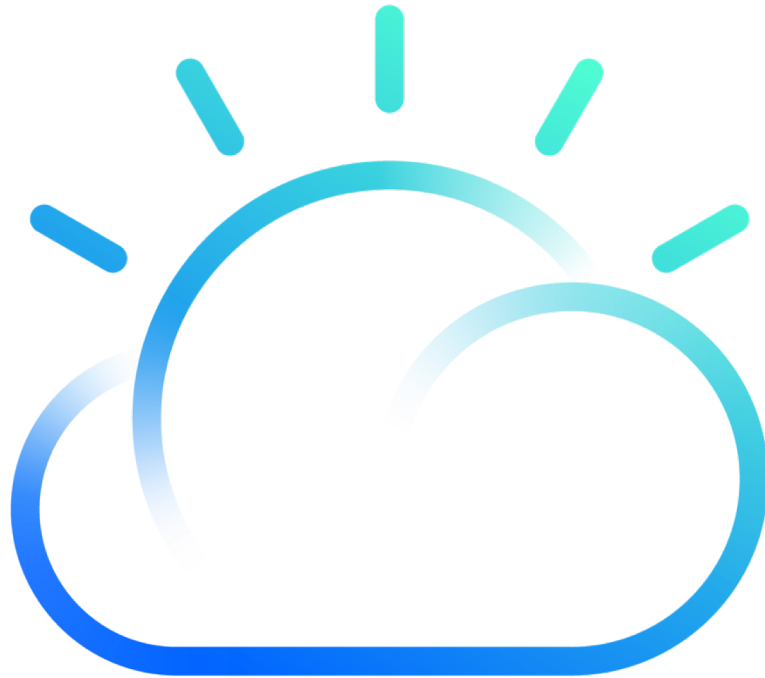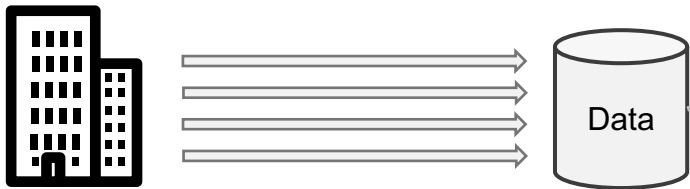# IBM Cloud Platform Overview for Data Scientists

Thomas Schaeck
Distinguished Engineer, Watson Studio

IBM

IBM Cloud

Data Science and ML in Enterprises

Enterprises generate TONS OF DATA
... requiring cataloging and governance

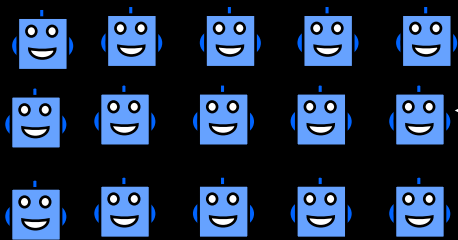Data must be cleaned and shaped for training

need good data   *avoid bias*
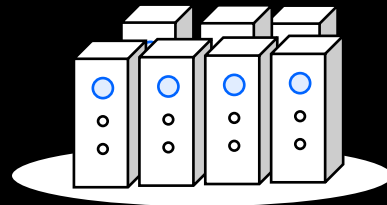
... then models must be created

... and trained on high performance compute

to select an optimal model...

...that must be hosted and monitored at scale

Why are enterprises struggling to capture the value of AI?

Data Access

- Data resides in silos & difficult to access

- Unstructured and external data wasn't considered

Data Governance

- If the data isn't secure, self-service isn't a reality

- Challenge understanding data 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

Tools&Infrastructure

- Need an environment that enables quick experiments and real outcomes

- Discrete tools present barriers to productivity

Watson AI and Data

Applications, solutions and services — Targeted solutions for enterprise businesses	Healthcare	Financial Services	Logistics		**Watson Studio & Knowledge Catalog**	IoT	Virtual agent

|					
Watson AI — Cognitive building blocks for developers	API Assistant	API Discovery	API Compare and Comply	API **Watson Machine Learning**	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 AI and data persistence	Ingestion →	Storage →	Analytics →	Deployment →	Governance
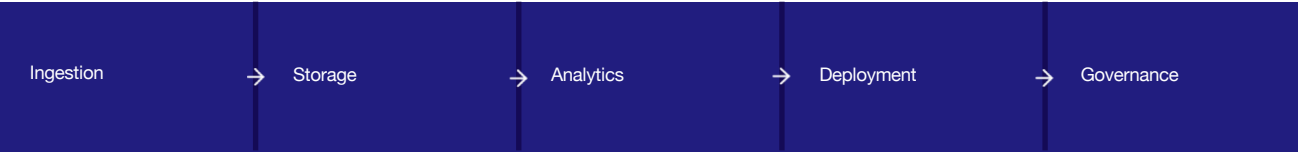
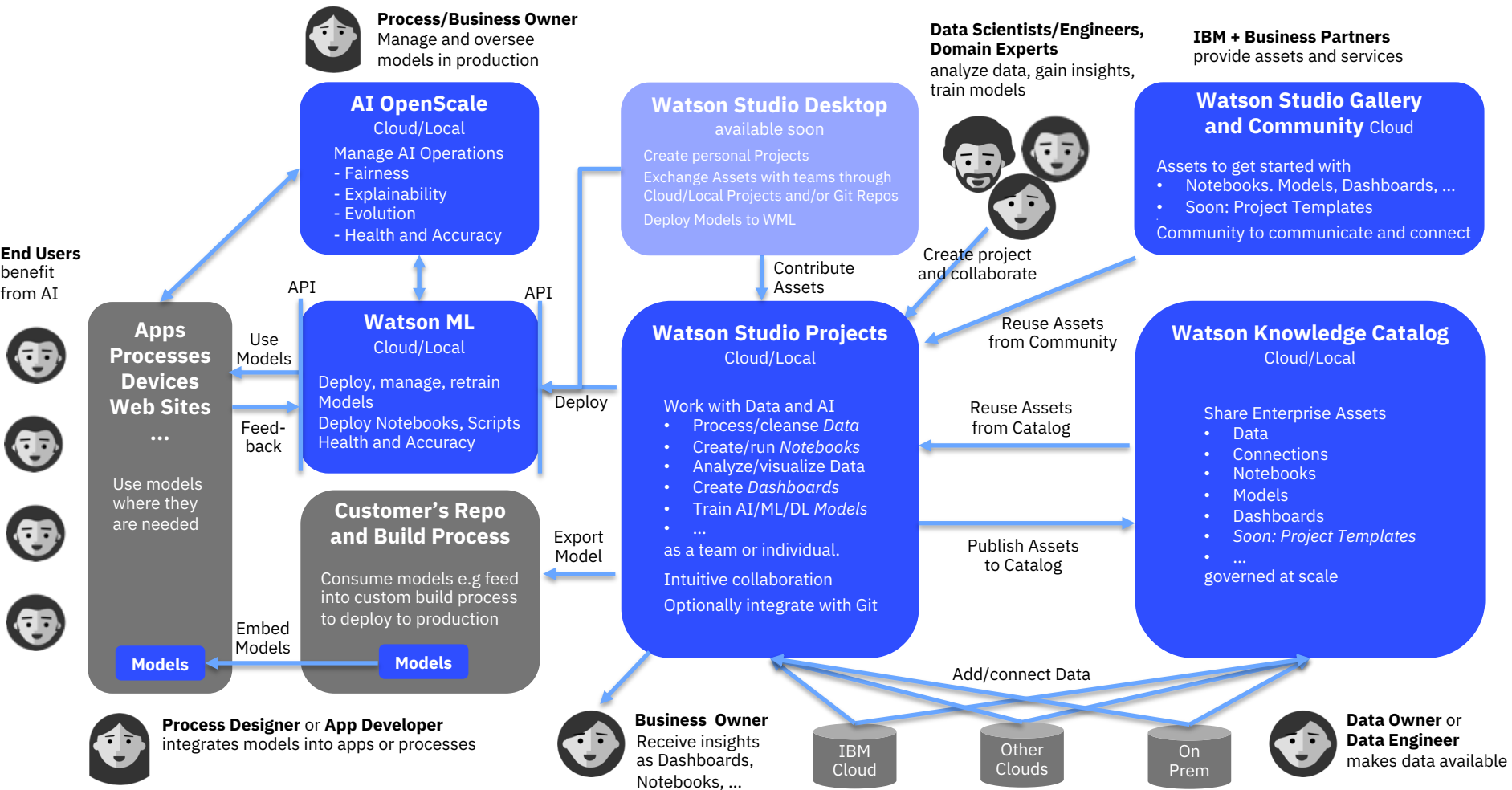**Cloud Infrastructure** — A highly scalable, security enabled infrastructure

Cloud Integration	Cognitive Micro-services	DevOps Tooling

Networking	Compute	Security	Dedicated Compute	Virtual Compute	Object Storage	…more

Public	Hybrid	Private

Collecting, organizing, analyzing Data → Training and deploying Models → Managing AI Ops

Process/Business Owner
Manage and oversee models in production

Data Scientists/Engineers, Domain Experts
analyze data, gain insights, train models

IBM + Business Partners
provide assets and services

AI OpenScale
Cloud/Local

Manage AI Operations
- Fairness
- Explainability
- Evolution
- Health and Accuracy

Watson Studio Desktop
available soon

Create personal Projects
Exchange Assets with teams through Cloud/Local Projects and/or Git Repos
Deploy Models to WML

Watson Studio Gallery and Community Cloud

Assets to get started with
- Notebooks. Models, Dashboards, …
- Soon: Project Templates
.
Community to communicate and connect

End Users
benefit from AI

API

API

Contribute Assets

Create project and collaborate

Reuse Assets from Community

Apps Processes Devices Web Sites …

Use models where they are needed

Use Models

Feed-back

Watson ML
Cloud/Local

Deploy, manage, retrain Models
Deploy Notebooks, Scripts
Health and Accuracy

Deploy

Watson Studio Projects
Cloud/Local

Work with Data and AI
- Process/cleanse *Data*
- Create/run *Notebooks*
- Analyze/visualize Data
- Create *Dashboards*
- Train AI/ML/DL *Models*
- …
as a team or individual.

Intuitive collaboration
Optionally integrate with Git

Reuse Assets from Catalog

Watson Knowledge Catalog
Cloud/Local

Share Enterprise Assets
- Data
- Connections
- Notebooks
- Models
- Dashboards
- *Soon: Project Templates*
- …
governed at scale

Publish Assets to Catalog

Customer's Repo and Build Process

Consume models e.g feed into custom build process to deploy to production

Export Model

Embed Models

Models

Models

Process Designer or **App Developer**
integrates models into apps or processes

Business Owner
Receive insights as Dashboards, Notebooks, …

Add/connect Data

IBM Cloud

Other Clouds

On Prem

Data Owner or **Data Engineer**
makes data available

Scaling AI in the enterprise

Connect to Data in IBM Cloud Data Services, On Prem Data, or Data on other Clouds

Store your data in the IBM Cloud   and/or connect your on-prem data or third-party cloud data

Cloud Object Storage   SQL Query

Cloudant
NoSQL DB

Compose for
MongoDB

Compose for
PostgreSQL
SQL

Compose for
Redis

Db2
Warehouse on
Cloud
SQL

Db2 on Cloud
SQL

Message Hub

Streaming
Analytics

...

IBM

IBM	IBM Informix	IBM	PostgreSQL on Compose	IBM	MySQL on Compose	IBM	IBM Db2 for i
IBM	IBM Cloudant	IBM	IBM Cloud Object Storage	IBM	Bluemix Object Storage	IBM	IBM Db2
IBM	IBM BigInsights HDFS	IBM	IBM Db2 Hosted	IBM	SoftLayer Object Storage	IBM	IBM PureData for Analytics
IBM	IBM Db2 for z/OS	IBM	IBM Db2 Warehouse on Cloud				

Third Party

Cloudera Impala	Salesforce.com	Apache Hive	Amazon Redshift
Microsoft SQL Server	Sybase IQ	Sybase	Oracle
Amazon S3	MySQL	Hortonworks HDFS	PostgreSQL
Pivotal Greenplum	Microsoft Azure SQL Database		

IBM Cloud SQL Query – Serverless SQL aaS
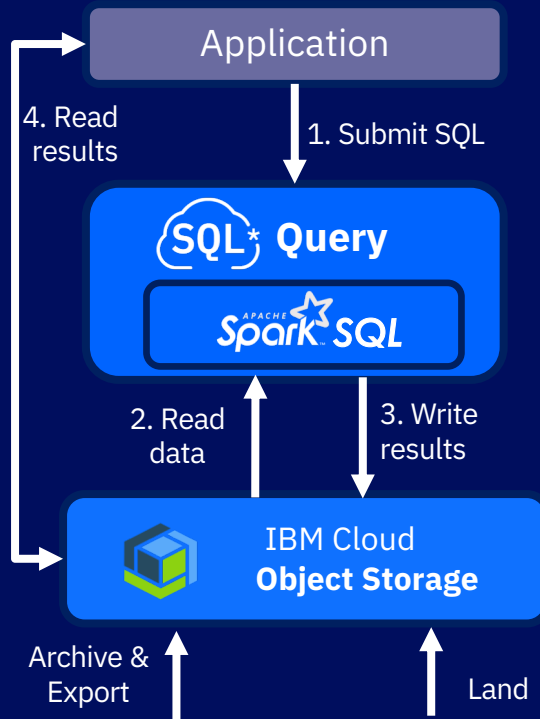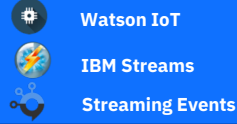
- Completely **self-service SQL** on cloud data
- **Serverless** scale-out SQL execution on **TB of data**
  - **Pay per query** (pennies per query)
  - Full **ANSI SQL** support, incl. **OLAP** functions & **Spatial** SQL
- Easy **SQL operationalization** for **developers**
  - **REST** API, **Cloud Function** integration, SDKs (**Python, Node.js**)
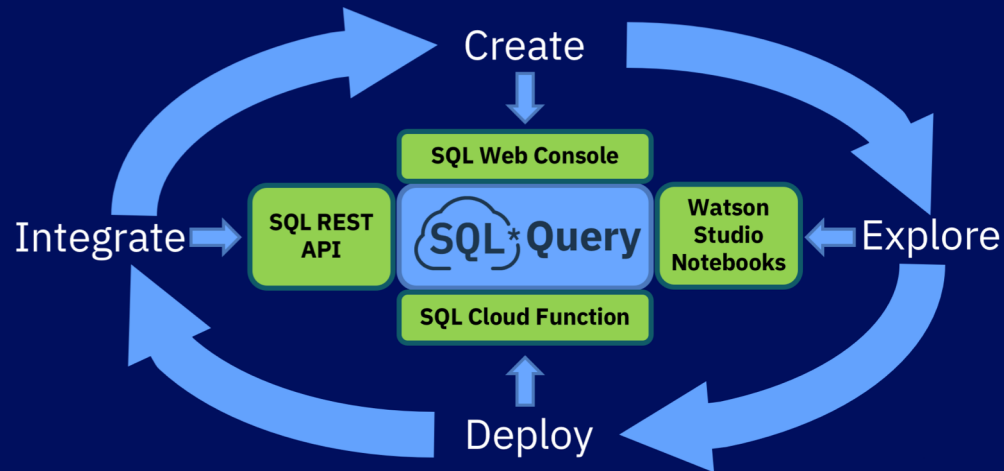
IBM Cloud Databases

A **fully managed** collection of **open source** databases and **persistence** services available through a consistent consumption, pricing, and interaction model

Eliminates the burden of **managing** infrastructure and system health as well as the associated capital expenses

Allows you to **focus on** your **business** objectives rather than the operations and support of your own platform stack

Responsiblities comparison

	On-prem	IaaS	DbaaS
	Performance	Performance	Performance
	Data Model	Data Model	Data Model
Client Managed (DIY)	Monitoring	Monitoring	Monitoring
	Backups	Backups	Backups
VS	High Availability	High Availability	High Availability
	DB Technology	DB Technology	DB Technology
IBM Managed	Security	Security	Security
	Networking	Networking	Networking
	Hardware	Hardware	Hardware

Avoid vendor lock-in and allows for easily coordinating multi-database strategies by using Open Source **DBaaS** like:

- **Postgres**
- **Mongo**
- **ElasticSearch**
- **Redis**
- **RabbitMQ**
- **etcd**
- **MySQL**
- **ScyllaDB**
- **JanusGraph**

IBM **Watson**   Germany

Community   Blog   Docs   Log In   Sign Up
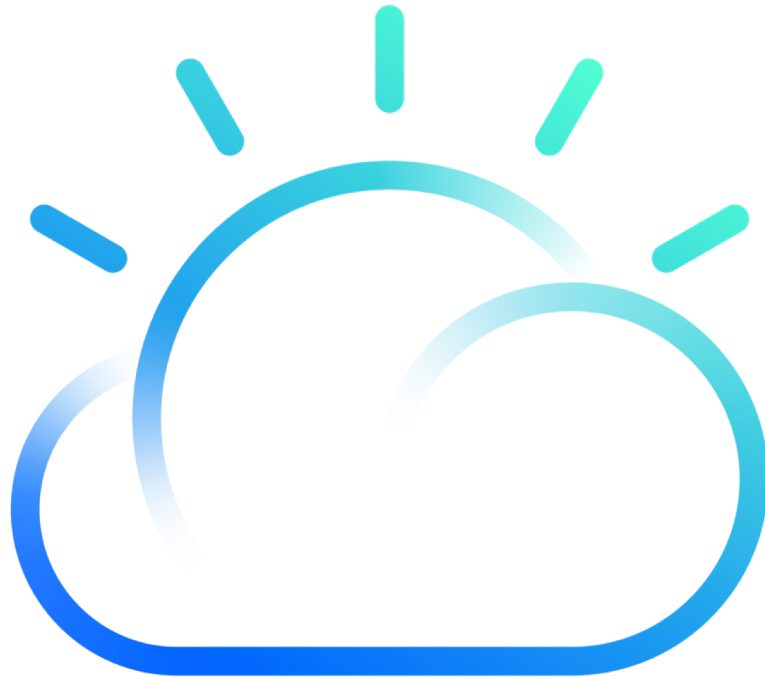
IBM **Watson**

Log In   Try it for Free

Explore our apps

Use IBM Watson to collaborate and build smarter applications. Quickly
visualize and discover insights from your data and collaborate across teams.

IBM **Watson Studio**

Democratize ML/DL to accelerate infusion of
AI in your business.

Learn more

IBM **Watson Knowledge Catalog**

Securely discover, catalog, and govern
enterprise data.

Learn more

Blog   Contact   Privacy   Terms of Use

Sign Up or use existing IBM Cloud Account – Watson Studio + Knowledge Catalog + Cloud Object Storage will be created for you

IBM Watson Studio

Try Watson Studio and Knowledge Catalog

Powered by IBM Cloud

Take advantage of machine learning an AI to analyze your data. Catalog your data to make it easy to find. Both applications are free and without time limit!

Your Watson Applications

Watson Studio

Machine learning and AI made easy! Solve your business problems in a collaborative environment.

Watson Knowledge Catalog

Connect the right data with the right people. Index, find, and protect your knowledge.

Activate Watson applications with your IBM Cloud account

Create an IBM Cloud Account

Email

☐ I accept the IBM Watson Studio terms and the IBM Watson Knowledge Catalog terms.

Next

Have an IBM Cloud account?

By clicking Log in, you agree to our terms and that you have read our Data Use Policy, including our Cookie Use.

Log in to activate Watson

Cookie Preferences

Create your own Projects to connect, analyze, visualize Data and create and train Models

Get started ▾

Welcome Thomas!

Watson Studio and Watson Knowledge Catalog are both part of IBM Watson Studio.

What's new? As part of the AI OpenScale beta program you can get started with AI operations today.

IBM **Watson**

Get started with key tasks

New project	Catalog and find data	Refine data	New notebook	Operationalize AI	Govern data

Hide ▲

Recently updated projects   View all (9)   ⊕ New project

NAME	ROLE	COLLABORATORS	DATE CREATED	LAST UPDATED
Armada TestFest7 Restricted Editor 01	Editor	SS AS +66	Sep 10, 2018	Sep 21, 2018
Armada TestFest7 Restricted Editor 02	Editor	AS YC +63	Sep 10, 2018	Sep 21, 2018
Armada TestFest7 Restricted Editor 03	Editor	AS YC +63	Sep 10, 2018	Sep 21, 2018

Watson Studio and Watson Machine Learning in the IBM Cloud Catalog

☰ IBM **Cloud**   Catalog   Docs   Support   Manage   🔍 Search for resource...   1716985 - Thomas Sc...   ⊙

Catalog

🔍 Search the catalog...   **Filter**

All Categories

- Compute
- Containers
- Networking
- Storage
- **AI** ❯
- Analytics
- Databases
- Developer Tools
- Integration
- Internet of Things
- Security and Identity
- Starter Kits
- Web and Mobile
- Web and Application

Watson Assistant (formerly Conversation)
Lite • IBM • EU Supported

Add a natural language interface to your application to automate interactions with your end users. Common applications include virtual agen...

AI OpenScale
Lite • IBM

IBM AI OpenScale is an enterprise-grade environment for AI infused applications that provides enterprises with visibility into how AI is...

Compare Comply
IBM • Beta

Process governing documents to convert, identify, classify, and compare important elements

Discovery
Lite • IBM • EU Supported

Add a cognitive search and content analytics engine to applications.

Knowledge Catalog
Lite • IBM

Discover, catalog, and securely share enterprise data.

Knowledge Studio
Lite • IBM • EU Supported

Teach Watson the language of your domain.

Language Translator
Lite • IBM • EU Supported

Translate text, documents, and websites from one language to another. Create industry or region-specific translations via the service's...

Machine Learning
Lite • IBM

IBM Watson Machine Learning - make smarter decisions, solve tough problems, and improve user outcomes.

Natural Language Classifier
IBM

Natural Language Classifier performs natural language classification on question texts. A user would be able to train their data and the predict...

Natural Language Understanding
Lite • IBM • EU Supported

Analyze text to extract meta-data from content such as concepts, entities, emotion, relations, sentiment and more.

Personality Insights
Lite • IBM • EU Supported

The Watson Personality Insights derives insights from transactional and social media data to identify psychological traits

Speech to Text
Lite • IBM • EU Supported

Low-latency, streaming transcription

Text to Speech
Lite • IBM • EU Supported

Synthesizes natural-sounding speech from text.

Tone Analyzer
Lite • IBM • EU Supported

Tone Analyzer uses linguistic analysis to detect three types of tones from communications: emotion, social, and language. This insight can...

Visual Recognition
Lite • IBM

Find meaning in visual content! Analyze images for scenes, objects, faces, and other content. Choose a default model off the shelf, or create your own...

Watson Studio
Lite • IBM

Embed AI and machine learning into your business. Create custom models using your own data.

Display a menu

IBM **Cloud**

Catalog   Docs   Support   Manage

Search for resource...

1716985 - Thomas Sc...

Catalog

Search the catalog...

Filter

All Categories

Storage

Compute

Containers

Networking

Storage

AI

Analytics

Databases

Developer Tools

Integration

Internet of Things

Security and Identity

Starter Kits

Web and Mobile

Web and Application

Block Storage
IBM

Persistent iSCSI based storage with high-powered performance and capacity up to 12TB.

File Storage
IBM

Fast and flexible NFS-based file storage with capacity options from 20GB to 12TB.

Object Storage
Lite • IBM

Provides flexible, cost-effective, and scalable cloud storage for unstructured data.

box
Third Party

Powering Content and data for your application. Whether you are building a line of business app, content management software or need to displa...

FEEDBACK

Display a menu

Database Services

Catalog   Docs   Support   Manage

Search for resource...

1716985 - Thomas Sc...

Search the catalog...   Filter

All Categories

Compute

Containers

Networking

Storage

AI

Analytics

Databases

Developer Tools

Integration

Internet of Things

Security and Identity

Starter Kits

Web and Mobile

Web and Application

Compose for etcd
IBM • Beta • EU Supported

etcd is a key/value store for distributed server configuration management.

Compose for JanusGraph
IBM • Beta • EU Supported

JanusGraph is a scalable graph database optimized for storing and querying highly-interconnected data

Compose for MongoDB
IBM • EU Supported

MongoDB is a JSON document store with a rich query and aggregation framework

Compose for MySQL
IBM • Beta • EU Supported

MySQL is a fast, easy-to-use, and flexible RDBMS.

Compose for PostgreSQL
IBM • EU Supported

PostgreSQL is an open source relational database. Consider using our new service, Databases for PostgreSQL.

Compose for RabbitMQ
IBM • EU Supported

RabbitMQ asynchronously handles the messages between your applications and databases

Compose for Redis
IBM • EU Supported

Redis is an open-source, blazingly fast, low maintenance key value store. Consider using our new service, Databases for Redis.

Compose for RethinkDB
IBM • EU Supported

RethinkDB is a JSON document based, distributed database with an integrated administration and exploration console.

Compose for ScyllaDB
IBM • Beta • EU Supported

ScyllaDB is a highly performant, in-place replacement for the Cassandra wide-column distributed database.

Databases for PostgreSQL
IBM

Postgres is a powerful, open source object-relational database that is highly customizable.

Databases for Redis
IBM

Redis is an open source (BSD licensed), in-memory data structure store, used as a database, cache and message broker.

Db2
Lite • IBM • EU Supported

A next generation SQL database. Formerly dashDB For Transactions.

Db2 Hosted
IBM

Db2 Hosted: Offers customers the rich features of an on-premise Db2 deployment without the cost,

Db2 Warehouse
IBM • EU Supported

Db2 Warehouse on Cloud is a flexible and powerful data warehouse for enterprise-level

Informix
IBM

IBM Informix on Cloud helps businesses gain a trusted view of data in a hybrid computing

SQL Query
Lite • IBM

Analyze data in Object Storage with ANSI SQL.

FEEDBACK

Display a menu