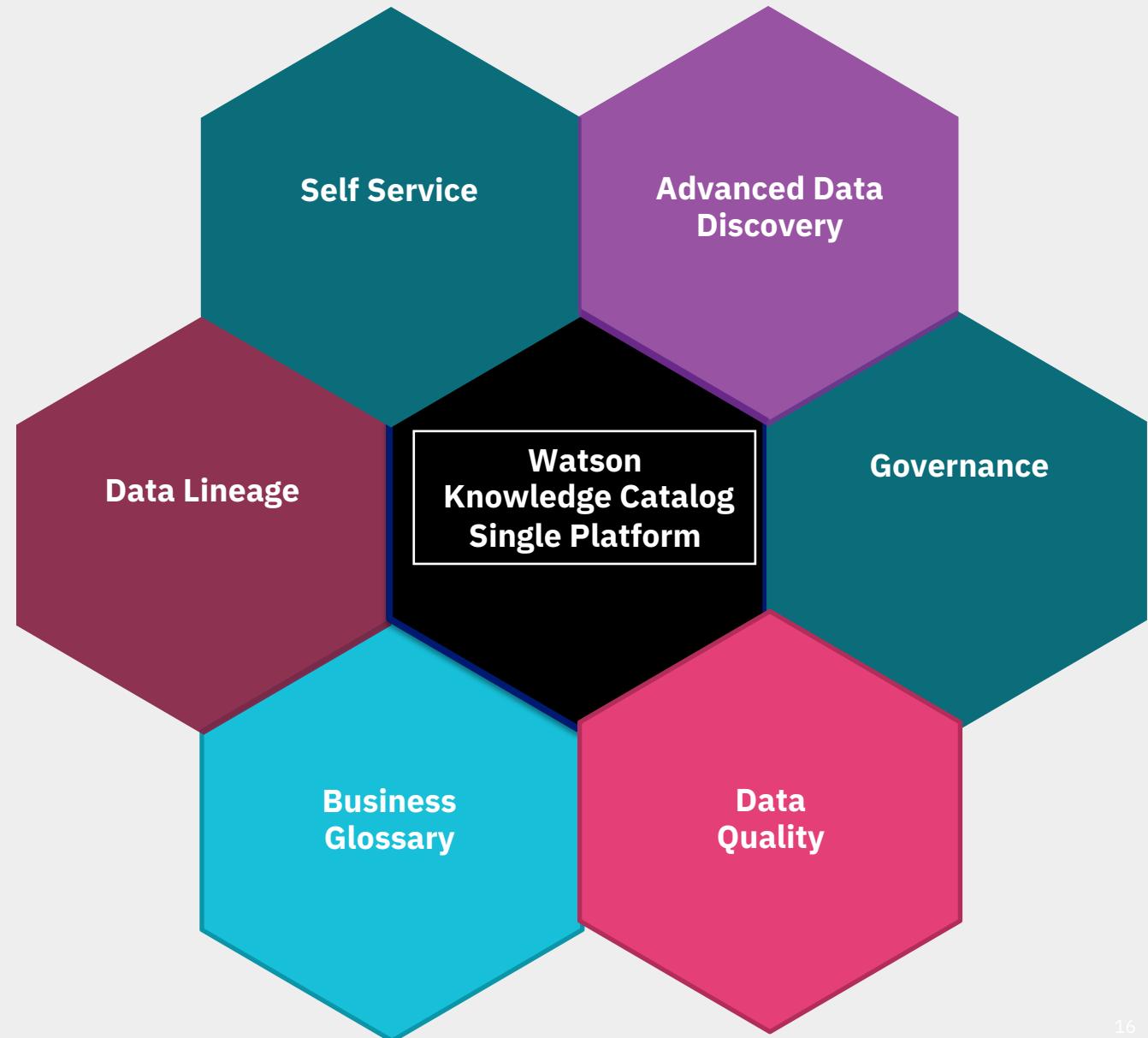
An intelligent enterprise data catalog powering self-service activation and discovery of data, AI models while governing them

- Interactively discover, cleanse and prepare your data, understanding data quality and distribution.
- Intelligent recommendations for knowledge workers for discovering relevant assets quicker across the organization through self-service access.
- Integrated data governance platform for Data Governance teams to define business glossary, policies and rules to govern and protect data and track data lineage and foster responsible stewardship of data
- Protect data & AI model misuse and confidently share assets based on organizational policies

# IBM Watson Knowledge Catalog Capabilities

An intelligent way to make data **accessible** and **valuable** to anyone who needs it.

- Bring the value of your governed data to the Data Citizen's
- Expose many types of assets for analytics and data science.  
Structured – Unstructured – Models – Notebooks - Dashboards
- Track the lineage of your assets throughout the enterprise, right up to consumption
- Open Metadata Exchange: Connect all of your metadata repositories using Open Metadata Repository Services (OMRS)
- Active policy engine providing trust and control over data
- Hybrid and Multi-cloud: Purchase on IBM Cloud, on-prem, or through other Cloud vendors (coming in 2019)



# Key Catalog Capabilities

## Seamless integration for Productive Use

- Watson Knowledge Catalog is fully integrated with Watson Studio meaning that once assets have been discovered users can easily drive productive use through integrated tools for data shaping, data movement, data science, AI and machine learning.

## Focus on Enabling AI & Analytics

- IBM's catalog supports all assets... not just data. Think dashboards, data science & machine learning models, connections, notebooks, etc. – All available in one easy to find self service experience.

## Based upon Open source

- Allows integration through opens standard to harvest, share and consume metadata from other open metadata environments. - Don't be locked in with proprietary metadata
- Allows the metadata to be consumed within Watson Studio using the open source tool of choice by developers.

## Modern Policy Activation

- IBM's Catalog contains an advanced policy enforcement engine to establish categories, policies & rules on data assets.
- Activation of these policies means data is masked or protected on the fly at the point of access.

## Structured & Unstructured

- IBM's catalog uses Watson Natural Language processing to extract and understand the value locked away in unstructured and structured data and presents it uniformly for consumption. Provides insights from unstructured data directly to knowledge workers.

## Your data, Your way

- Providing choice & flexibility for data location gives the opportunity for highly regulated industries to benefit from IBM Cloud and Watson.
- Connect to data sources anywhere, not just in Hadoop.

## Discover what you never knew you had

- Powerful discovery tools will unlock and provide context to data you didn't know you had or understand its value.
- Intelligent algorithms will detect how data is related, who is using it and where it may be duplicated.

## Using AI to power your AI

- Watson Recommends – Our AI powered data recommendation engine unlocks data for insights.
- Our ML based classification engine will automatically decide which data classes are discovered within your data and self categorize it for you.

## Fit for purpose experiences

- Our self service portal provides unique capabilities to ensure the knowledge worker can quickly find the right asset, quickly.
- Rich enterprise governance and stewardship interfaces ensure that the IT team can easily manage their data landscape.

# IBM Has Invested Several Billion Dollars Over Years To Build One Integrated Platform For Delivering Business Ready Data To Knowledge Workers

## Business Ready Data Supply Chain

### Data Integration

#### **Enterprise Data Integration**

IBM provides the most comprehensive, most integrated, and most scalable data integration platform provides the core data integration backbone running in the largest banks, telcos, retailers, insurance companies, etc..

### Data Quality

#### **Enterprise Data Quality**

IBM provides the most complete, most scalable, and most integrated data quality platform supporting data profiling; creation, execution, and monitoring of data validation rules; and data matching and consolidation.

### Data Governance

#### **Enterprise Data Governance**

IBM provides an enterprise data governance platform that supports business and technical users on data governance teams. It eliminates the cost and complexity of integrating a stand-alone data governance platforms with data integration, quality, and consumption tooling.

### Data Consumption

#### **Enterprise Data Consumption**

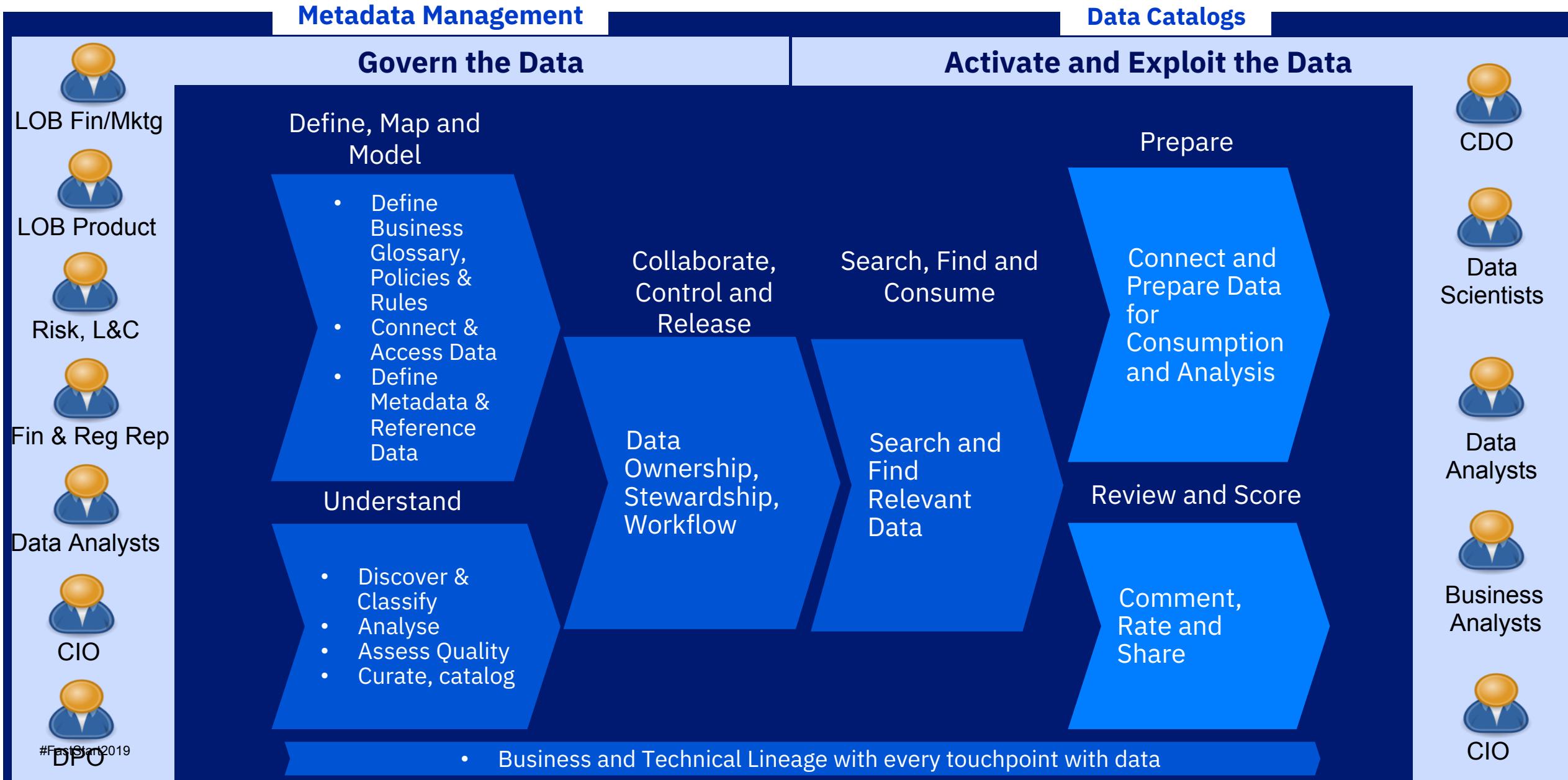
IBM provides an enterprise data catalog that delivers self-service capabilities for Knowledge Workers (business analysts, data scientists) to search and explore information, to preview and refine information, to act upon information in a secure environment driving new Data Science or Analytics.

### **Whole Is Greater Than The Sum Of The Parts**

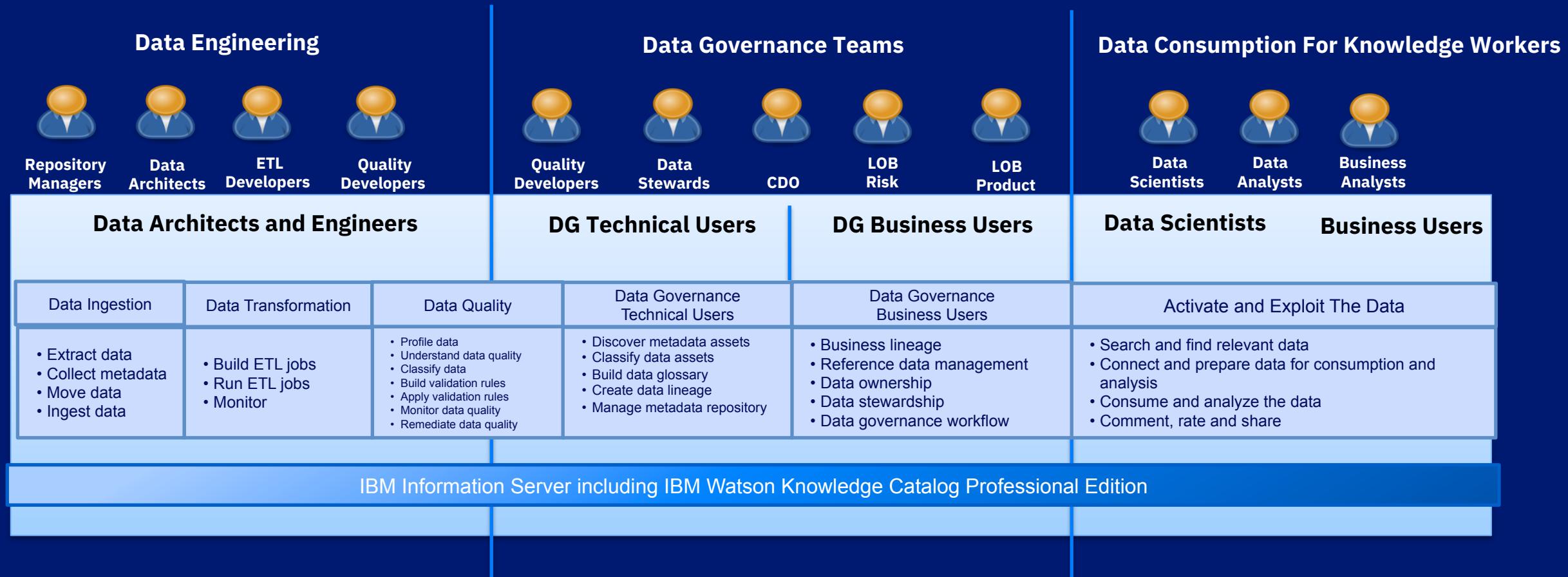
By providing the most complete platform for Enterprise Data Integration, Quality, Governance, and Consumption, IBM delivers Business Ready Data to Knowledge Workers without the manual, expensive, and impossible task of integrating stand-alone data integration, data quality, data governance, and data consumption solutions.

**End-to-End. Unique. Fully Integrated.**

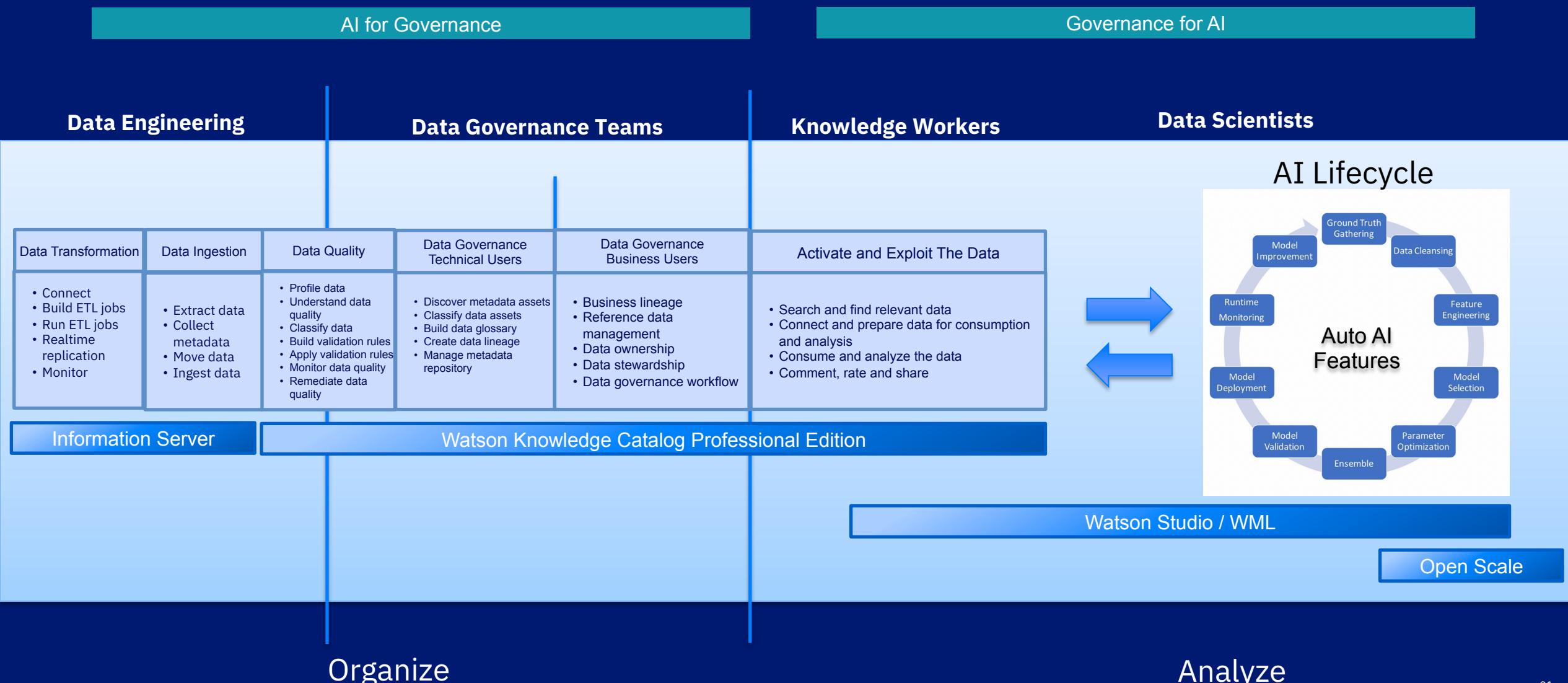
# Supporting the journey from collection to consumption



# IBM is providing the most complete platform for delivering Business Ready Data to Knowledge Workers



# The Most Complete Platform For Enterprise Data Integration, Quality, Governance, And Consumption – Supporting Data Science & Analytics



# Policy Activation – Key to delivering data monetization

## **Policy Activation**

Activate your governance program through Machine Learning based capabilities to detect, enforce and remediate your policies against your data.

## **Detection**

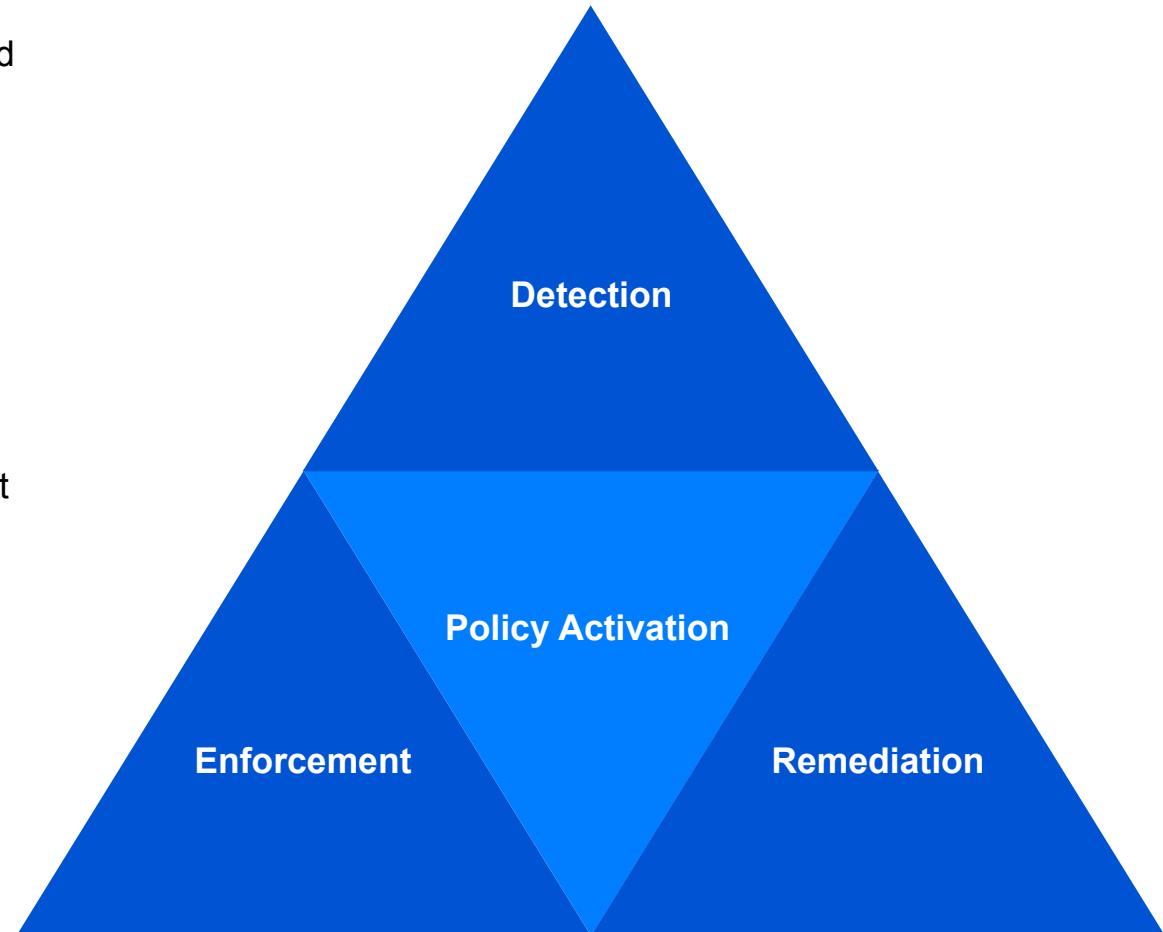
Detect what your data is, and where it is. Understand the key entities within the data and who is trying to use it and for what purpose.

## **Enforcement**

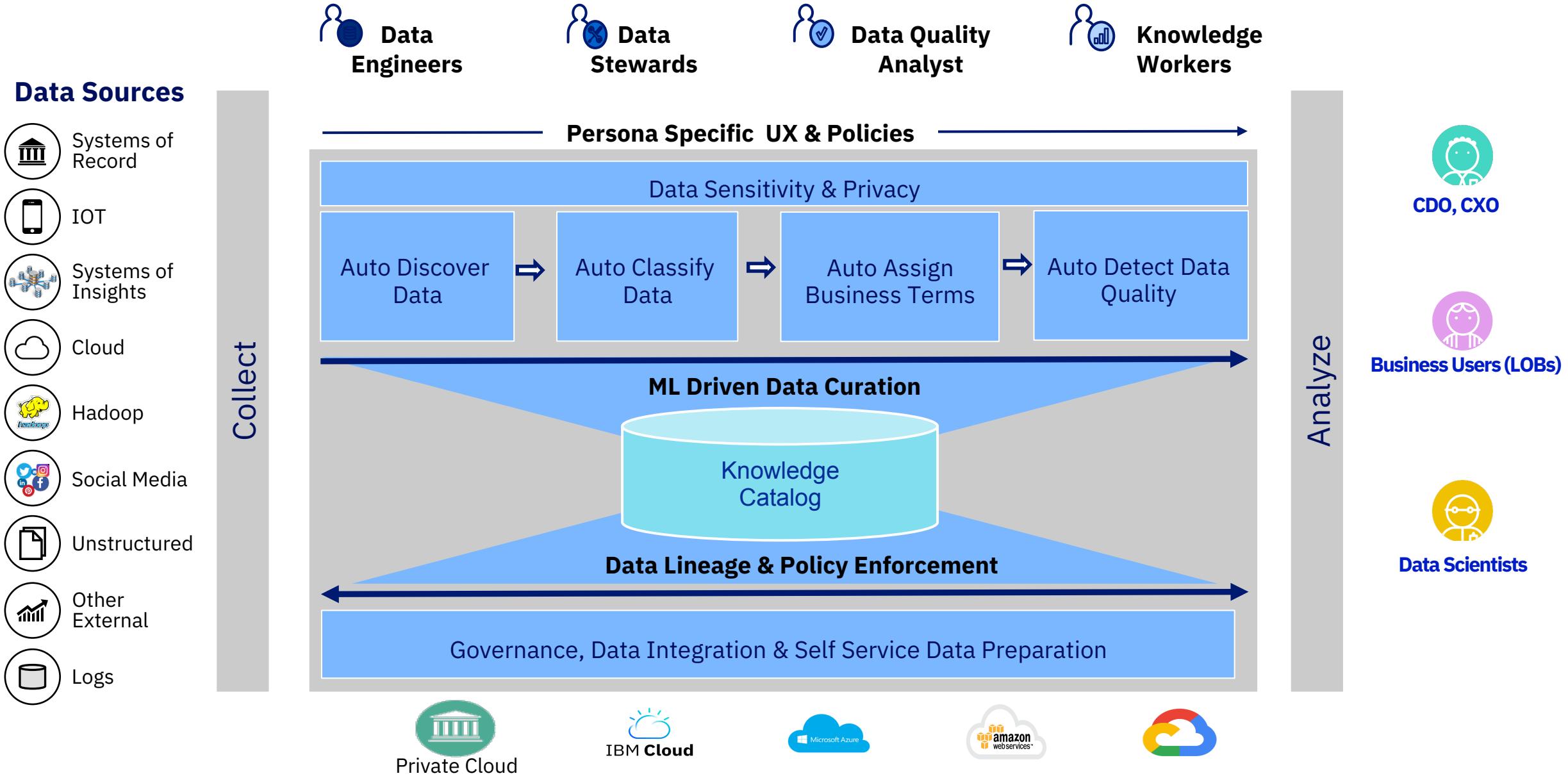
Enforce the rules defined by the governance program to ensure that data is masked or users are denied access at the point of use.

## **Remediation**

Allow issues with the data to be flagged as incorrect or requiring changes. Integrated workflow capabilities ensure the issue is flagged to the right individual to be corrected. All them time tracking what was remediated by who.



# IBM Data Curation Strategy



# IBM Unified Governance & Integration

## Data Sources

Systems of Record

IOT

Systems of Insights

Cloud

Hadoop

Social Media

Unstructured

Other External

Logs

Collect

## Organize

### Core Integration Services

Enterprise ETL

Multi-Cloud Data Integration \*

Data Replication

### Automated Data Curation Services

Auto Discover & Classify Data

Auto Detect Sensitive Data

Auto Analyze Data Quality

Auto Assign Business Terms

### Metadata Management & Data Catalog

Knowledge Catalog

On prem  
On Public Cloud  
On ICP4D

### Self-Services Interaction

Search & Find Relevant Data

Tagging, Comments Annotations

Share & Collaborate

Self-Service Data Preparation

### Core Governance Services

Policy Management & Enforcement

Model Governance & Bias Reporting \*

Business Glossary Management

Data Quality Management

Data Lineage

Data Archival & Disposal

Consent Management

Entity Management & Resolution

Analyze

## Users



CDO



Governance Officers



Data Quality Analyst



Data Steward



Data Scientist



Business Users



Data Engineer



Private Cloud



IBM Cloud



Microsoft Azure



amazon web services\*



\* Future roadmap

# A modern architecture for today and the future

*Build upon native cloud value leveraging open frameworks and multicloud*

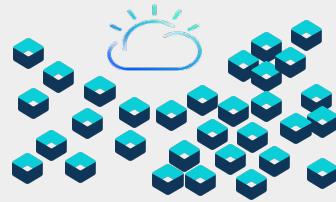
## Build for agility and scalability

### Microservices



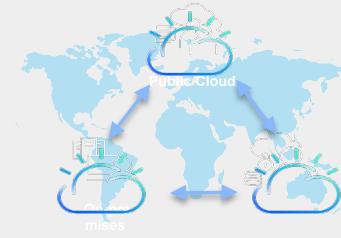
An architecture of loosely coupled data services, easily refactored to create containerized workloads

### Containerized workloads



Stand-alone workloads composed of micro-services and data that are flexibly deployed, orchestrated and managed

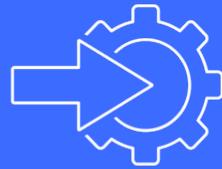
### Multicloud provisioning



Agile provisioning of containerized workloads in multi-Cloud environments and consumption of Cloud services



# Key Benefits



## Augmented Data Management

Automate the collection, curation and understanding of your data. Detect business entities, quality, bias and validity of your data to build an automatic level trust in your data.



## AI Powered Data Prep & Findability

Utilize automatic suggestions on how data can be best prepared for use. Intelligent suggestions to guide users to the best data for faster, smarter analytics and AI.



## Policy Activation for Data Governance

Autonomous enforcement of data and AI governance policies, providing automatic decisions to mask and protect data for a particular purpose.



## AI Governance

Support your adoption and roll out of trusted AI. Ensure that your models remain compliant for quality and bias.

*Automate the mundane*

*Use the data to get the data*

*Activate your governance program*

*Govern AI across your enterprise*

# Data Curation is critical to Knowing your Data, so that you can monetize it ... at scale

Messy & Unclassified; Hard to access & use



Discover  
Ingest  
Classify  
Organize  
Integrate



Automate with ML

Organized; Ready for use



Knowledge Catalog  
(Insights about your  
information Assets)

# Data Consumption is the final step to drive value from your governance initiative

Messy & Unclassified;  
Hard to access & use



Discover  
Ingest  
Classify Organize  
Integrate

Automate with ML

Organized; Ready for use



Share  
Find  
Curate  
Consume

Automate with ML

Consumed; Delivering Value



## Whole Is Greater Than The Sum Of The Parts

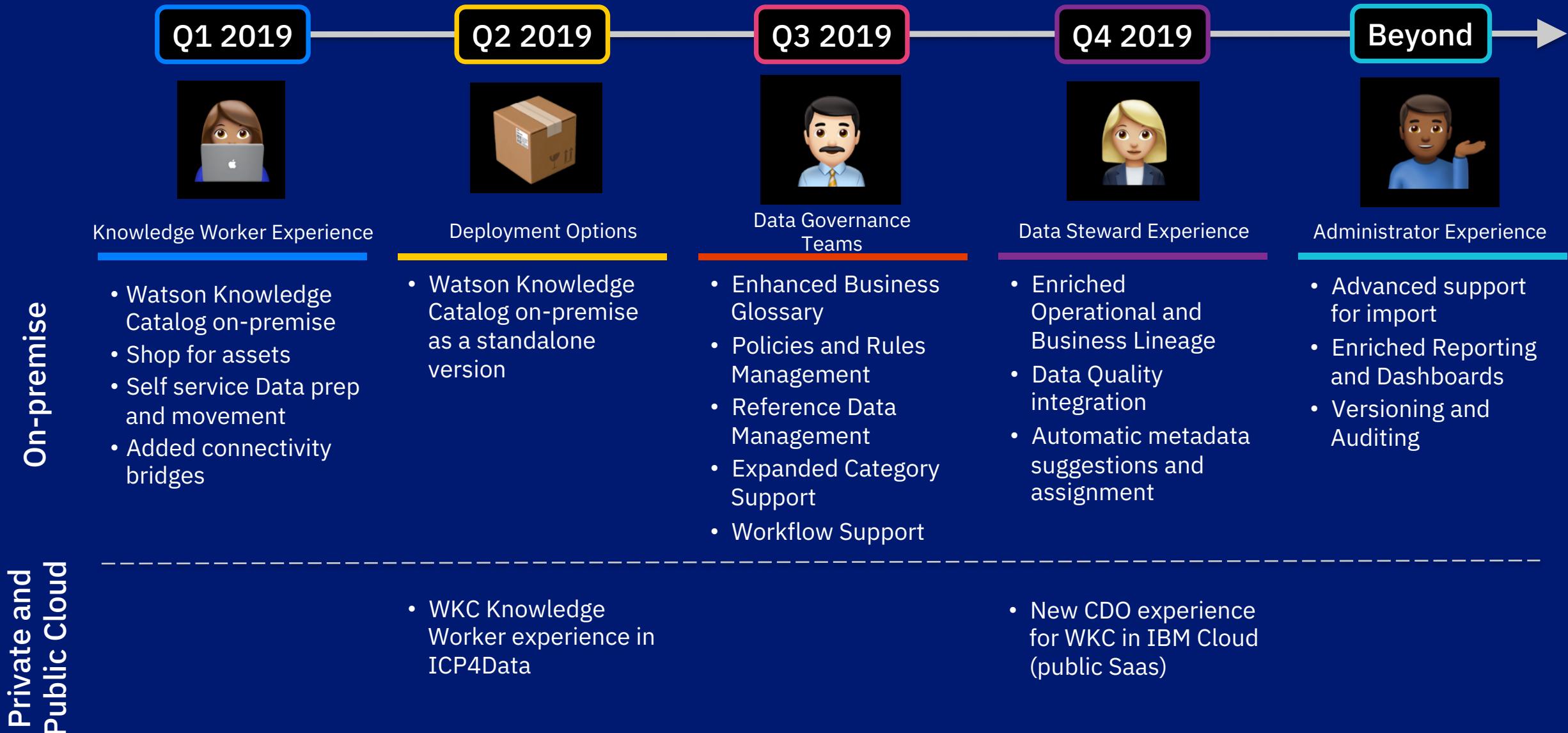
By providing the most complete platform for Enterprise Data Integration, Quality, Governance, and Consumption, IBM delivers Business Ready Data to Knowledge Workers without the manual, expensive, and impossible task of integrating stand-alone data integration, data quality, data governance, and data consumption solutions.

Unleash the value of your data governance initiative by empowering Data Citizens to deliver value from data Infused with Machine Learning to automate and simplify the understanding of data

# Quick View: Persona Based Roadmap

# Watson Knowledge Catalog

2019 Roadmap and Strategic Vision



# Summary:

- IBM provides the most comprehensive Data Governance platform
- Watson Knowledge Catalog expands this further with self service to support data monetization
  - The market leading data catalog
- Your competitive advantage is your data, put it in the hands of the people that can extract the value.
- The need to embrace, AI is real. Your competitors are doing it and already baking into business processes.
- Ensure your chosen platform supports governance across the entire data & AI lifecycle.

# 5 Things you can do now to know more

1. Visit: <https://www.ibm.com/cloud/watson-knowledge-catalog>
2. Watch this video: <https://youtu.be/1pxpLpKyLJI>
3. Read this blog:  
<https://www.ibmbigdatahub.com/blog/data-catalogs-of-the-future-within-an-automated-and-intelligent-world>
4. Download the Forrester Machine Learning Data Catalogs Report :  
<https://www.ibm.com/account/reg/us-en/signup?formid=urx-35284>
5. Try for free:  
[https://dataplatform.cloud.ibm.com/registration/stepone?apps=data\\_catalog&context=wdp](https://dataplatform.cloud.ibm.com/registration/stepone?apps=data_catalog&context=wdp)

# Snapshot

**Watson Knowledge Catalog: Search and Explore**

Catalogs / Enterprise

Add to Catalog

Browse Assets Access Control Settings

What assets are you looking for?

Watson Recommends Highly Rated Recently Added Expand

Filter Asset types Tags

Available Assets

Showing 25 of 53 assets

Catalogs / Enterprise Add to Catalog

Browse Assets Access Control Settings

What assets are you looking for?

Watson Recommends Highly Rated Recently Added Collapse

Data Asset (39) Connection (11) Notebook (2) Model (1)

Insurance (17) Great-Outdoors (13)

NAME OWNER

2017 J.D. Power U.S. Auto Clai... (User icon)

Auto Insurance Claim (User icon)

Auto Insurance Claims (User icon)

Auto Insurance Customers (User icon)

Auto Insurance Policies (User icon)

1 0 1 Data Asset Insurance Driver

Owner: Dirk deRoos Added: Sep 12, 2018 9:32 AM

Insura... | Demo

5 stars 0 reviews

1 0 1 Data Asset Great Outdoors Retailer Sales Recent

Owner: Ricardo Buglio Added: Feb 20, 2018 2:44 PM

Great... | Sales | Retail

5 stars 0 reviews

1 0 1 Data Asset Princeton Video Customer Sales

Owner: Ricardo Buglio Added: Oct 31, 2017 9:22 AM

Prince... | Sales

5 stars 0 reviews

Filter Available Assets

# Watson Knowledge Catalog: Review and Understand

Catalogs / Enterprise / CUSTOMER\_DEMOGRAPHIC

Add to Catalog

Overview Access Review Profile Lineage

Data Asset

## CUSTOMER\_DEMOGRAPHIC

Remove Download Add to Project

EMAIL_ADDRESS	PHONE_NUMBER	NATIONAL_ID	CREDITCARD_NUMBER	CREDITCARD_TYPE	CREDITCARD_EXP
Type: String	Type: String	Type: String	Type: String	Type: String	Type: String
Email Address	US Phone Nu...	US Social Se...	Credit Card Number	Organization Na...	Date
wfronsek1@source...	434-553-8337	231-24-6500	7df7ea4aac5e3493ea3127e	VISA	5/23
dcoyekj@pcworld.co...	760-277-6466	649-96-7557	db7d354d8a0d2042c969dd	Diners Club	3/21

Description  
Customer churn data asset

Added: Jan 10, 2019 11:09 AM  
Format: CSV  
Size: 322 KB

Tags  
Customer | Churn

Reviews  
 1 review

Catalogs / Enterprise / CUSTOMER\_DEMOGRAPHIC

Add to Catalog

Schema: 28 Columns | 2066 Rows | 2 Columns anonymized [i](#)

Preview: 1000 rows | Last refresh: 1 day ago | Refresh

STATE	STATE_CODE	ZIP_CODE	EMAIL_ADDRESS	PHONE_NUMBER	NATIONAL_ID	CREDITCARD_NUMBER	CREDITCARD_TYPE	CREDITCARD_EXP
Ohio	OH	44646	wfronsek1@source...	434-553-8337	231-24-6500	7df7ea4aac5e3493ea3127e	VISA	5/23
Washington	WA	98373	dcoyekj@pcworld.co...	760-277-6466	649-96-7557	db7d354d8a0d2042c969dd	Diners Club	3/21
California	CA	91604	kwaggatdu@typepx...	915-586-6081	962-90-5618	c6656b4696af2bcad045	American Express	8/19
New York	NY	11570	vcokeriy@shareasal...	817-205-8994	613-45-8368	2802c0eabcf2fec9d4bc9	American Express	12/23
Oregon	OR	97214	ppotegi@cnet.com	901-313-6753	873-61-4046	e3c9064b11ac1a8a446fd	Discover	4/23
Virginia	VA	23220	lmebs68@mozilla.o...	318-710-5442	950-96-4106	11776326fed27128f89ac	JCB	7/20
Colorado	CO	80021	cburcher1o@tuttoci...	214-501-0431	708-76-8249	81471b1239519e9b69c8	VISA	7/20
Ohio	OH	43147	tgeeves9w@nsw.go...	212-160-2937	805-98-0233	738f72ee7485cdcd47a3c	JCB	10/19
California	CA	94541	ovincentqy@weibo.i...	520-774-2490	061-31-7031	978312a843698d357512	JCB	12/22
Ohio	OH	43035	ptythertonix@phoc...	608-176-4288	353-89-2233	5af9d98a2efda868e10949	Diners Club	12/22

Catalogs / Enterprise / CUSTOMER\_DEMOGRAPHIC

Add to Catalog

Overview Access Review Profile Lineage

Current profile  
164 classifiers

Last profile  
10 Jan 2019 - 11:43 am [View Log](#)

Columns 28 Rows 2,066 Delete Update Profile

ID	CUSTOMER	NAME	COUNTRY
Type: Smallint	Type: Varchar	Type: Varchar	Type: Varchar
• I	• I	• T	• CC
100%	0%	0%	100%
1922 - 1954	CH15677	Marie Machen	Ohio
1491 - 1518	E271780	Bernadette Jeffrey	Washington
371 - 397	CB44654	Dalston Lamberton	California
1899 - 1925	BP28492	Nero Martell	New York
1222 - 1248	GE50763	Job Trevena	Oregon
529 - 555	QT25383	Dulcine Leap	Virginia
1411 - 1437	AU96286	Gerhardine Brockhouse	Colorado
1873 - 1898	JV89183	Tiphany Samter	Ohio
503 - 528	FH83537	Terrye Dunleavy	California
• US	• US	• US	• US

Schema: 28 Columns 2 Columns anonymized [i](#)

Preview: 1000 rows | Last refresh: 4 hours ago | Refresh

STATE	STATE_CODE	ZIP_CODE	EMAIL_ADDRESS	PHONE_NUMBER	NATIONAL_ID	CREDITCARD_NUMBER	CREDITCARD_TYPE	CREDITCARD_EXP	CREDITCARD_C
Ohio	OH	44646	wfronsek1@source...	434-553-8337	231-24-6500	7df7ea4aac5e3493ea3127e	VISA	5/23	4952
Washington	WA	98373	dcoyekj@pcworld.co...	760-277-6466	649-96-7557	db7d354d8a0d2042c969dd	Diners Club	3/21	2218
California	CA	91604	kwaggatdu@typepx...	915-586-6081	962-90-5618	c6656b4696af2bcad045	American Express	8/19	1600
New York	NY	11570	vcokeriy@shareasal...	817-205-8994	613-45-8368	2802c0eabcf2fec9d4bc9	American Express	12/23	7608
Oregon	OR	97214	ppotegi@cnet.com	901-313-6753	873-61-4046	e3c9064b11ac1a8a446fd	Discover	4/23	5886
Virginia	VA	23220	lmebs68@mozilla.o...	318-710-5442	950-96-4106	11776326fed27128f89ac	JCB	7/20	4901
Colorado	CO	80021	cburcher1o@tuttoci...	214-501-0431	708-76-8249	81471b1239519e9b69c8	VISA	7/20	9952
Ohio	OH	43147	tgeeves9w@nsw.go...	212-160-2937	805-98-0233	738f72ee7485cdcd47a3c	JCB	10/19	6027
California	CA	94541	ovincentqy@weibo.i...	520-774-2490	061-31-7031	978312a843698d357512	JCB	12/22	8397
Ohio	OH	43035	ptythertonix@phoc...	608-176-4288	353-89-2233	5af9d98a2efda868e10949	Diners Club	12/22	9587

# Watson Knowledge Catalog: Glossary Terms

Business Glossary

Published Draft Archived

Business Glossary

Sort by: T

What are you looking for?

# A B C D E F G H I J K L M N O P Q R S T U V

Confidential Information

Data that is classified as confidential

Confidential

Last modified: Jan 2, 2018

Personally Identifiable Information

Data that is classified as Personally Identifiable Information (PII).

Business Glossary / Sensitive Information

Overview Related content

Sensitive Information

Sensitive information that is protected against unwarranted disclosure.

Term details

Creator: Ricardo Buglio  
Date created: Oct 31, 2017  
Last editor: Ricardo Buglio  
Last modified: Jan 10, 2019

Owner

Ricardo Buglio

Tags

Sensitive

Associated classifier or term

Sensitive Personal Information

Add Terms

Business Glossary

Sort by: Term name

What are you looking for?

# A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Confidential Information

Data that is classified as confidential

Confidential

Last modified: Jan 2, 2018

Personally Identifiable Information

Data that is classified as Personally Identifiable Information (PII).

PII

Last modified: Apr 23, 2018

Sensitive Information

Sensitive information that is protected against unwarranted disclosure.

Sensitive

Last modified: Jan 10, 2019

Business Glossary / Sensitive Information

Overview Related content

Sensitive Information

Archive

Related content by type

Related content

Assets (9) View all

ASSET NAME	CATALOG NAME	LAST MODIFIED
Auto Insurance Customers		Mar 30, 2018
CUSTOMERS	Enterprise	Nov 9, 2018
CUSTOMER_DEMOGRAPHIC	Enterprise	Jan 10, 2019
Insurance Driver	Knowledge Catalog	Nov 16, 2018

16 Total

9 Assets 1 Policies 6 Rules

# Watson Knowledge Catalog: Data Rules and Policies

Policy Manager / Data Privacy

+ Add ▾



## Data Privacy

### Description

Enterprise data privacy policies pertaining to data classified as Sensitive, Classified or Personally Identifiable Information that have to be protected and comply with government and industry regulations.

Creator: Ricardo Buglio    Created: Jul 31, 2017    Contains: 3 published policies and 4 rules

What are you looking for?

NAME	TYPE	STATUS	CONTAINS	LAST MODIFIED	⋮
▼ Sensitive Information	Policy	Published	2 Items	Jan 10, 2019	⋮
Anonymize Government Identity Information	Rule	Published	--	Jan 10, 2019	⋮
Anonymize Financial Account Information	Rule				Data Dashboard
➤ Confidential Information	Policy				
➤ Personally Identifiable Information	Policy				

Data Dashboard

Policy enforcements over time ⓘ

Dec 16, 2018 - Jan 16, 2019



Action \* ⓘ

then anonymize data

in columns containing

Government Identities ✖

Select how to anonymize data:

Redact

BEFORE  
452-821-1120

Replace data with Xs.

Substitute

BEFORE  
452-821-1120

Replace data with values that don't match the original format.

Mask

BEFORE  
452-821-1120

Replace data with similarly formatted values.

[Visit the documentation](#) to learn more about data anonymization.

### Data assets containing personal or restricted data

#### Confidential Information ⓘ

2

[View all](#)

#### Personally Identifiable Infor... ⓘ

2

[View all](#)

#### Sensitive Information ⓘ

9

[View all](#)

### Operational policies ⓘ

3

Data policies

4

Data policy rules

### Automatic enforcement ⓘ

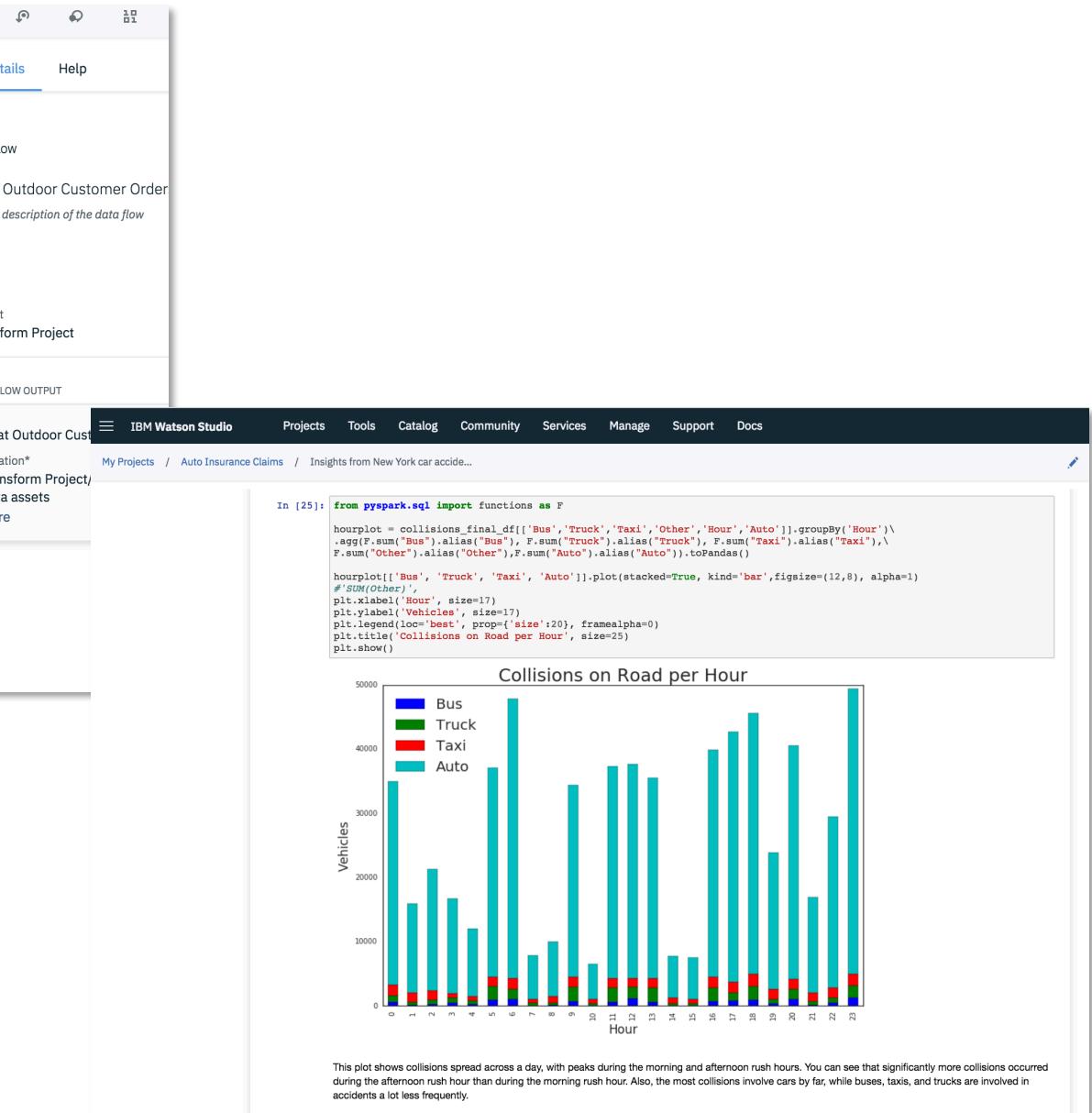
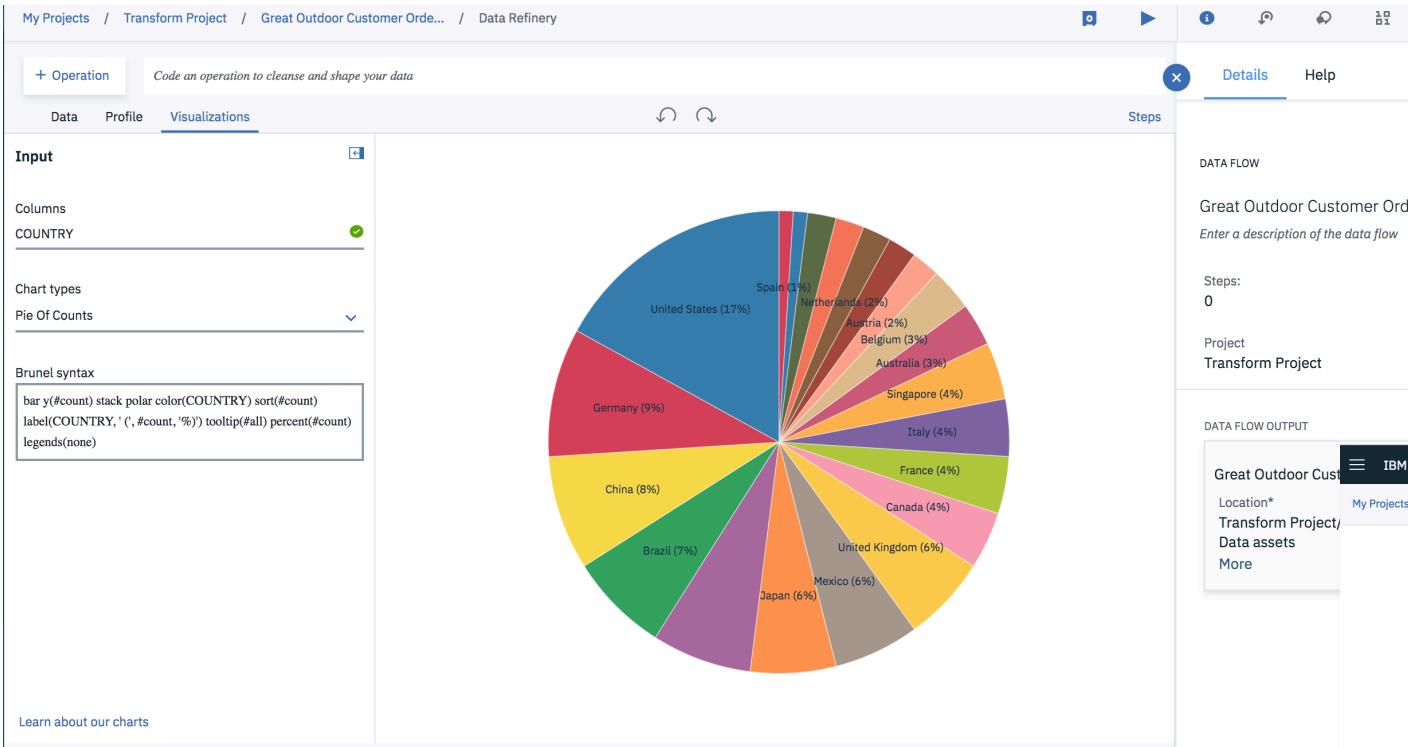
230

enforcements in last 30 days

71.71%

from last month

# Watson Knowledge Catalog: Data Activity



# Data Rule Definitions

IBM Information Server    Catalog    Connections    Monitoring    Quality    Management    Help    AI

Workspaces    Data rules    Data sets    Discover    Analysis    Report    Exceptions    Select a workspace

DataLakeWorkspace  
Workspace with optimal settings to run a quick analysis.

Owner: isadmin  
Datasets: 5    Analyzed: 0%  
Low quality: 0    Data rules: 0

InDepthAnalysisWorkspace  
Workspace with optimal settings to run a detailed analysis.

Owner: isadmin  
Datasets: 4    Analyzed: 0%  
Low quality: 0    Data rules: 0

PIIWorkspace  
TestWorkspace

IBM Information Server    Catalog    Connections    Monitoring    Quality    Management

Workspaces    **Data rules**    Data sets    Discover    Analysis    Report    Exceptions

DataLakeWorkspace

**Rule definition**

Name	Description	Status	Created by	Created on	Terms	Actions
All	Global category for all QualityComponents			1/11/2019, 5:45:53 PM		
Published Rules						
01 Personal Identity				1/11/2019, 5:44:23 PM		
Age				1/11/2019, 5:44:23 PM		
ChildInRangeString	String data Child: Age >= 0 and < 18; applied to string age data	ACCEPTED	isadmin	1/11/2019, 5:44:24 PM		
AdultInRangeNumeric	Adult: Age >= 18 and < 125; applied to numeric age data	ACCEPTED	isadmin	1/11/2019, 5:44:23 PM		
AgeInRangeNumeric	Age: Age >= 0 and < 125; applied to numeric age data	ACCEPTED	isadmin	1/11/2019, 5:44:22 PM		
AgeInRangeString	String data Age: Age >= 0 and < 125; applied to string age data	ACCEPTED	isadmin	1/11/2019, 5:44:23 PM		
AdultInRangeCalc	Derived Age Adult: Age >= 18 and < 125; applied to derived age calculated as the absolute value of the difference between the current date and date of birth	ACCEPTED	isadmin	1/11/2019, 5:44:24 PM		
ChildInRangeCalc	Derived Age Child: Age >= 0 and < 18; applied to derived age calculated as the absolute value of the difference between the current date and date of birth	ACCEPTED	isadmin	1/11/2019, 5:44:25 PM	⋮	
ChildNotMarriedString	If Child (string) then Marital Status = 'N'; applied to string age data	ACCEPTED	isadmin	1/11/2019, 5:44:25 PM		