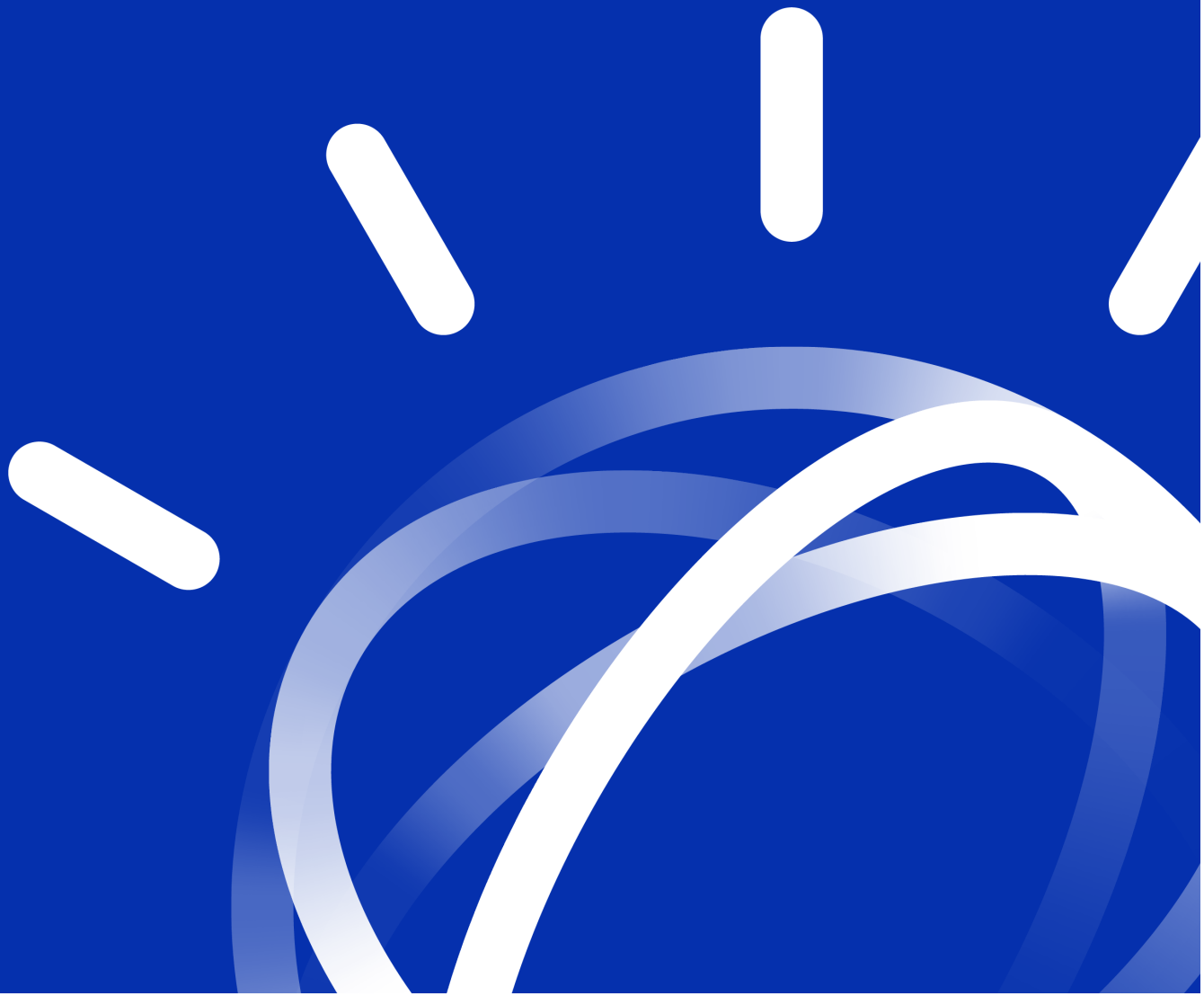
IBM Watson Studio
Overview

March 2018



Machine Learning is algorithm selection

Deep learning is neural network design

AI is systems architecture

Why are enterprises struggling to capture the value of AI?

Data

- Data resides in silos & difficult to access
- Unstructured and external data wasn't considered

Governance

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

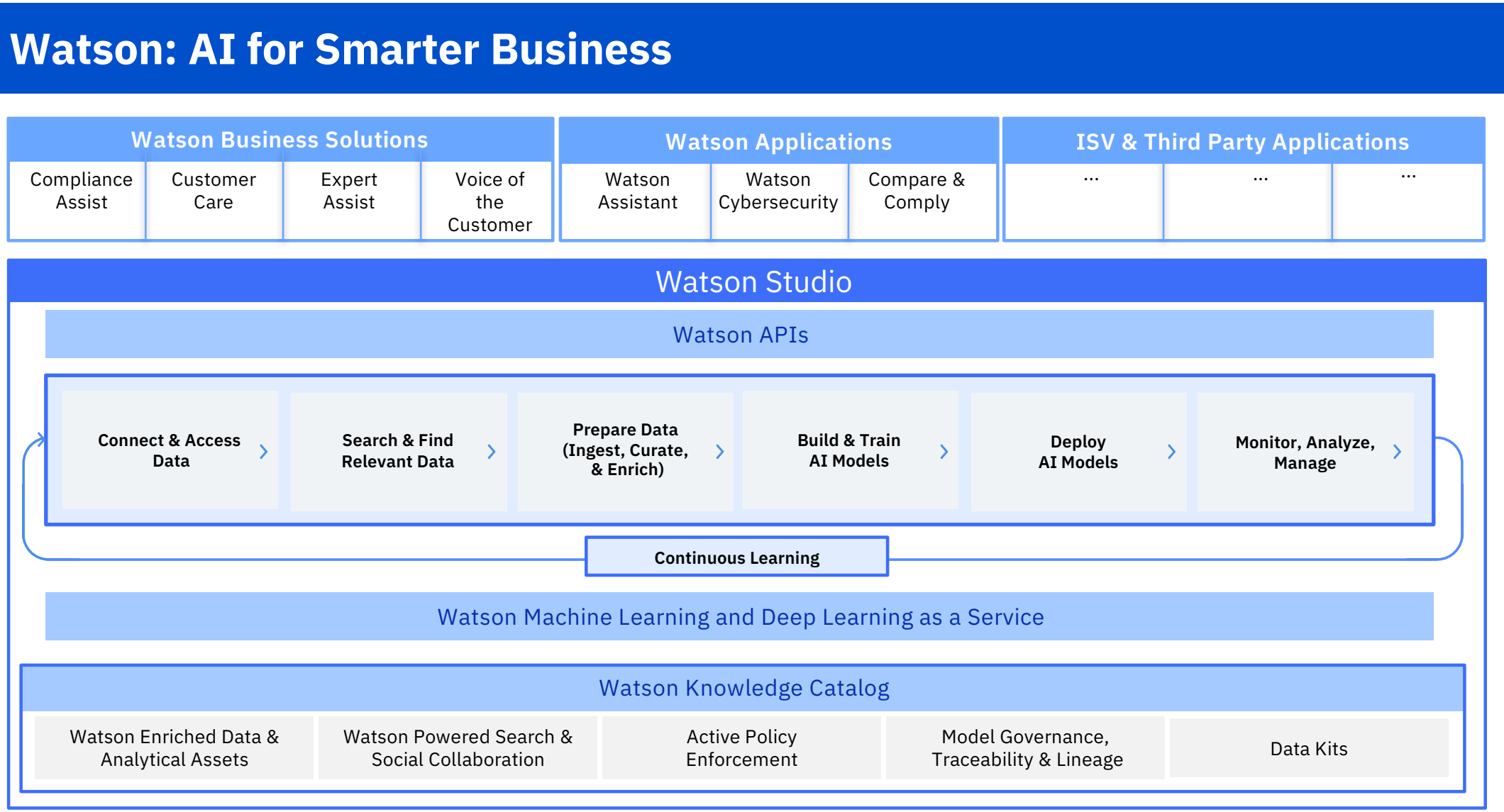
Skills

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

Tools & Infrastructure

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

Watson: AI for Smarter Business



Watson Studio

Watson Studio accelerates the machine and deep learning workflows required to infuse AI into your business to drive innovation. It provides a suite of tools for data scientists, application developers and subject matter experts to collaboratively and easily work with data and use that data to build, train and deploy models at scale.

AI Requires Teamwork

- AI is not magic
- AI is algorithms + data + team

Watson Studio

Built for AI teams – enabling team productivity and collaboration

Tanya
Domain Expert

Her Job:
To transfer knowledge to Watson for a successful user experience.

What she does:
- Range of domain knowledge and uses that to teach Watson and develop a custom models
- As Tanya gains more experience she optimizes her knowledge to teach Watson to design better end-user experiences.

Sometimes known as:
Subject matter expert, content strategist.

Mike
Data Scientist

His Job:
Transform data into knowledge for solving business problems.

What he does:
- Runs experiments to build custom models that solve business problems.
- Use techniques such as Machine Learning or Deep Learning and works with Tanya to validate success of trained models.

Sometimes known as:
ML/DL engineer, Modeler, Data Miner

Ed
Data Engineer

His Job:
Architects how data is organized and ensures operability

What he does:
- Builds data infrastructure and ETL pipelines. Works with Spark, Hadoop, and HDFS.
- Works with data scientist to transform research models into production quality systems.

Sometimes known as:
Data infrastructure engineer

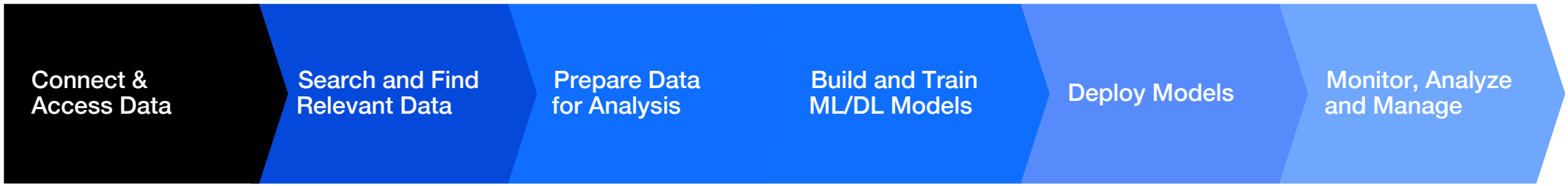
Deb
The Developer

Her Job:
Builds AI application that meet the requirements of the business.

What she does:
- Starts PoCs which includes gathering content, dialog building and model training
- Focus is on app building for the team or company to use. Will handle ML Ops as needed

Sometimes known as:
Front-end, back-end, full stack, mobile or low-code developer

Watson Studio
Supporting the end-to-end AI workflow



Connect and discover content from multiple data sources in the cloud or on premises. Bring **structured** and **unstructured** data to one toolkit.

Find data (structured, unstructured) and AI assets (e.g., ML/DL models, notebooks, Watson Data Kits) in the **Knowledge Catalog** with intelligent search and giving the right access to the right users.

Clean and prepare your data with **Data Refinery**, a tool to create data preparation pipelines visually.
Use popular open source libraries to prepare unstructured data.

Democratize the creation of ML and DL models. Design your AI models **programmatically** or **visually** with the most popular **open source** and IBM ML/DL frameworks or leverage transfer learning on **pre-trained** models using **Watson tools** to adapt to your business domain. Train at scale on **GPUs** and **distributed** compute

Deploy your models easily and have them **scale automatically** for online, batch or streaming use cases

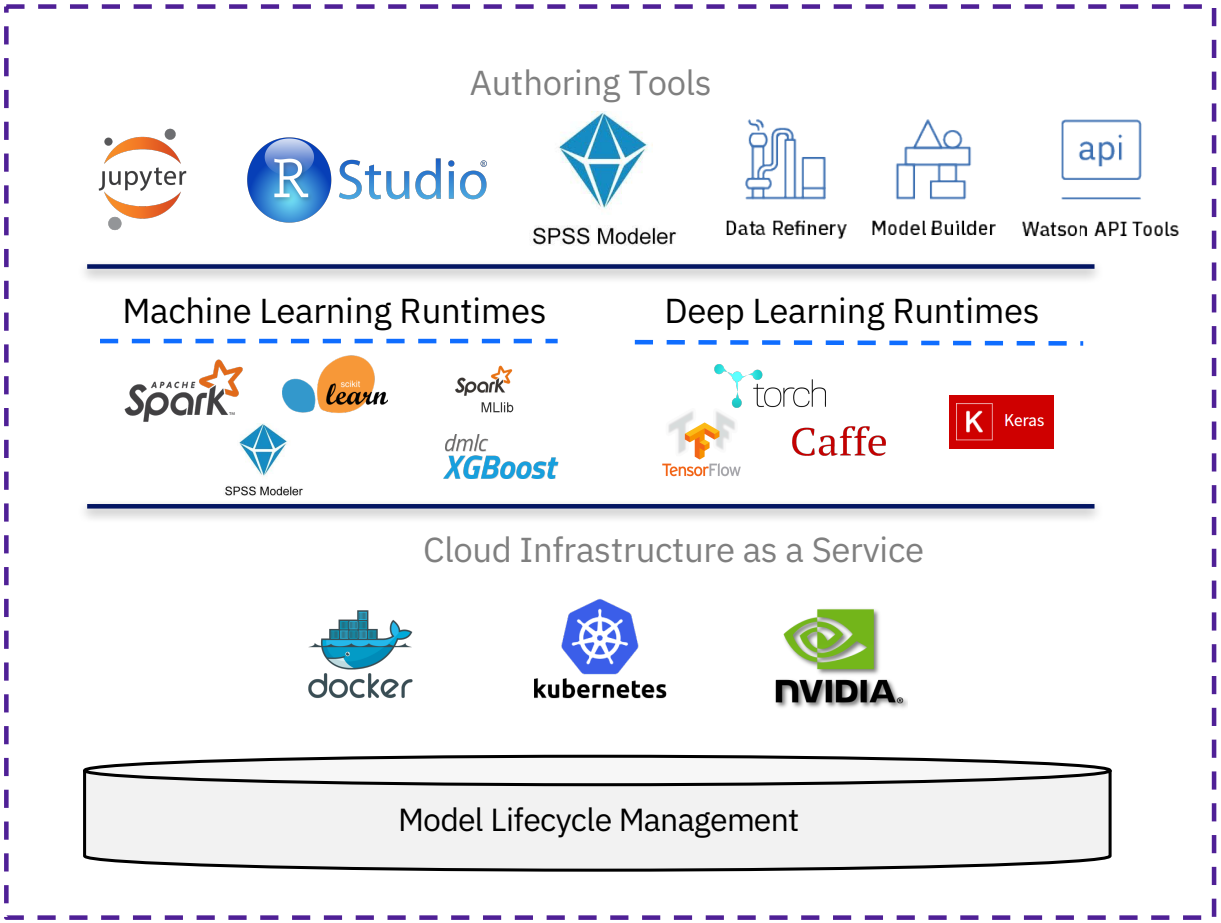
Monitor the performance of the models in production and trigger automatic retraining and redeployment of models. Build **Enterprise Trust** with Bias Detection, Mitigation Model **Robustness** and Testing Service Model **Security**.

Watson Studio
Tools for supporting the end-to-end AI workflow

- Create, collaborate, deploy, and monitor
- Best of breed open source & IBM tools
- Code (R, Python or Scala) and no-code/visual modeling tools

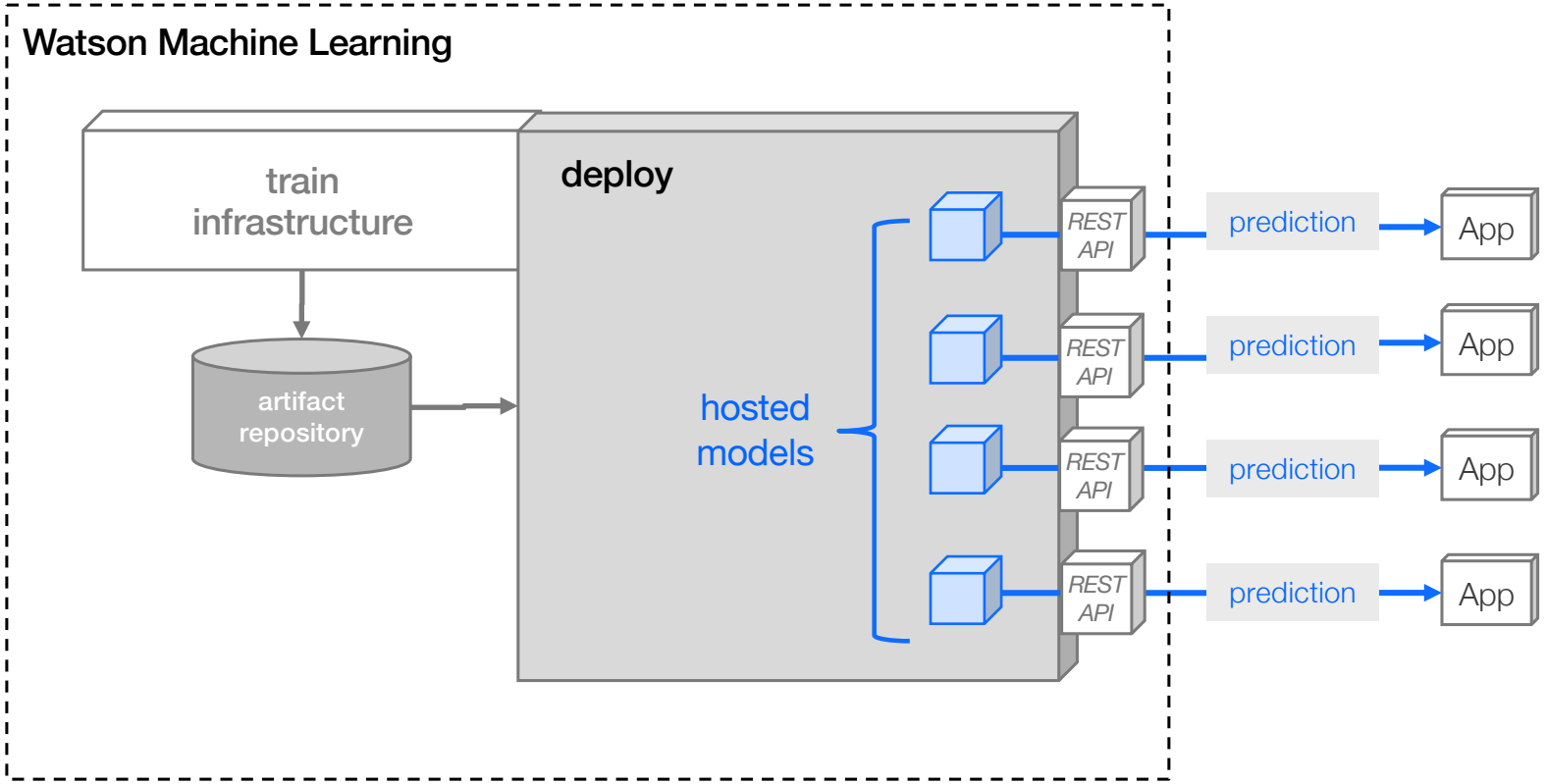
- Most popular open source frameworks
- IBM best-in-class frameworks

- Fully managed service
- Container-based resource management
- Elastic pay as you go cpu/gpu power



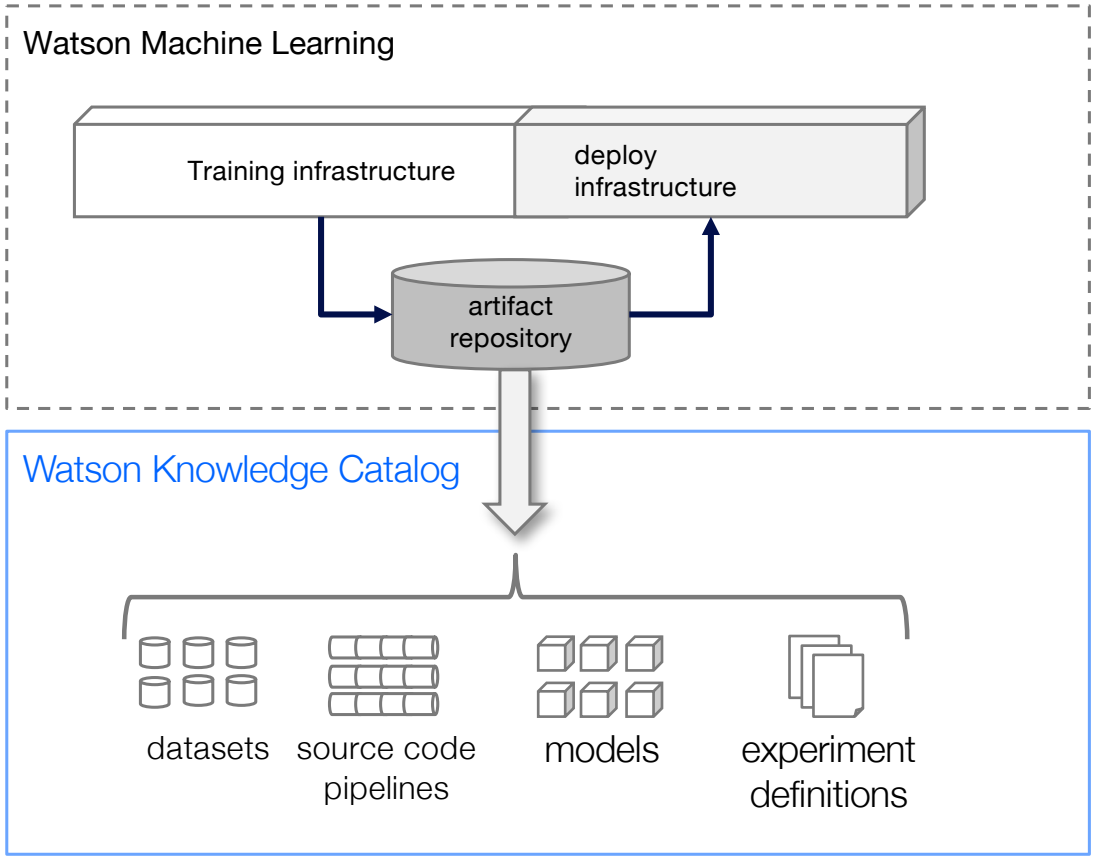
Deploying Trained Models

Download your trained models or deploy your models within Watson Machine Learning



AI Asset Lifecycle Management

Use the Watson Knowledge Catalog and Watson Studio to manage your AI assets or manage them yourself



Model Explanations
In May 2018, the General Data Protection Regulation (GDPR) takes effect and grants consumers the legal “right to explanation” from organizations that use algorithmic decision making.

Audit Trails
Tracking prediction to each model’s unique heritage is critical to regulatory compliance. Enforcing access controls for model sharing and deployment ensure ensures data security and application stability.

Watson Studio
Differentiating Capabilities

Integrated Collaboration Environment

- Data Scientists, Subject Matter experts, Business Analysts & Developers all in one environment to accelerate innovation, collaboration and productivity
- Built-in learning to get started or go the distance with advanced tutorials

Choice of Tools for the full AI lifecycle

- Best in-breed open source and IBM tools that support the end-to-end AI lifecycle
- Choice of code or no-code tools to build and train your own ML/DL models or easily train and customize pre-trained Watson APIs

Support for all levels of expertise

- Use Watson smarts and recommendations for the best algorithms to use given your data, OR
- Use the rich capabilities and controls to fine tune your models

Experiment centric DL workflow

- Monitor batch training experiments then compare cross-model performance without worrying about log transfers and scripts to visualize results.
- You focus on designing your neural networks. We'll manage and track your assets.

Model lifecycle & management

- Deploy models into production then monitor them to evaluate performance.
- Capture new data for continuous learning and retrain models so they continually adapt to changing conditions.

Integrated with Knowledge Catalog

- Intelligent discovery of data and AI assets that enables reuse & improves productivity
- Seamlessly integrated for productive use with Machine Learning and Data science
- Powerful governance tools to control and protect access to data

Watson Studio

Inject AI on your business

Making AI a **Team Sport**

Learn

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

- Community with free sample Notebooks, Models, Open Data Sets to import and start working right away.
- Learn with hundreds of tutorials and get informed about the latest techniques with articles
- Share and bookmark your favorite community assets

Create

The best of open source and IBM Watson tools to create state-of-the-art data products

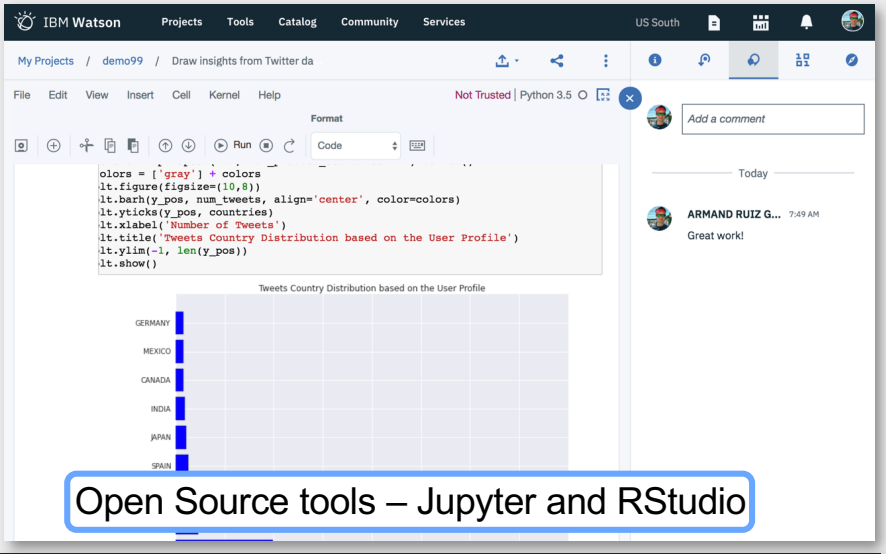
- Code in Python, R and Scala in Jupyter Notebooks or Rstudio.
- Access the most popular ML & DL frameworks.
- Create Models and design neural networks in minutes without coding with Visual Modeling tools.
- Start with Watson pre-trained models and customize them with your own data
- Create compute environments on demand and scale/customize them as needed.

Collaborate

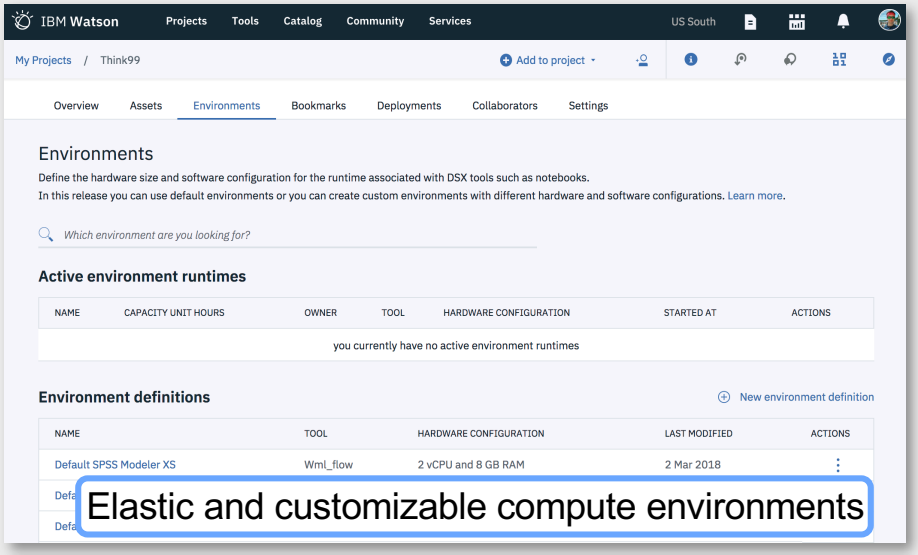
Community and social features that provide meaningful collaboration

- Create Projects and add your colleagues to it to work all together
- Control the permissions with Admin/Editor/Viewer access
- Add comments and track the activity of your colleagues in Projects
- Track the workflow from experimentation to production
- Version control of your assets
- Data Scientists, Subject Matter experts, Developers all in one environment to accelerate innovation and collaboration

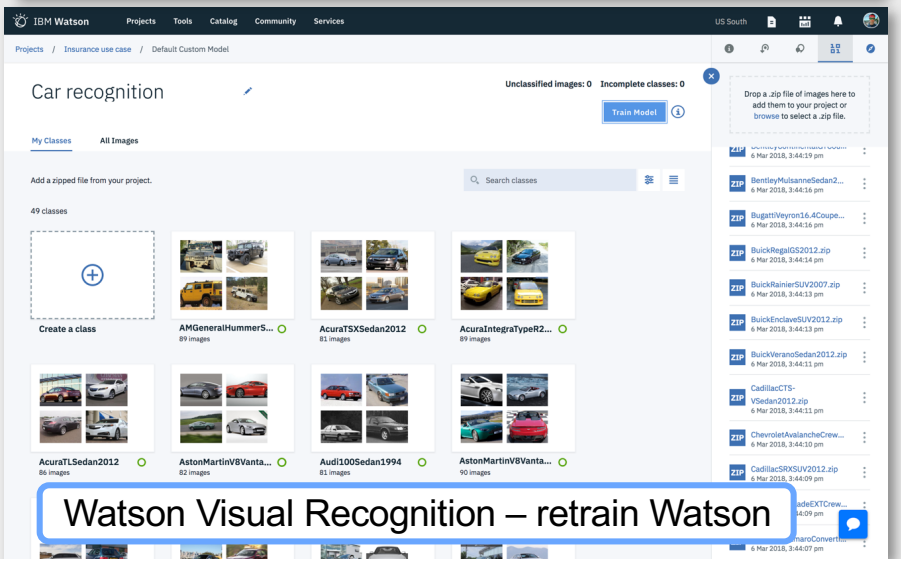
Watson Studio



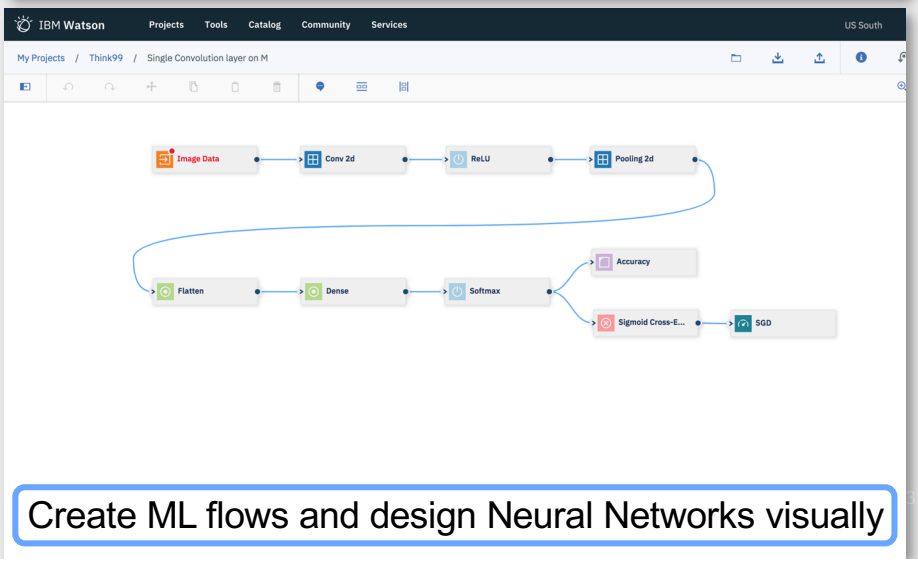
Open Source tools – Jupyter and RStudio



Elastic and customizable compute environments



Watson Visual Recognition – retrain Watson



Create ML flows and design Neural Networks visually

Watson Knowledge Catalog

Unlock tribal knowledge and unleash your knowledge workers

An **intelligent asset catalog** for a 360 degree view of your data & AI assets

Discover

Intelligent discovery of data and AI assets with advanced classification and profiling to provide context

- Intelligent data classification and profiling that determines what the data is and how it should be used
- Quickly build a 360 view of all assets and provide them for AI and Analytics
- Crawlers to auto discover usage information of data to understand how data is used

Catalog

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

- A business friendly shopping portal for your enterprise data
- Integrated with other platform solutions to facilitate self service analytics and AI
- Access controls and security
- Seamlessly integrated for productive use with Data prep, movement, dashboarding, Machine learning and Data science

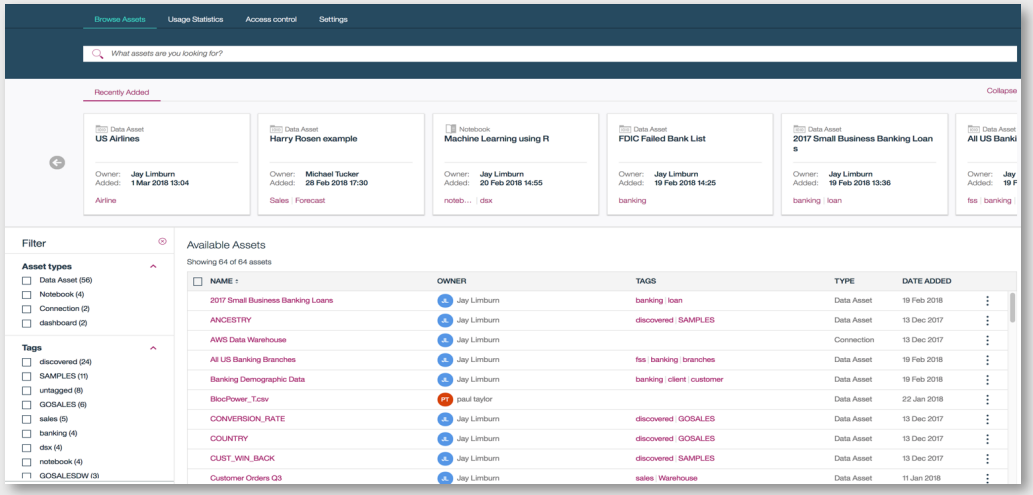
Activate

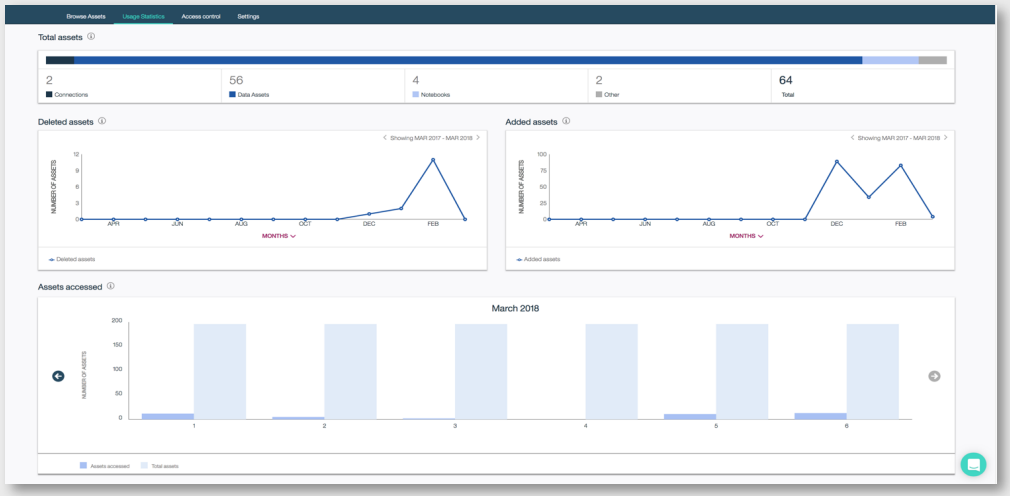
Powerful governance tools to control and protect access to data with visibility to data use

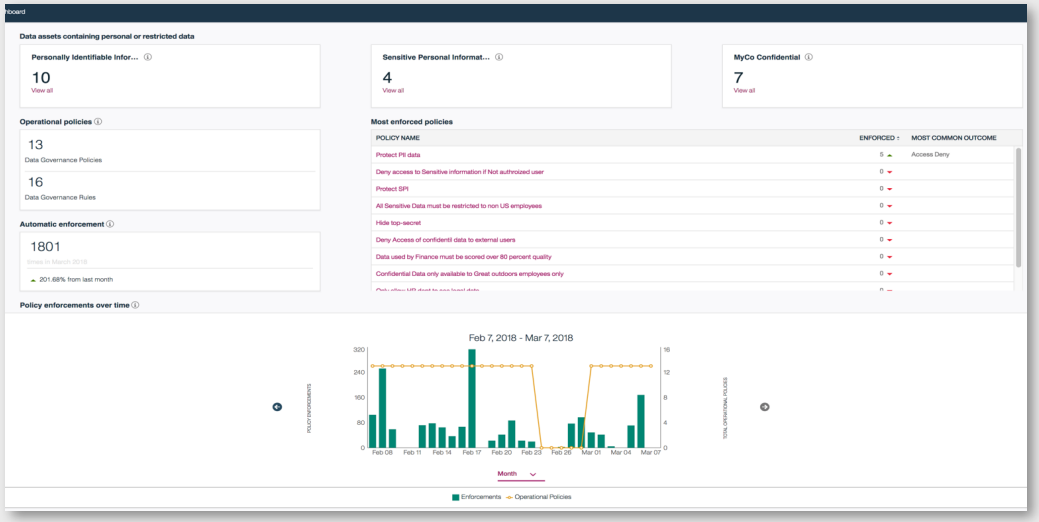
- Business Glossary to define business terms and map them to technical assets
- Policy Engine to author, activate and enforce business policies and rules
- Governance and Insights dashboards to understand how data is used and how the governance program is impacting it

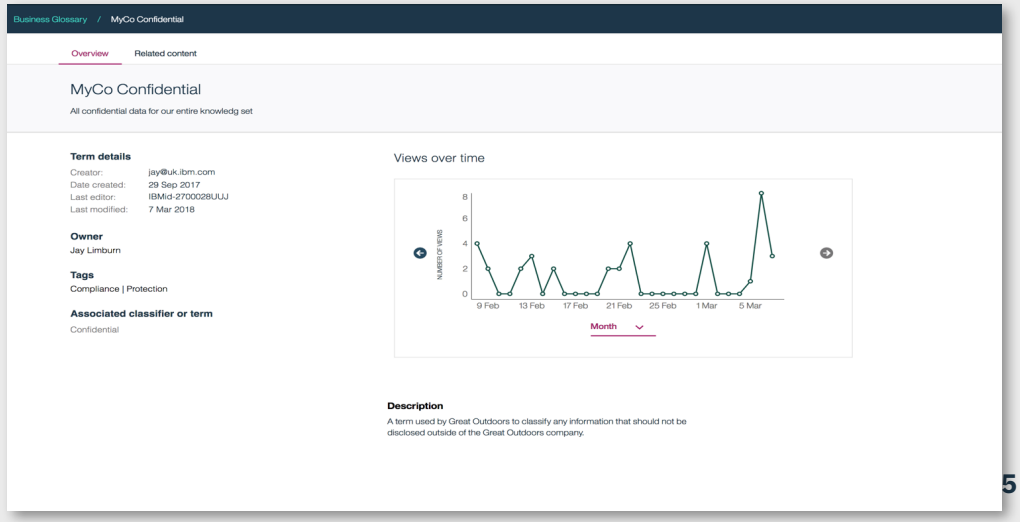
Watson Knowledge Catalog

Unlock tribal knowledge and unleash your knowledge workers









Data Refinery

Making data fit for use

Simplifying **Data Preparation** to Accelerate Data Insights

Connect

Access to your data regardless of where it resides

- Support for cloud and on-premises data
- 33+ connectors to IBM, non IBM and third party data sources
- Support for Files, Relational, No SQL and Hadoop data sources…
- Shared connections across all Watson Data Platform experiences
- Seamless integration with Data Catalog

Wrangle

Profile, shape and cleanse data and make it fit for use

- Intuitive self-service interface for data exploration, visualization and refinement
- Combine data from multiple sources
- Profile, classify and enrich data
- Robust set of data transformation functions for the non-technical and technical user
- No coding or technical ETL skills required

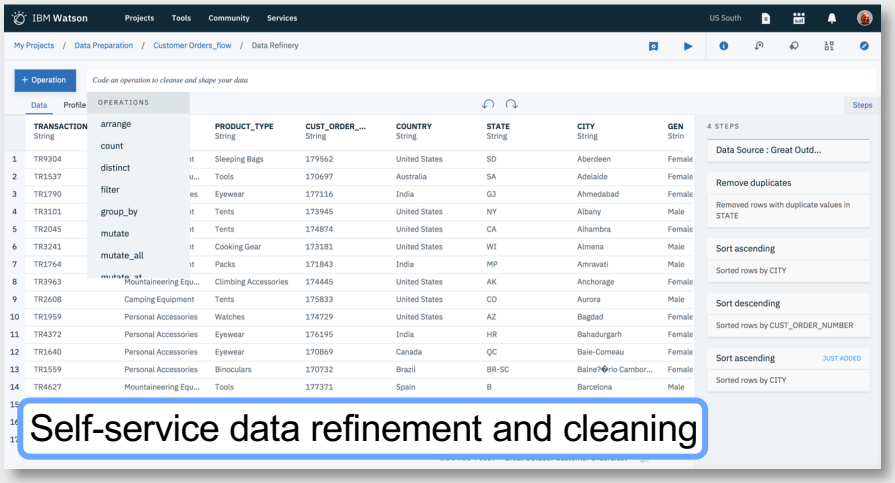
Deploy

Operationalize for repeatability, scalability and monitoring

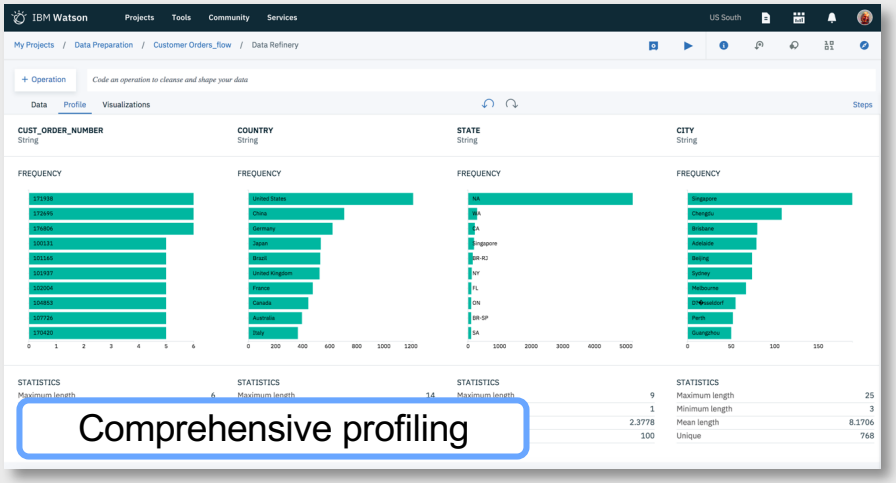
- Visual creation, modification and reuse of data flows
- Scheduling of data flow executions for repeatable outcomes
- Monitoring and notification of results
- Scale on demand to support any sized workload

Data Refinery

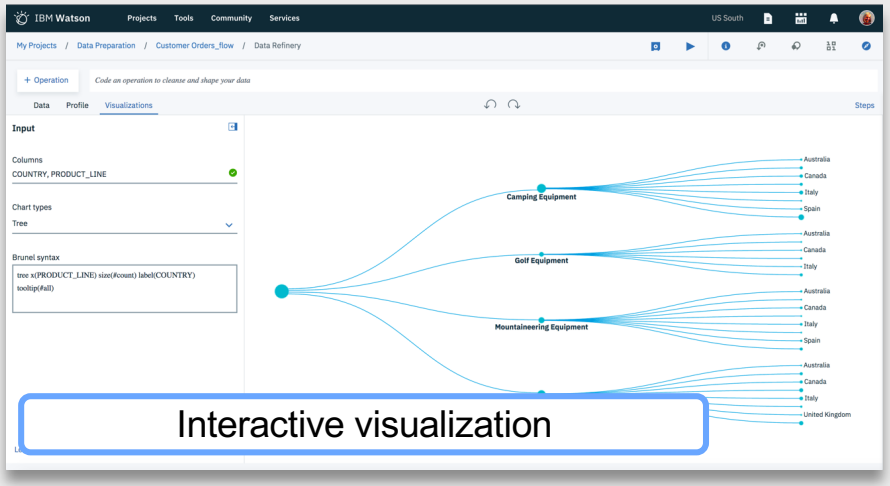
Making data fit for use



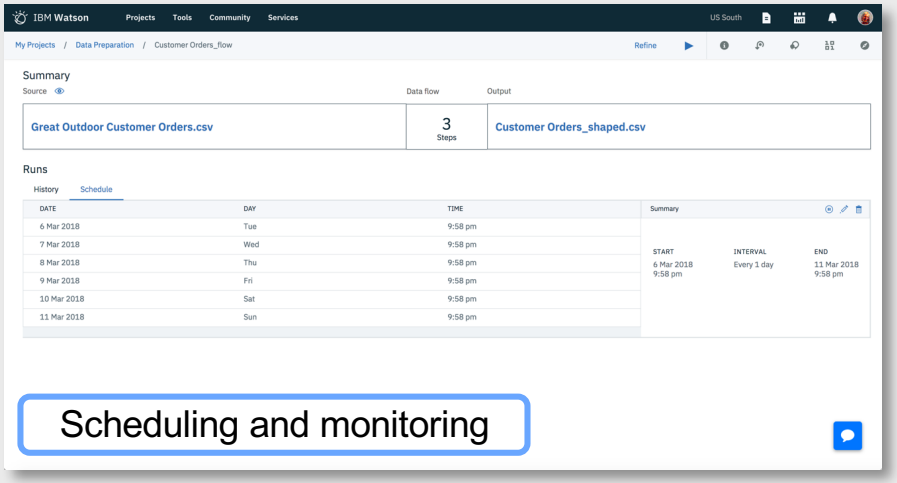
Self-service data refinement and cleaning



Comprehensive profiling



Interactive visualization



Scheduling and monitoring

Watson Machine Learning

Simplifying deployment & management of ML models in production apps

Embed AI in your Business

Train

Train custom ML and DL models at scale

- Use your favorite tools and frameworks such as SPSS, Tensorflow, Keras, Scikit Learn or xgboost.
- Track the progress and performance of your training jobs with the Experiment Assistant interface.
- Create new models without writing a single line of code using the neural network modeler

Deploy

Move models into production where they improve and remain relevant

- Portable models – deploy in the cloud, on devices or on premise
- Import models trained somewhere else and deploy in the cloud
- Transfer models to connected devices with support of Core ML, Tensorflow Lite
- Make millions of predictions in seconds

Automate

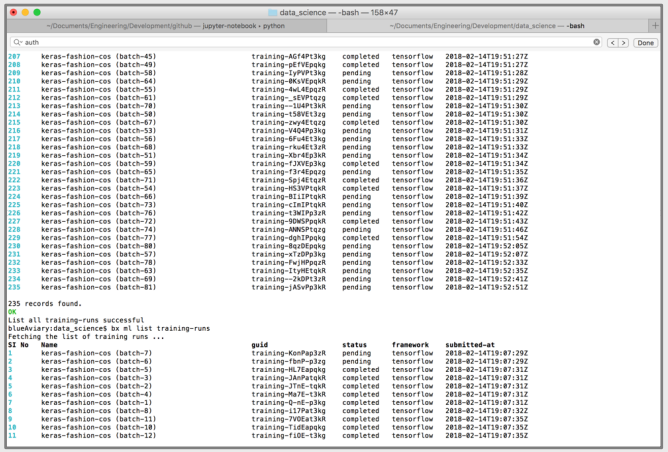
Build better models faster while they continue to learn after

- Build better models faster with automatic parameter tuning
- Automate the retraining of models with our feedback loop capabilities
- Automate the deployment of models in products
- Automatic algorithm selection for ML models with our CADS technology from Research

Watson Machine Learning

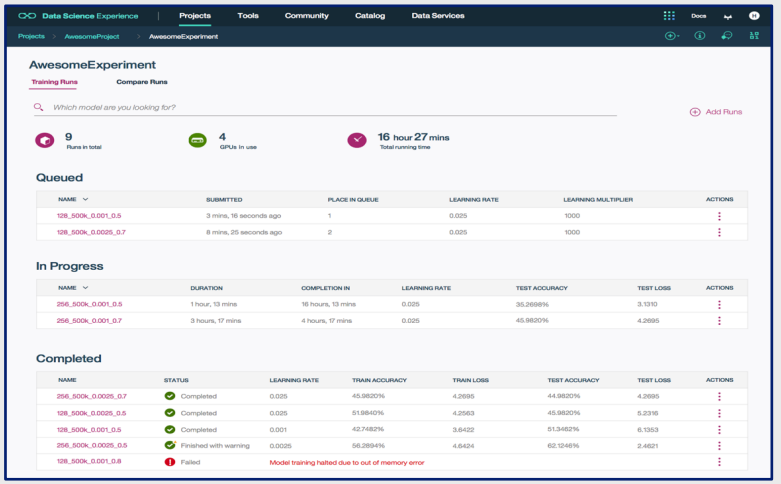
Simplifying deployment & management of ML models in production apps

Python client, command line interface (CLI) or UI? You choose the tooling that best fits your existing workflows. Training history and assets are tracked then automatically transferred to the customer's Object Storage for quick access.

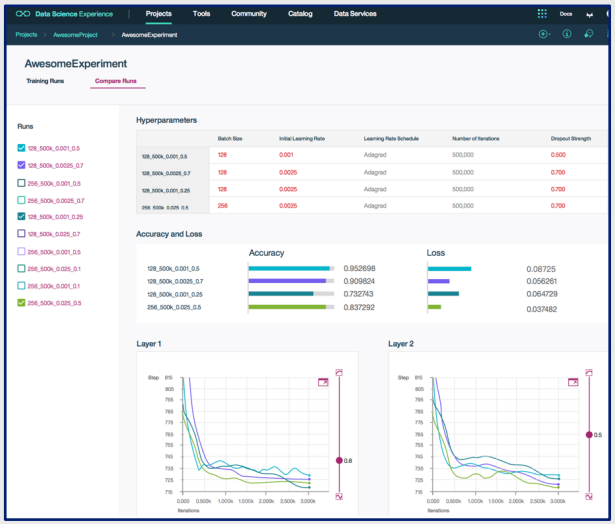


Train neural networks in parallel across NVIDIA GPUs.

Pay only for what you use. Auto-deallocation means no more remembering to shutdown your cloud training instances.

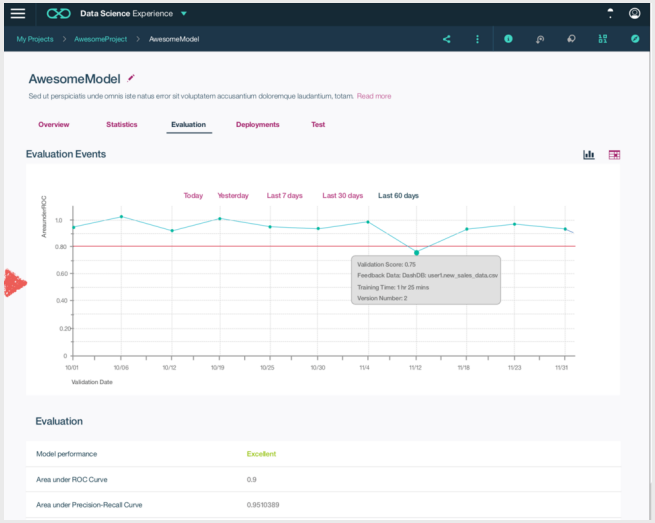


Monitor batch training experiments then compare cross-model performance without worrying about log transfers and scripts to visualize results. You focus on designing your neural networks. We'll manage and track your assets.



Deploy models into production then monitor them to evaluate performance.

Capture new data for continuous learning and retrain models so they continually adapt to changing conditions.



IBM Analytics Engine

A **flexible framework to deploy and manage** Hadoop/Spark at scale

Open

Open source stack based on industry leading Hadoop/Spark distribution

- Standard open source APIs to build and deploy portable applications on Spark and Hadoop components
- Integrate with other Hadoop and Spark ecosystem frameworks for advanced analytics use cases

Scalable

Architecture based on separate compute and storage for enhanced scalability, reliability and cost efficiency

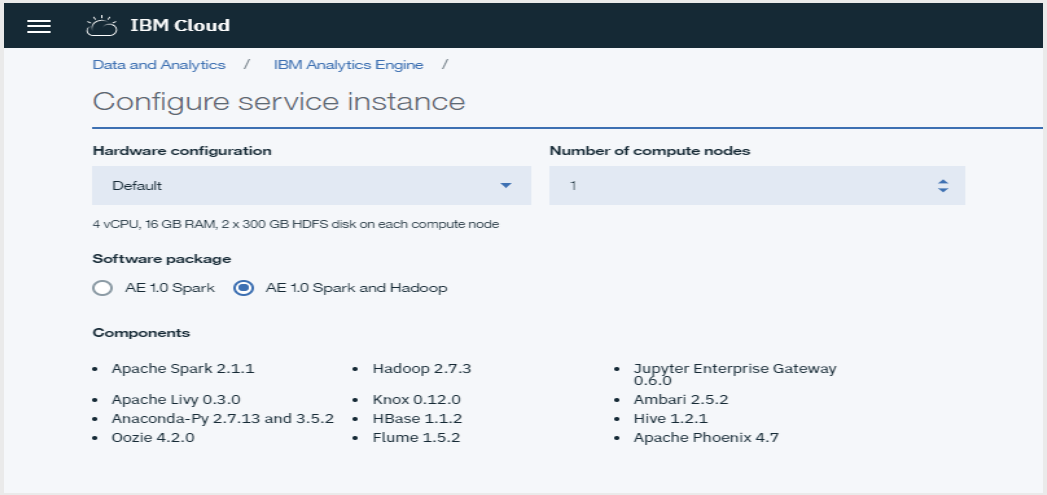
- Persist data in secure, scalable IBM Cloud Object Storage with options to replicate across regions
- Scale compute and storage independently for better economics
- Separate compute and storage ensure no data-loss in cases of cluster failure

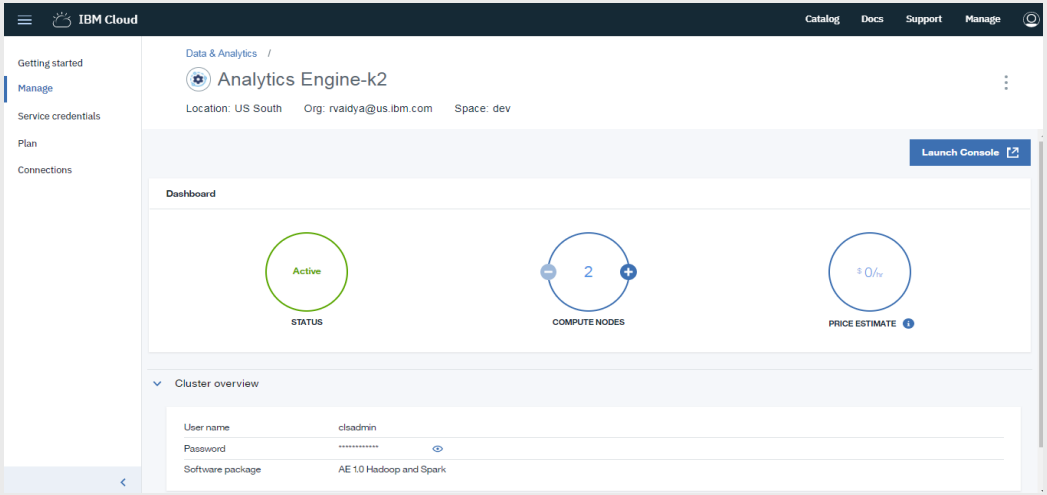
Flexible

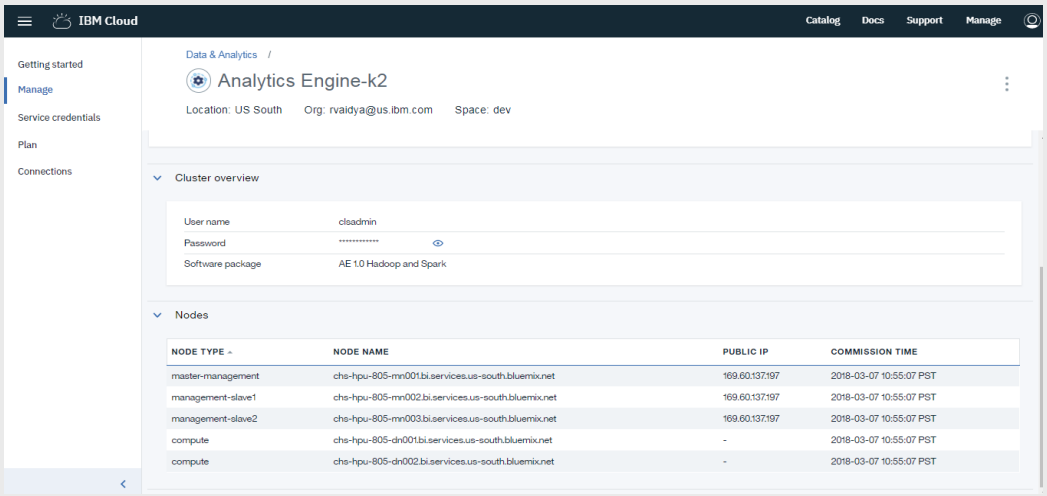
Options to deploy analytics application and operationalize administration

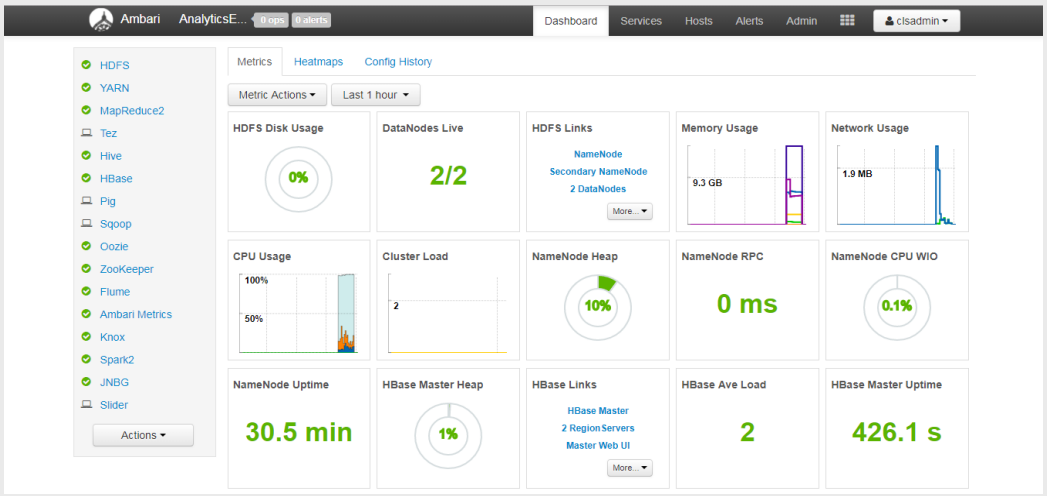
- Spin up, within minutes, use case specific clusters using different instance sizes for different use cases
- Access and administer through multiple interfaces – Cloud Foundry CLI, REST APIs on public interface, and GUI
- Enhanced flexibility for configuring and clusters, including installing 3rd party libraries through bootstrap scripts

IBM Analytics Engine









Dynamic Dashboards

Making insights available to all

Simplifying **Data Exploration** to initiate analysis and share results

Visualize

Represent analytic results as compelling interactive graphics

- Immediately view effective visualizations fitting the data selected.
- Choose your favorite graph or chart type from palette of options
- Apply Filters to individual or multiple visualizations.
- Customize appearance with individual settings or dashboard themes.

Share

Deliver dashboards across the enterprise, providing insights where business decisions are made.

- Collaborate on dashboards with other project users.
- Share interactive dashboards through a public URL for use by others outside the platform.

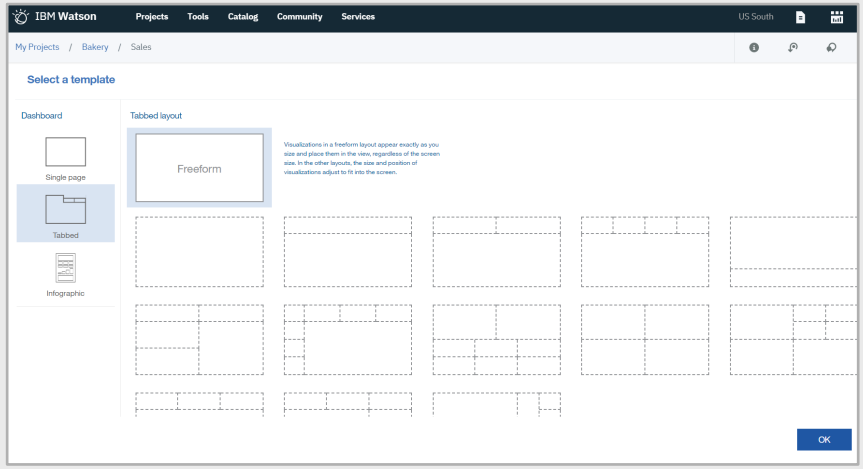
Discover

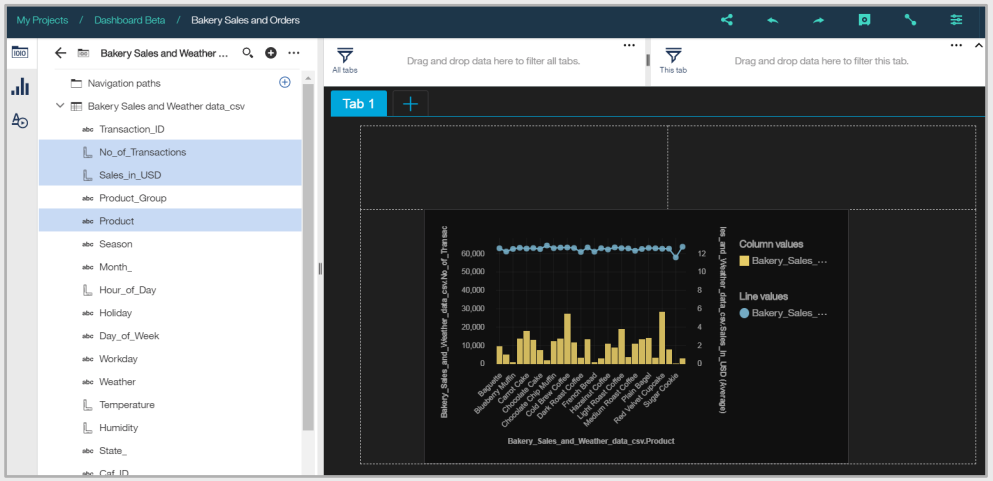
Utilize enterprise data assets with confidence through projects and the Watson Knowledge Catalog.

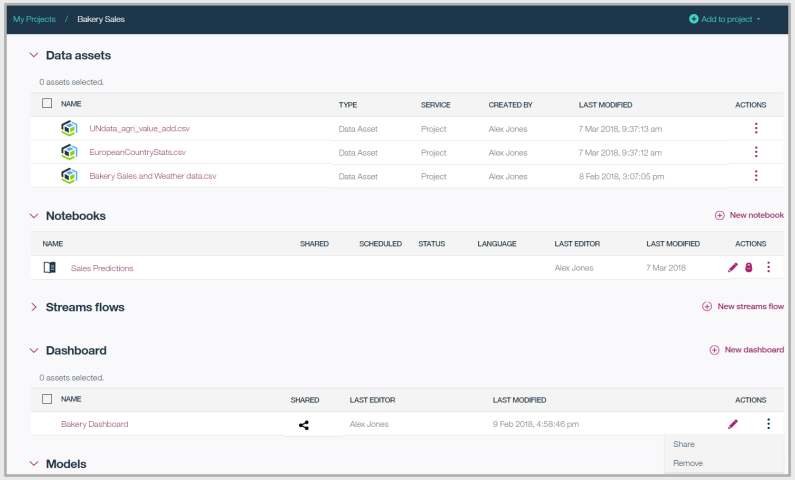
- Access data assets associated with your project
- Data sources may be local uploaded files or connections to relational sources.
- Search for data through the data catalog, then explore using dashboards.

Dynamic Dashboards

Making insights available to all







IBM **Watson**®

Get started today www.ibm.com/watson