IBM Watson Data Platform

Jay Limburn
Distinguished Engineer & Director of Offering Management
October, 2017

Why do companies struggle to deliver value?

Data

- Data resides in silos
- Detailed data was never stored
- Unstructured and external data wasn't considered
- Difficult to access

Governance

- If the data isn't secure, self-service isn't a reality
- Understanding lineage and getting to a system of truth

Skills

- Data Science skills are in low supply and high demand
- Nurturing new data professionals is challenging

Infrastructure

- Need an environment that enables a "fail fast" approach
- Discrete tools present barriers to progress

Knowledge workers spend

80%

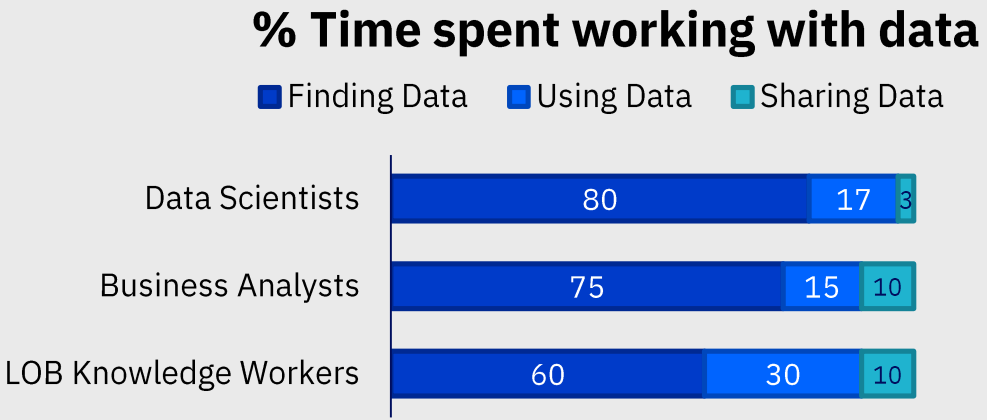
of their time searching for the correct data

.....imagine if that time was spent doing Data Science and Analytics!

The Data Lake Fallacy

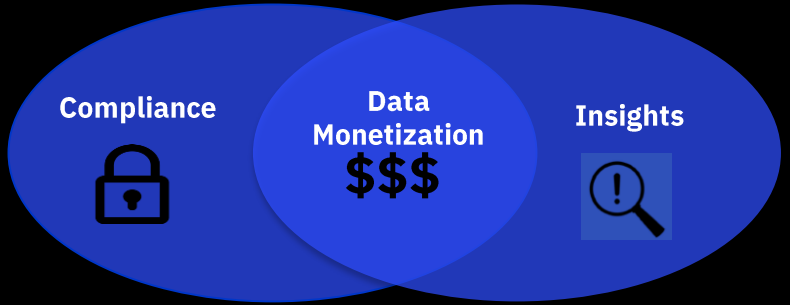
Enterprise Data Lakes are not delivering on their promise

- Inability to easily find data
- Lack of trust in how the data will be used
- Obstacles to finding and sharing
- Too difficult to ingest sources



Users spent significantly more time finding the correct data, rather then extracting value from it.

Chief Data Officers need answers!



WHERE is all my data?

Compliance

WHO is using it?

Compliance

HOW can I monetize it?

Insights

WHAT can I do to improve it?

Insights

Watson Data Platform

an integrated platform of tools, services and data that helps companies accelerate their shift to become data-driven organizations

Drive governance policy effectiveness while tracking how data is used and its value to the company

Access powerful tools to prepare data to tease out the insights they're looking for, without IT involvement

Data Steward

Data Scientist

Data Engineer

App Developer

Easily build data pipelines that power dashboards and data platforms while ensuring high quality

Make the insights immediately actionable and add intelligence to apps in straightforward manner

Watson Data Platform

The IBM foundation for data innovation

Applications, solutions and services

Targeted solutions for enterprise businesses

Data Science

IoT

Virtual agent

AI

Cognitive building blocks for developers

API Conversation	API Discovery	API Compare and Comply	API DLaaS	API Tone Analyzer	API Personality Insight
API Visual Recognition	API Speech	API Document Conversion	API Nat Language Understanding	API Nat Language Classifier	…more

Data

Tools to prepare data for cognitive

Ingestion → Storage → Analytics → Deployment → Governance

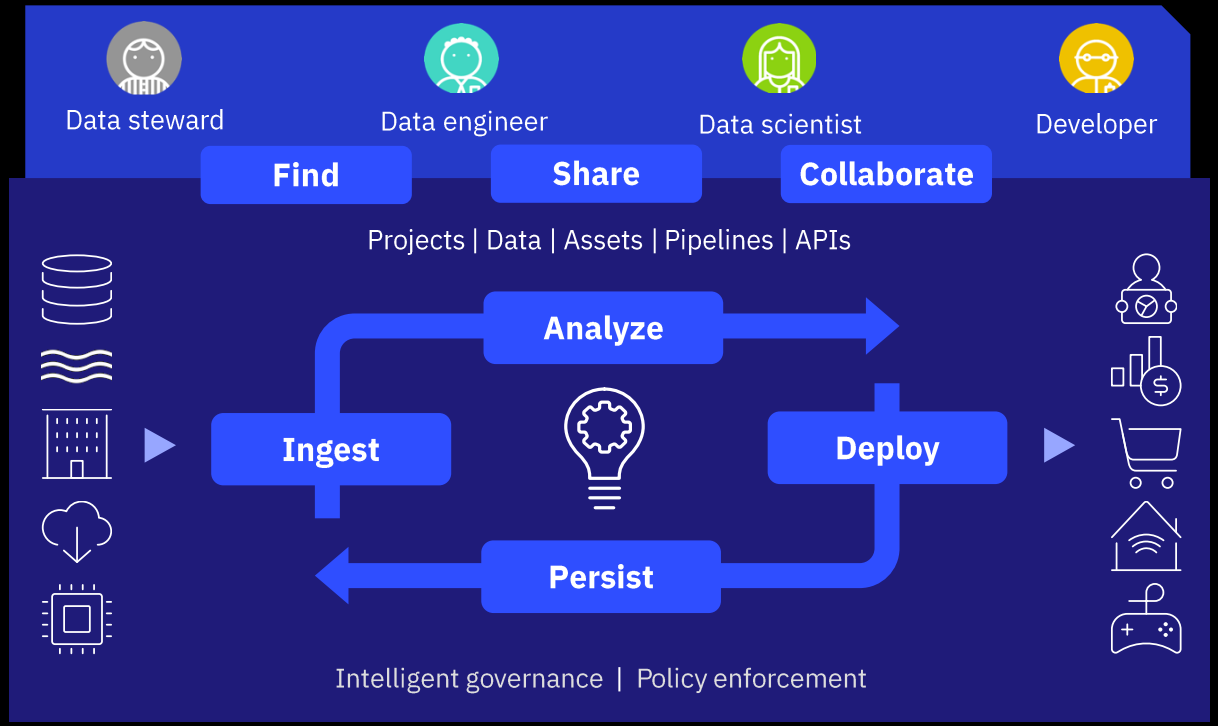
Cloud Infrastructure

A highly scalable, security enabled infrastructure

Cloud Integration		Cognitive Micro-services			DevOps Tooling	
Networking	Compute	Security	Dedicated Compute	Virtual Compute	Object Storage	…more
Public		Hybrid			Private	

Watson Data Platform

Making data into insights is a team sport

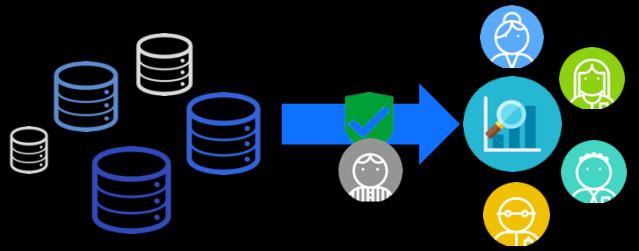


Our Core Tenets

1. Intelligent by Design
2. Collaborative for data Professionals
3. Self-service access to trusted data
4. Best in class streaming and real-time analytics
5. Open and Extensible

Watson Data Platform

Democratized data and analytics



Projects and communities

enable team collaboration ensuring the enterprise builds upon the insights its members develop and springboards from those that peer organizations and experts share.

Common tools

including open source notebooks and tools, shapers and data visualizations that enable the team to quickly investigate any volume of data

Catalog

to organize, find and understand data that enables every project team to quickly discover and use data they can trust

Data access

to disparate data sources that makes it easy to connect to and gather data from anywhere without friction

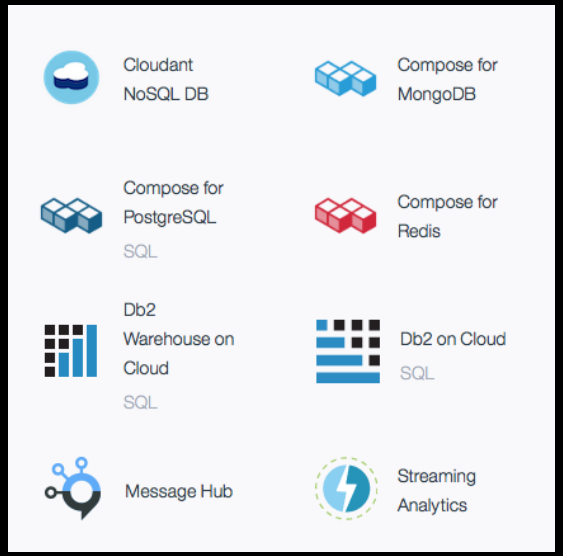
Enterprise grade governance

enabling regulatory compliance with automatic policy enforcement that enables data professionals to focus on delivering valuable insights rather than managing and organizing data

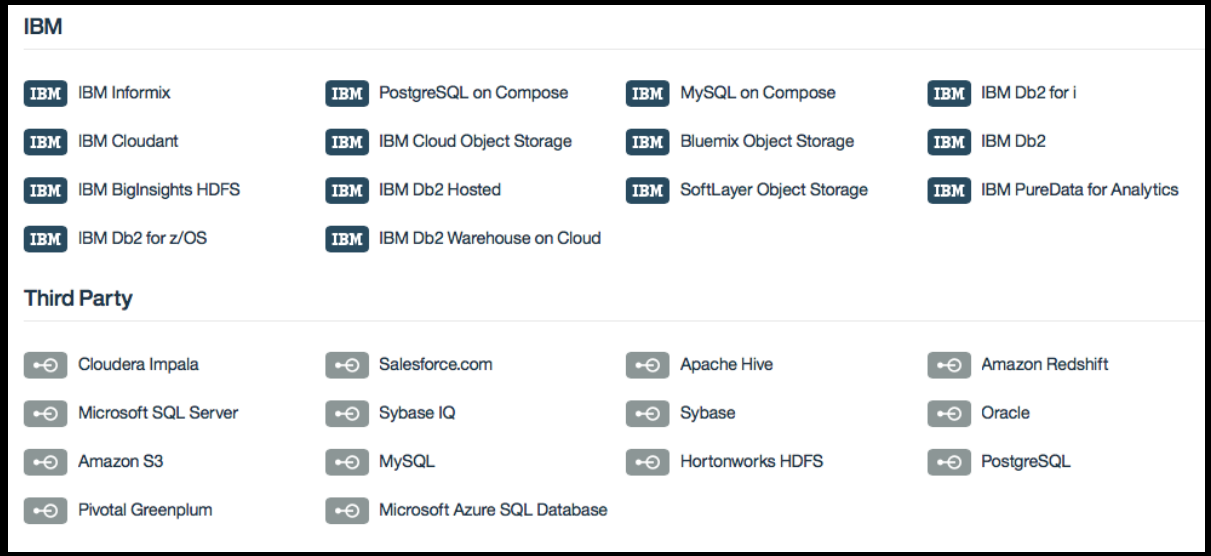
Watson Data Platform

Data at your fingertips

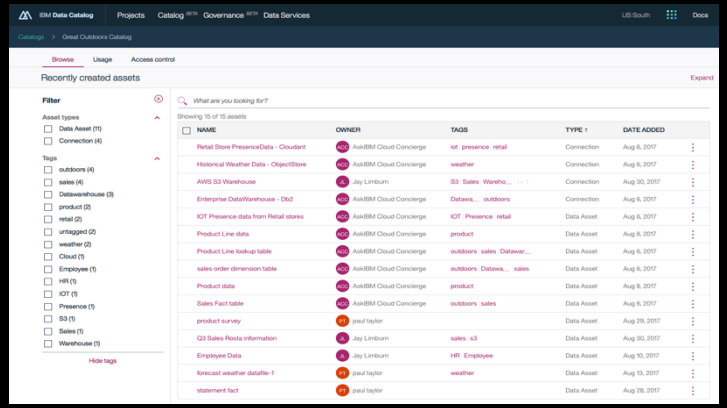
In the IBM Cloud

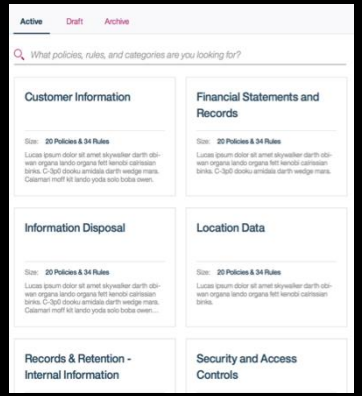


Where your data is



Data Catalog





Unlock tribal knowledge to unleash your data professionals

Discover

Intelligent discovery of data, advanced classification and profiling to provide context

Catalog

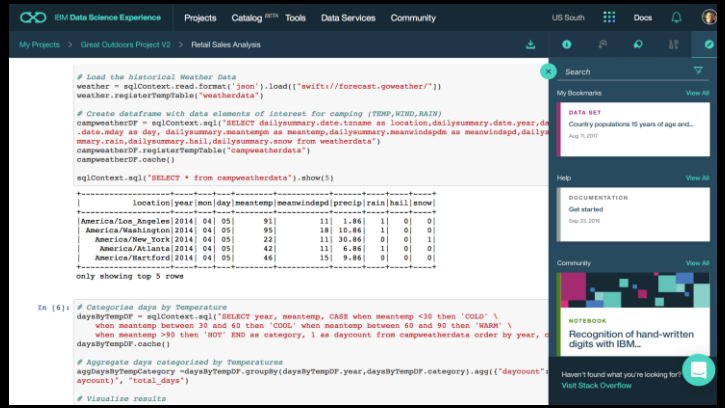
A rich metadata index of all data, with social collaboration and enhanced findability

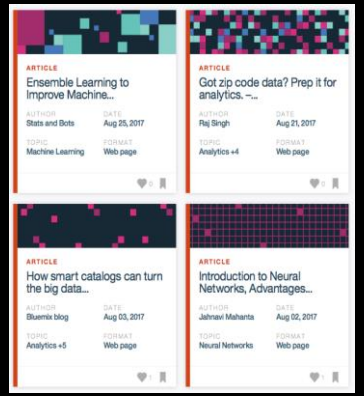
Govern

Powerful governance policy tools to control and protect access to data with visibility to data use

Watson Data Platform

Data Science Experience





Making data science a team sport

Learn

Built-in learning to get started or go the distance with advanced tutorials

Create

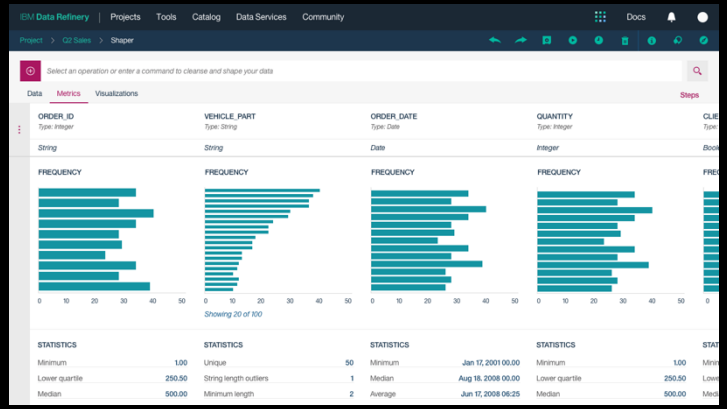
The best of open source and IBM value-add to create state-of-the-art data products

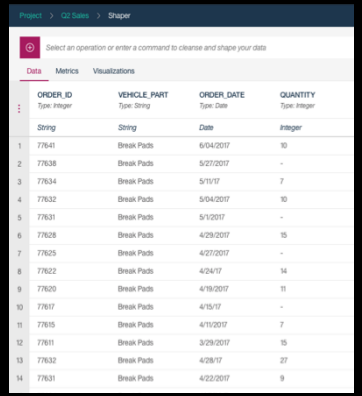
Collaborate

Community and social features that provide meaningful collaboration

Watson Data Platform

Data Refinery





A Breakthrough Approach to Explore and Prepare Data

Wrangle

Interactively explore, resolve quality issues, enrich, classify, standardize, summarize and join data

Flow

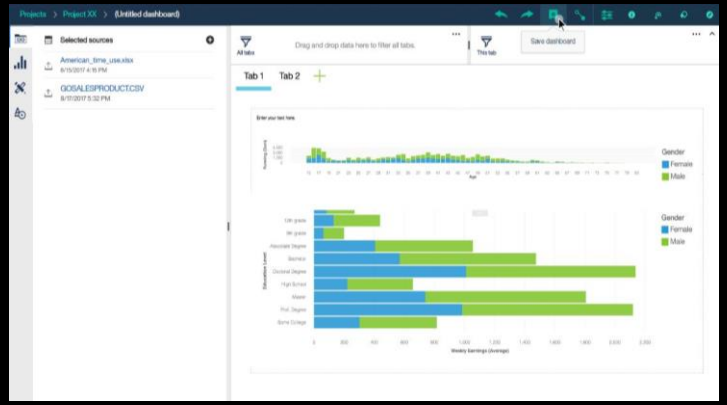
Create data flows visually, schedule for repeatability, monitor and notify

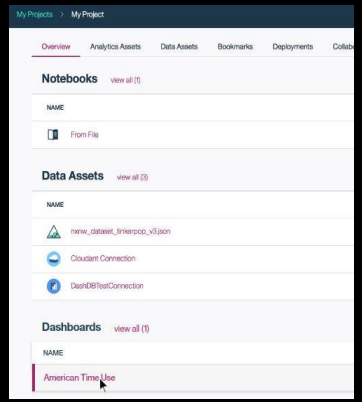
Adapt

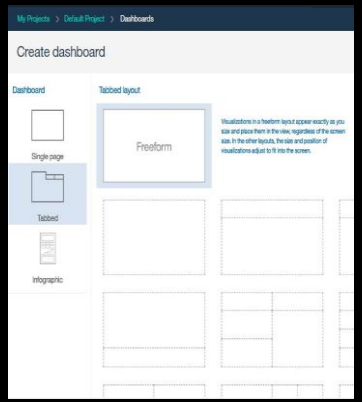
Connect to 30+ cloud and on-premises stores and scale on demand with cataloging and governance

Watson Data Platform

Dashboards







Making Insights Available to All

Visualize

Represent analytic results as compelling interactive graphics.

Share

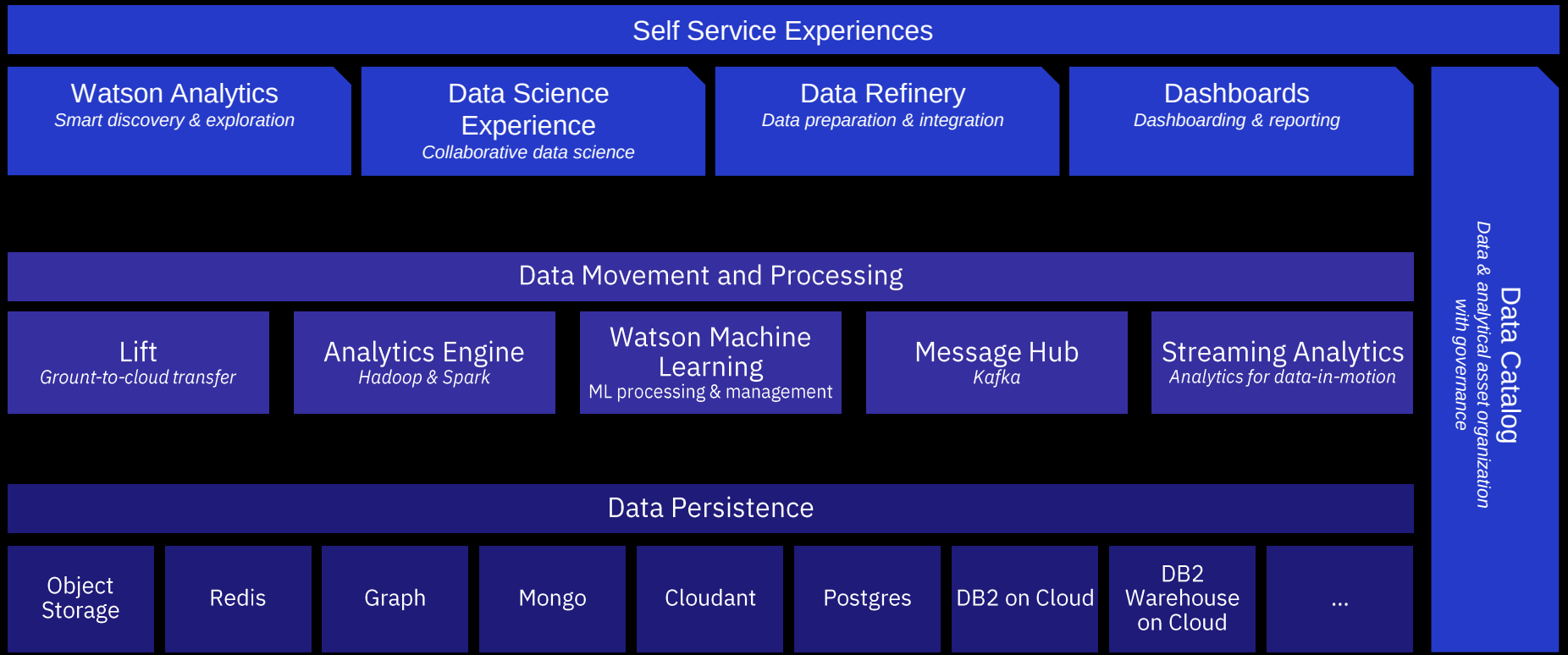
Deliver dashboards across the enterprise where analytical insights are required to make business decisions.

Discover

Utilize data assets from the entire organization with confidence by accessing them through the Data Catalog.

Watson Data Platform

Services of the integrated platform



IBM Data Catalog

Powered by the Watson Data Platform

360 View of Data

Data Science

Allow Data Scientists to spend more time doing Data Science!

Business problem

Data Scientists can’t easily locate the data they need. Projects typically start with an exhaustive effort to find locate and gain access to the data they need. What’s worse is they don’t even know what other data could help them find a different answer.

Solution

Watson Data Platform allows data to be easily indexed and classified. Intelligent search algorithms can quickly guide your Data Scientists to the best data for their purpose. Once found they can move the data into their sandboxes. Better still Data Scientists and share and collaborate on their models allowing for reuse and accelerating the initiation of new projects.

Compliance Office

Transform your business to become Data Driven

Business problem

Too much data, not enough use. How can the compliance office monetize their data, yet ensure it is protected and governed at the same time? The Compliance Team don't have the tools or understanding to drive the cultural change required to unlock their organizations value in their data.

Solution

Watson Data Platform provides an easy way to build and index of all the assets across your business. It builds an intelligent model for those assets and provides context around those asset to make them easy to find and understand how that data should be used, accessed and managed.

Business Analysis

Turn every worker into a knowledge worker that thinks about data first

Business problem

Tribal knowledge is locked away in department silos. Business analysts know this data is locked away but don’t know how they can get hold if it to improve their analytics. They know their analytics are not as accurate as they need to be.

Solution

Watson Data Platform makes it easy for tribal knowledge to be shared and accessible across the enterprise. Every worker can easily collaborate and curate assets to aid the finadability of data for all. It makes finding data a delightful online shopping experience for all and allows them to provide their own assets back into the marketplace.

Data Engineering

Optimize your Data Engineers by reducing duplication and providing self service access to data allowing them to focus on the most important jobs

Business problem

Your data engineers are working as hard as they can but it takes too long to provide data to the business when they request it. The complexities of moving, protecting and cleansing the data takes time and the business teams can’t wait.

Solution

Watson Data platform provides self service access to data. The intelligent catalog allows all users to know what data exists and the tools to allow them to go and get it. The business users no longer need to wait on IT to get the data they need and the Data Engineers can focus on the business critical tasks essential for business growth.

Application Development

Supercharge your next generation application development

Business problem

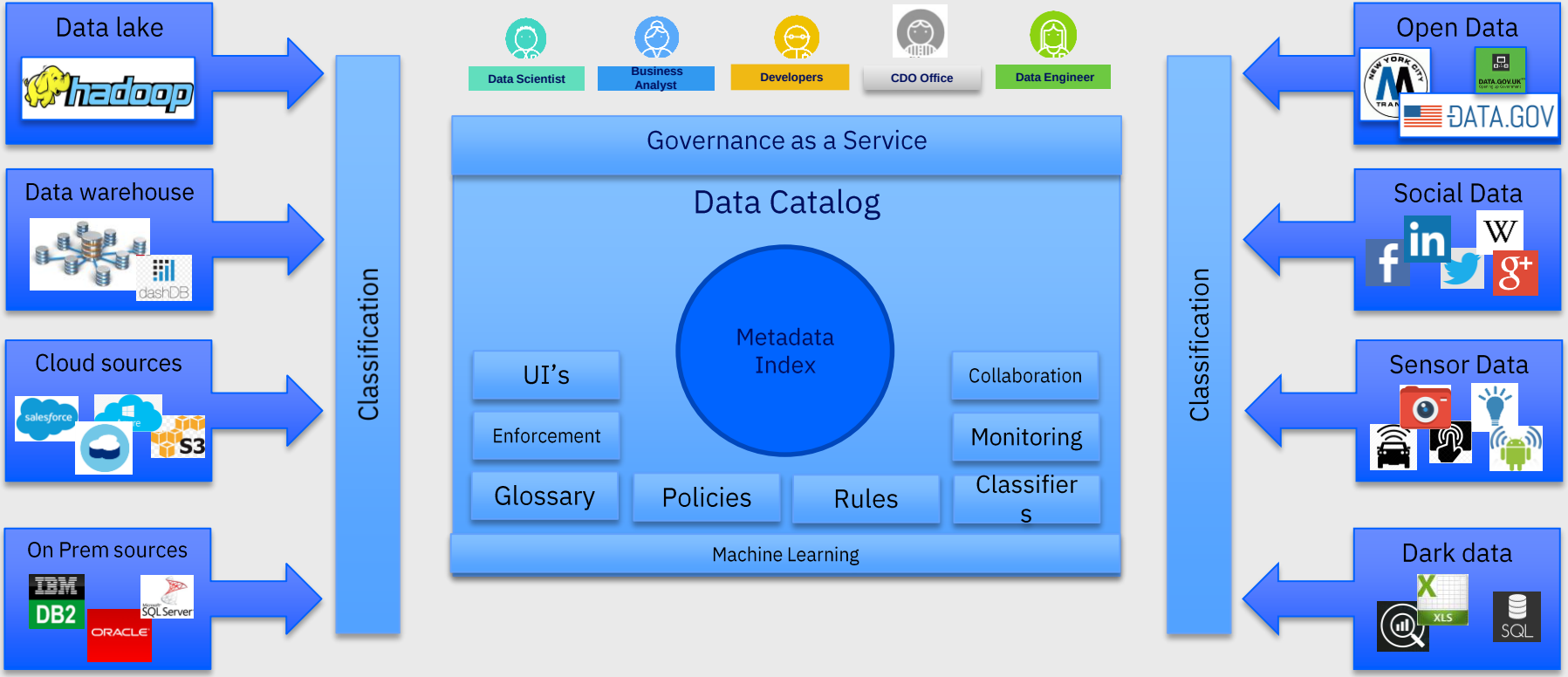
Data is growing, changing, being duplicated, being moved. Application Developers add to this problem by creating new silos of data to power their applications. In addition they are unable to understand what data should be used to power their applications

Solution

IBM Data Catalog provides an intelligent index of all data so that application developers can be guided towards the right assets for consumption within their next generation cognitive applications, but more importantly any new data is cataloged and made available to knowledge workers for further insights.

Intelligent Metadata Index

- Providing 360 view of **all** data for **all** users
- Data can reside in original systems, but users can discover it quicker join together for smarter analytics



Governance Framework

- Govern Data throughout its lifecycle
- Automatically discovery and provide context to data
- Ensure it is used in a compliant manner
- Aid findability of data
- Take action to improve the data and its use
- Empower your knowledge workers to be more effective with their tools

A fabric that supports the data governance objectives of our clients throughout the analytics lifecycle

Define

Policy Management

Rule Authoring

Term Definitions

Discover

Automated Metadata discovery

Manual ingest of data

Classify

Automatic profiling of data to relate to data policies

Assignment of Business Terms to Technical Assets

Share

Publish to catalog

Allow knowledge workers to curate and collaborate with data

Enforce

Automated policy enforcement throughout its lifecycle within the catalog

Monitor

Understand how data is used and found.

Take action to improve data reuse

IBM Data Catalog

Powered by the Watson Data Platform

ALL Users
ALL data
EVERYWHERE

.....Never before has so much information been available to so many

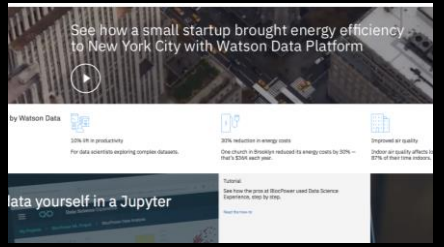
Continue your WDP Exploration

https://www.ibm.com/analytics/us/en/watson-data-platform/

Follow our thought leadership series:

https://www.ibm.com/blogs/bluemix/tag/ibm-data-catalog/

See how a small startup brought energy efficiency to NYC with WDP



https://www.ibm.com/analytics/us/en/watson-data-platform/blocpower/

Browse & Share the WDP Tutorial

https://ibm.biz/wdptutorial