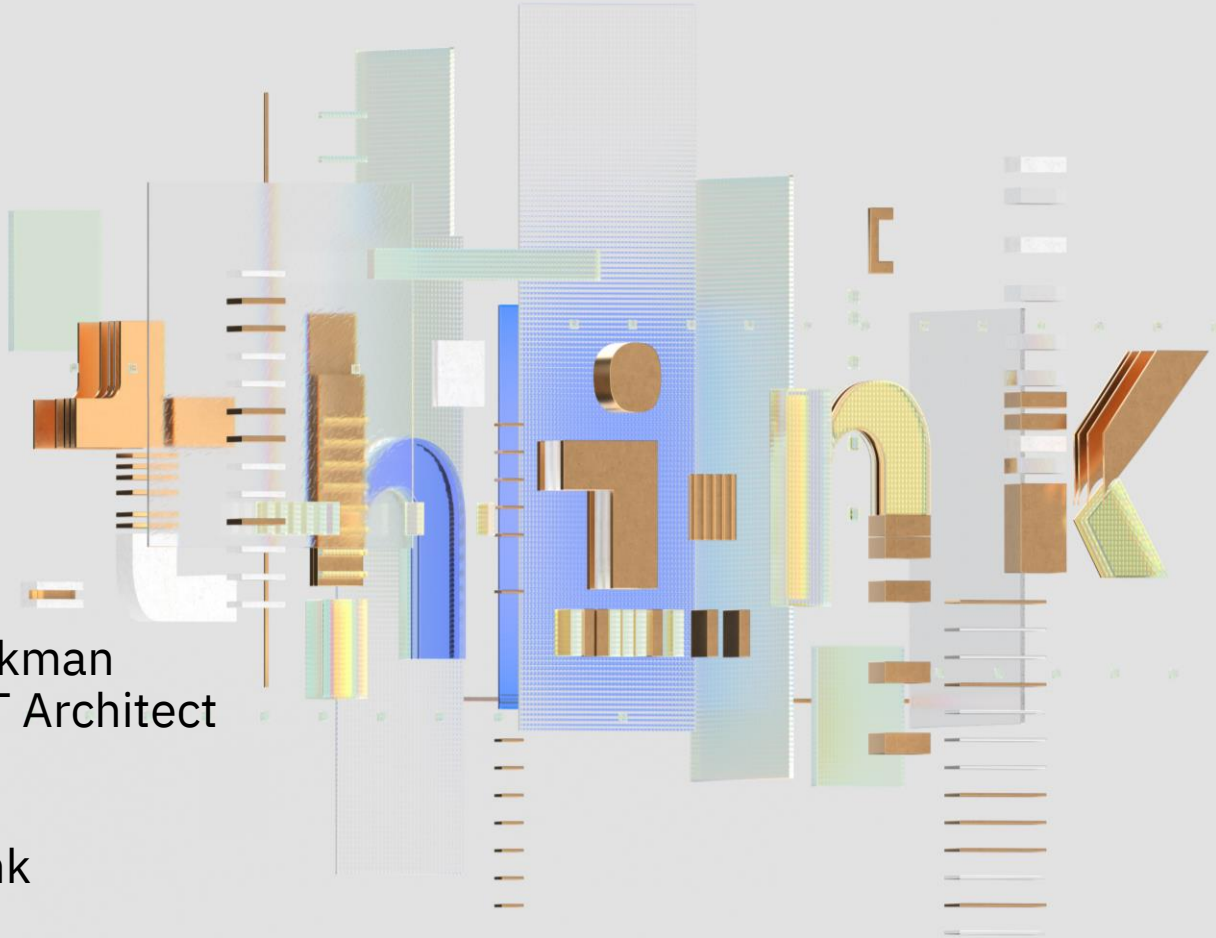
Introduction to Watson Studio



Bernard Beekman
Executive I/T Architect
US Federal

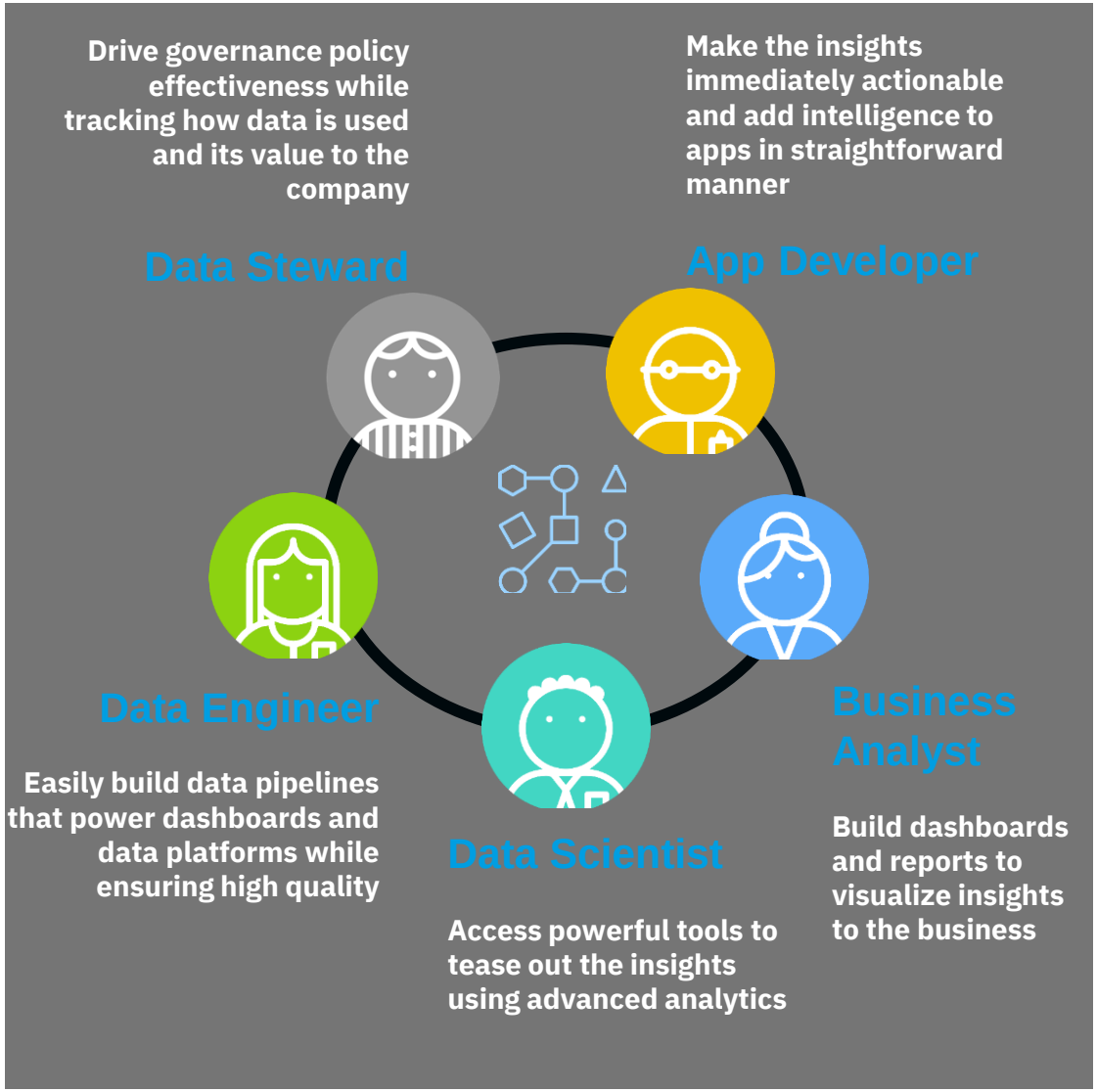
Michael Cronk
I/T Architect
US Federal

IBM takes an Enterprise Approach to Machine Learning

- Freedom of Choice
  - Choose programming languages, open source libraries, IBM value-add capabilities
  - Code/Click
  - Machine Learning/Deep Learning/Decision Optimization.
- Operationalize Machine Learning
  - Manage complete ML lifecycle – Build, Deploy, Manage, Scale, Monitor, Re-train
- Hybrid ML
  - Build where you want, deploy where you want
- Governance
  - Ensure that right people get access to the right data

IBM Watson Studio Platform

An integrated platform of tools, services, data, and metadata that help companies or agencies accelerate their shift to be data-driven organizations.



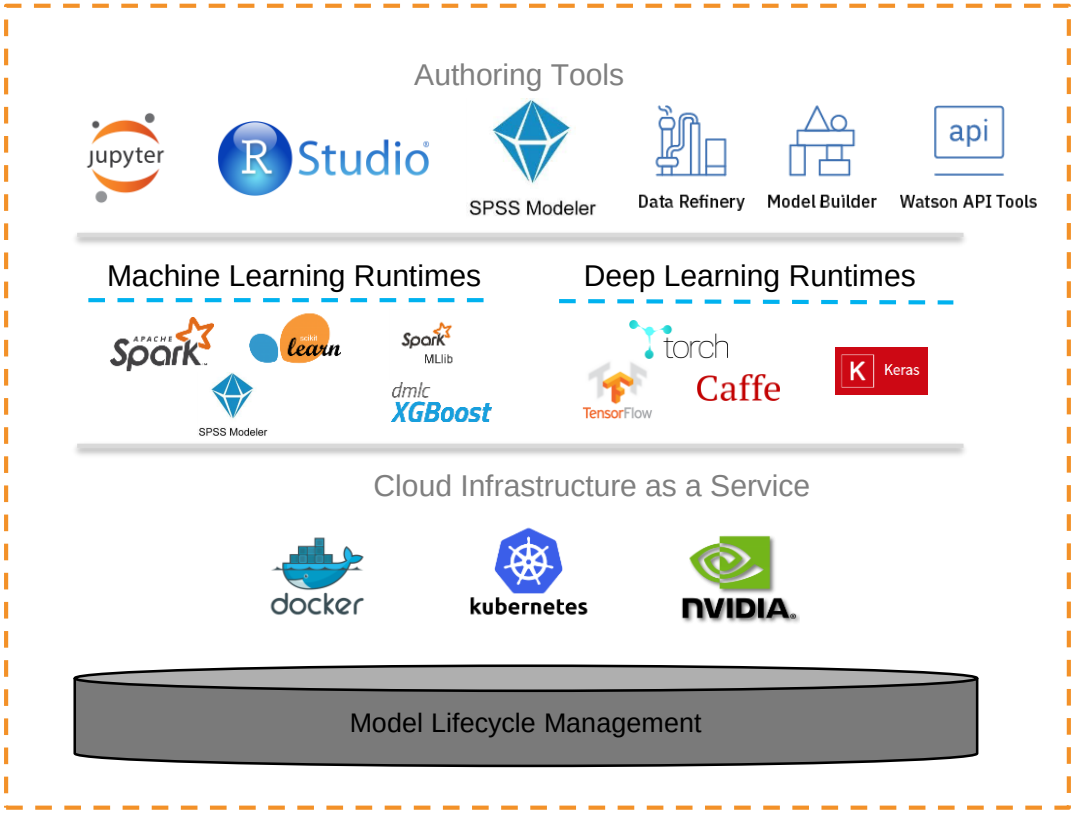
Watson Studio Deployment Options

- Watson Studio on IBM Cloud
  - Managed offering provided by IBM
- Watson Studio Local
  - On-premise – Private Cloud
  - IBM Cloud, AWS, Azure
- Watson Studio Desktop
- IBM Cloud Private for Data

Watson Studio Tools

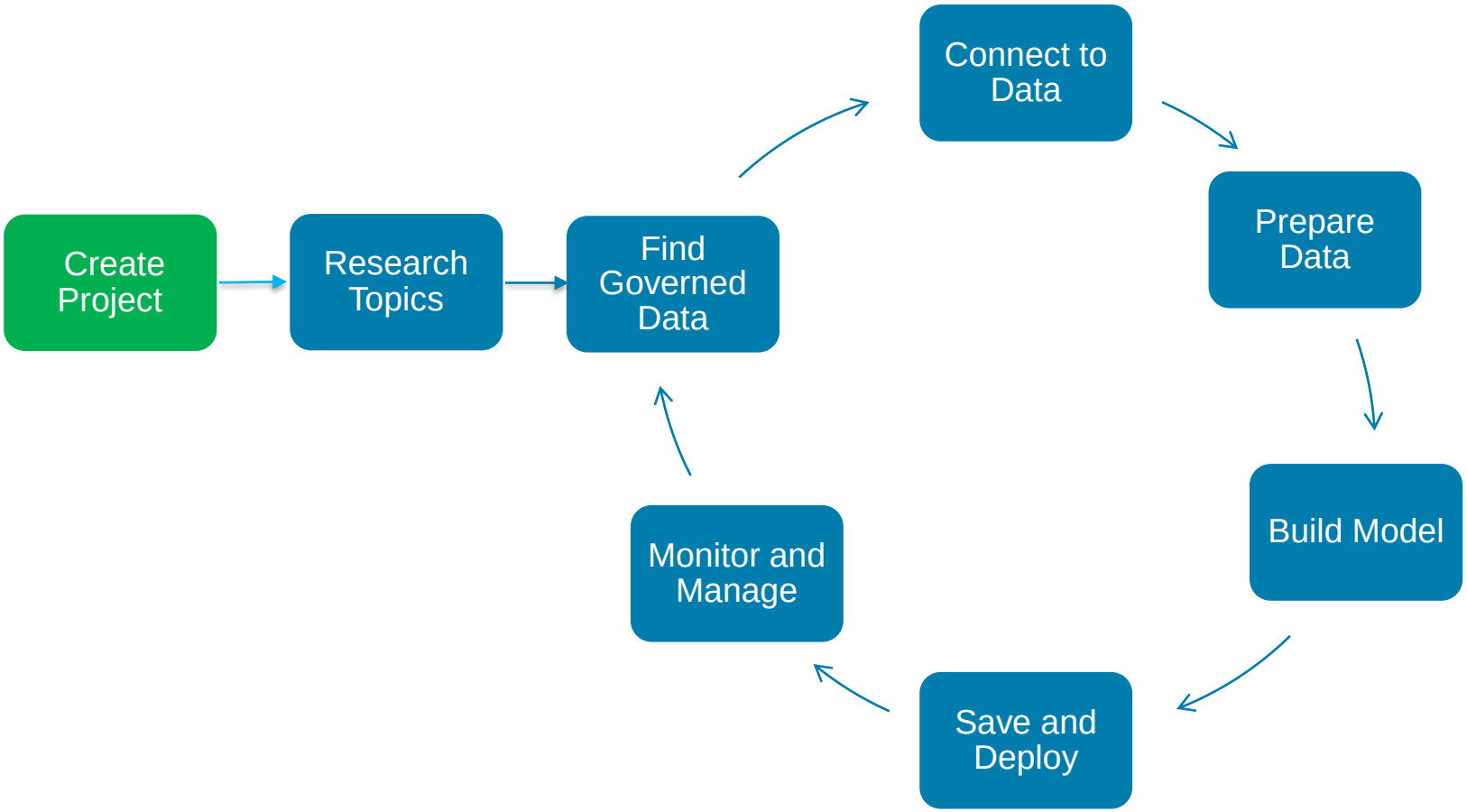
- Create, collaborate, deploy, and monitor
- Best of breed open source & IBM tools
- Code (R, Python or Scala) and no-code/visual modeling tools
- Open Source and IBM libraries/frameworks

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



Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



Watson Studio Project Features

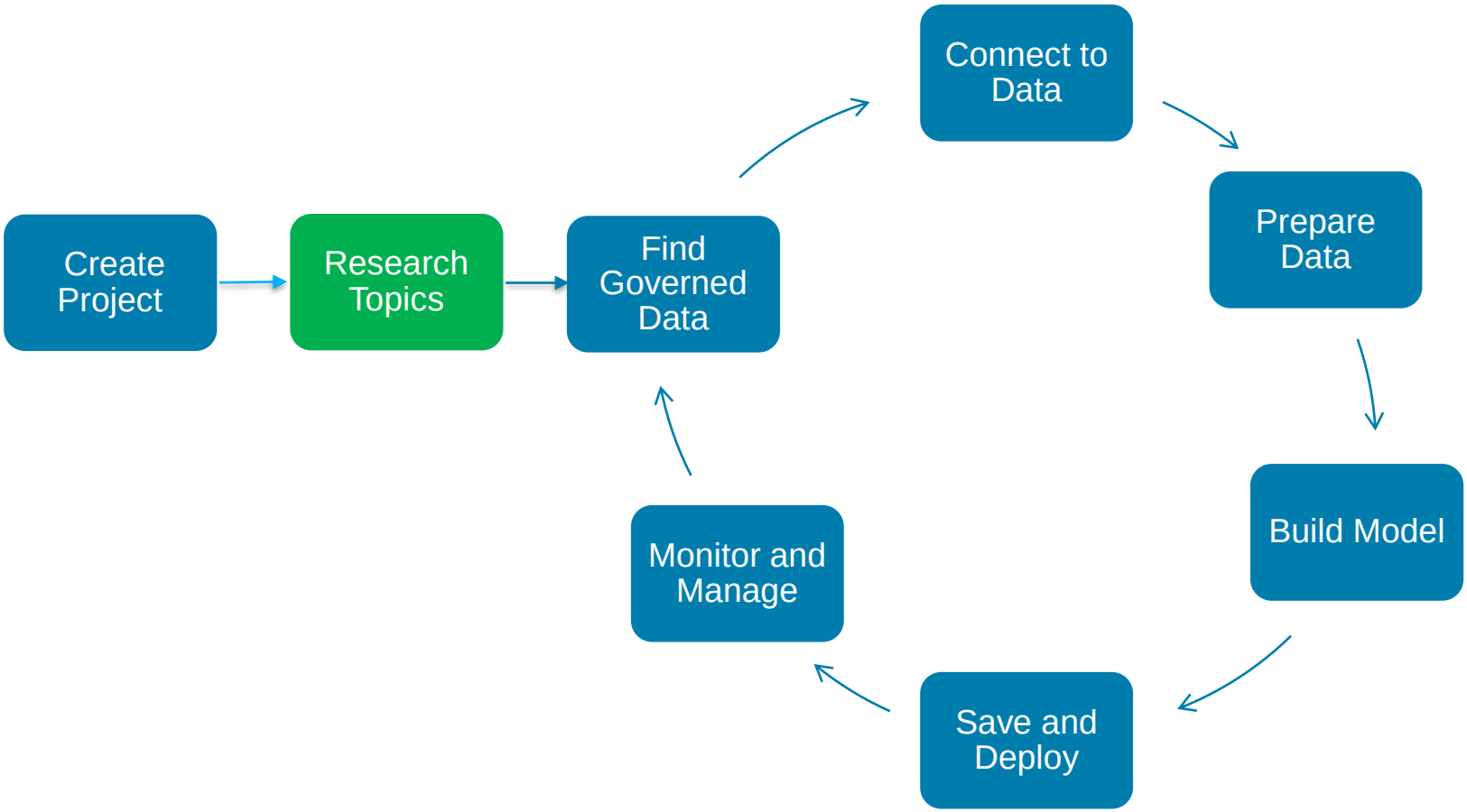
Making Data Science a Team Sport

Create Project

- Organizes resources to achieve a particular data analysis goal
- Support role-based collaboration (Admin, Editor, Viewer)
- Assets from all IDEs can be included in one Watson Studio project: notebooks, files, data sources, flows, models, etc.

Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



Watson Studio Community Card Features

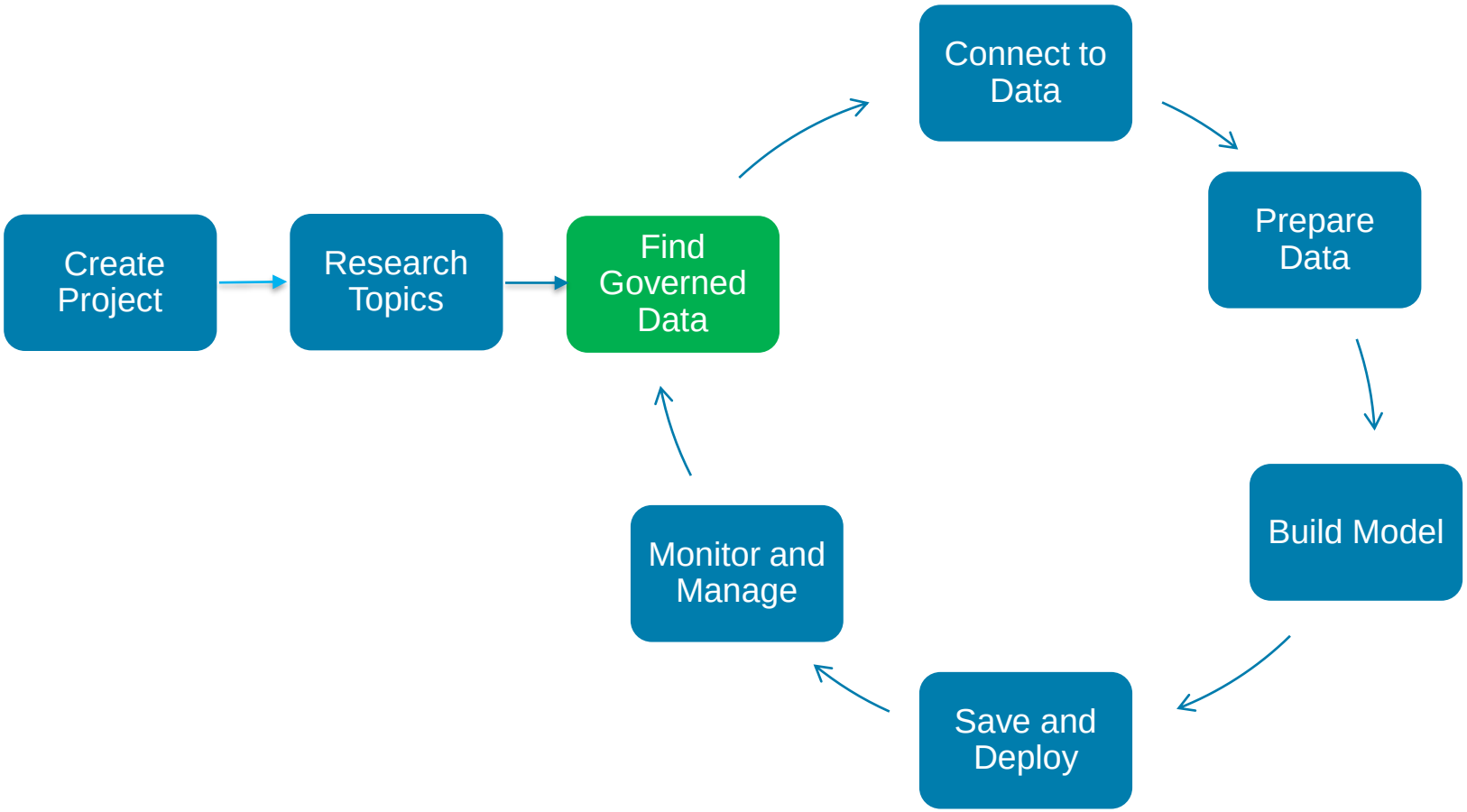
Built-in learning to get started

Research Topics

- Community Card Feature includes curated articles, tutorials, notebooks, data sets, and papers
- Bookmark in Projects
- Copy notebooks or Data Sets into projects
- Continuously updated

Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



Watson Knowledge Catalog Features

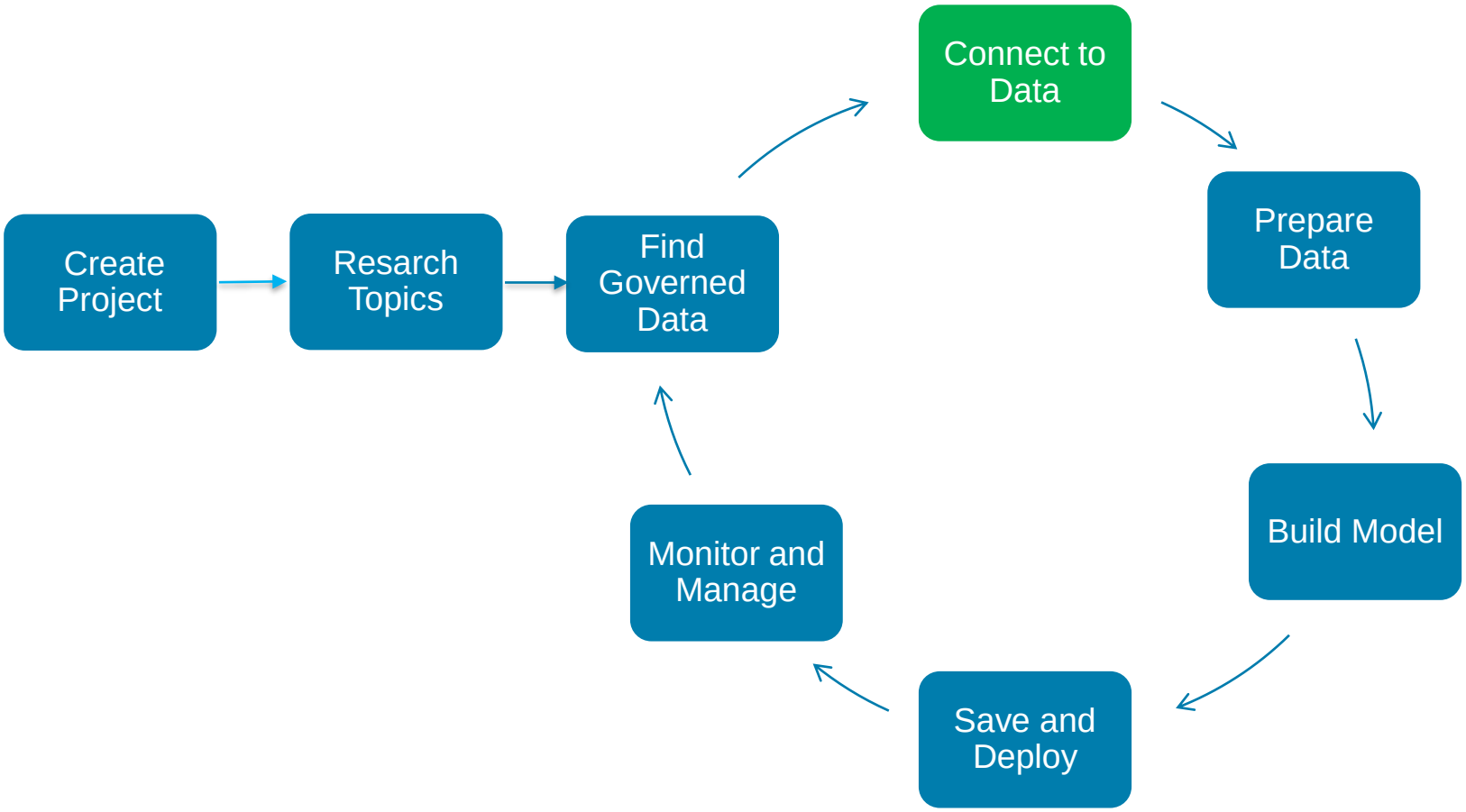
Unlock tribal knowledge and unleash knowledge workers

Find Governed Data

- **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.
- Discover assets, profiling, classification
- Policy, rule authoring
- Policy, rule enforcement
- Asset Usage Statistics

Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



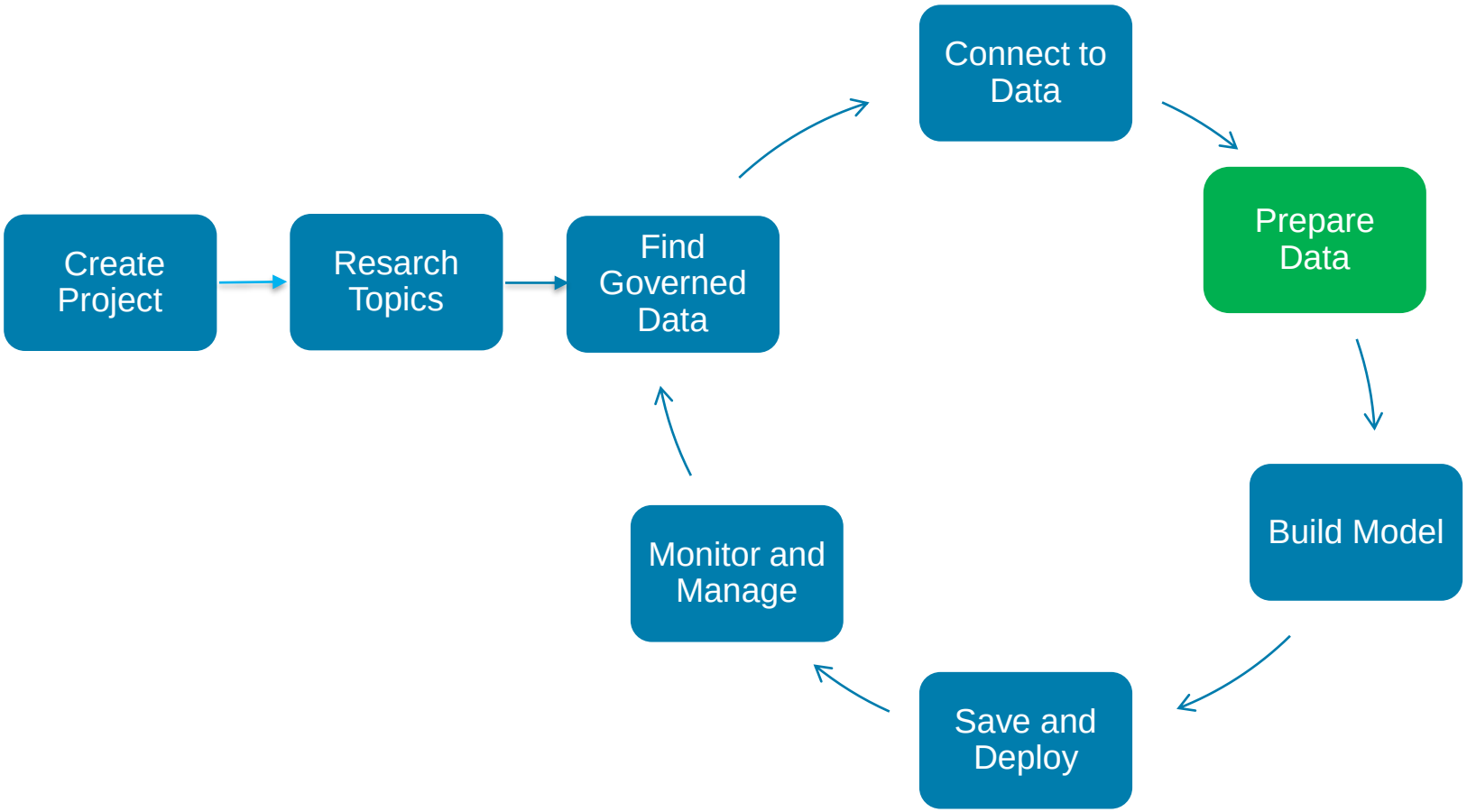
Watson Studio Connection Features

Connect to Data

- Upload files to Cloud Object Storage
- 39 Connectors to Structured and Unstructured, On-prem and Cloud data sources. ,
- Wizard based connection definition and code generation

Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



Watson Studio Data Refinery Features

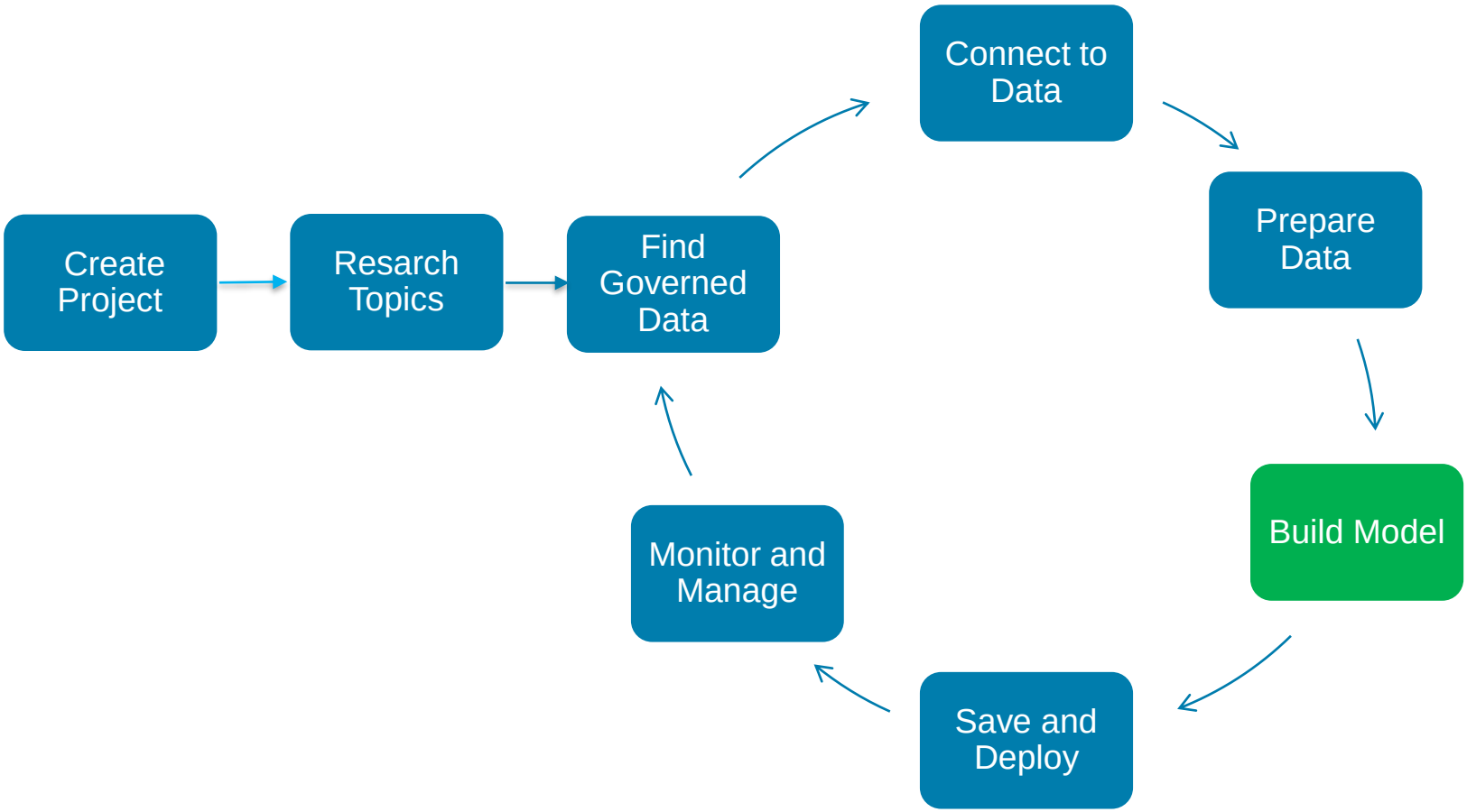
Making Data fit for use

Prepare Data

- Data Refinery tool to profile, visualize, and shape data.
- Create data preparation pipelines via point and click capability on subset of data
- Run the pipeline on all the data
  - Manually (on demand)
  - Automated (scheduled)

Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



Watson Studio Model Building Features

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

Build Model

Open Source Tools

- Jupyter Notebooks**
- RStudio and Shiny
- Libraries- scikit-learn**, XGBoost**, Spark**, TensorFlow**, Caffe, Keras, PyTorch

IBM Tools

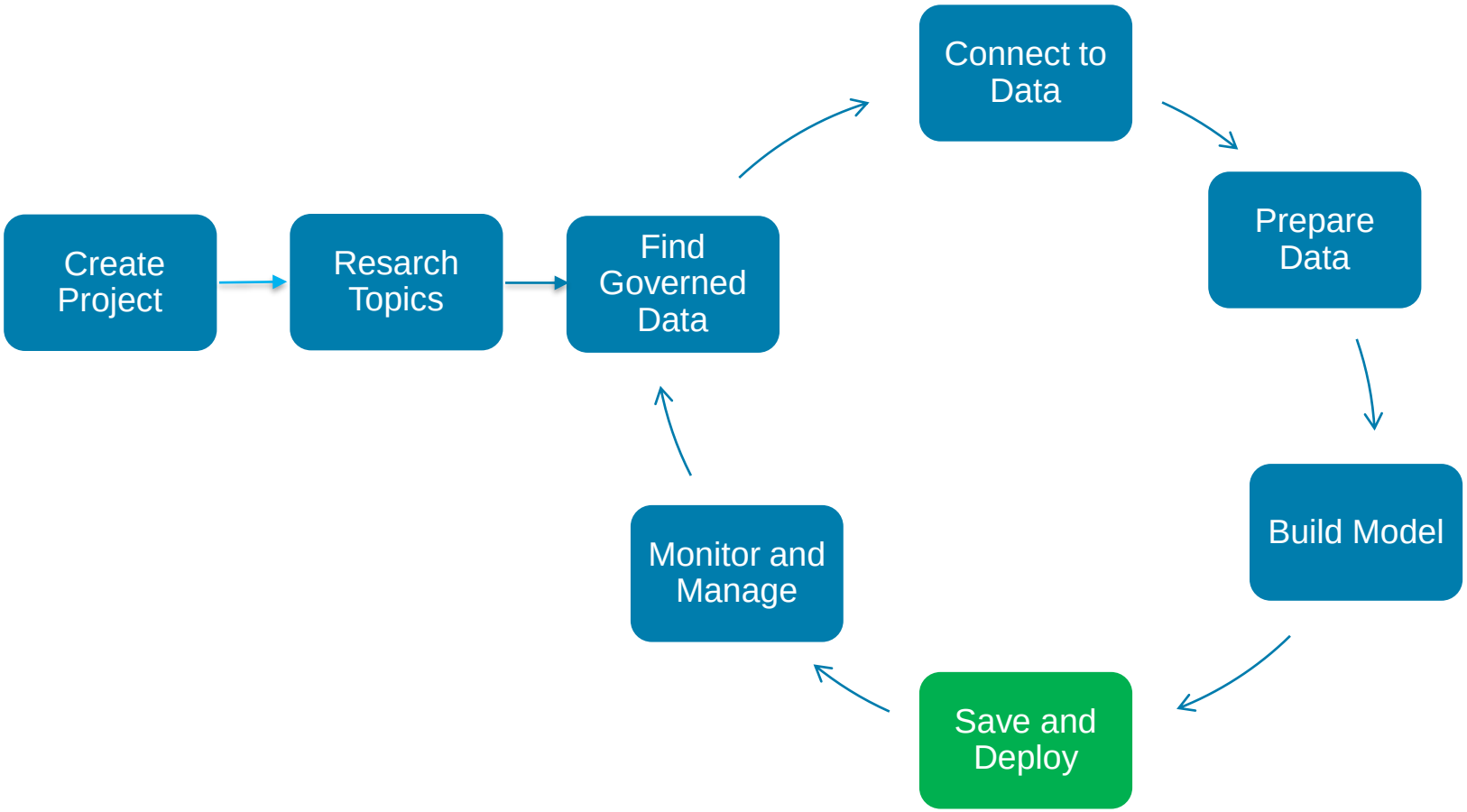
- Model Builder **
- SPSS Modeler
- Neural Network Modeler**
- Experiment Builder**
- Natural Language Classifier Model
- Visual Recognition Model

Train at scale on **GPUs** and **distributed** compute

** in hands-on labs

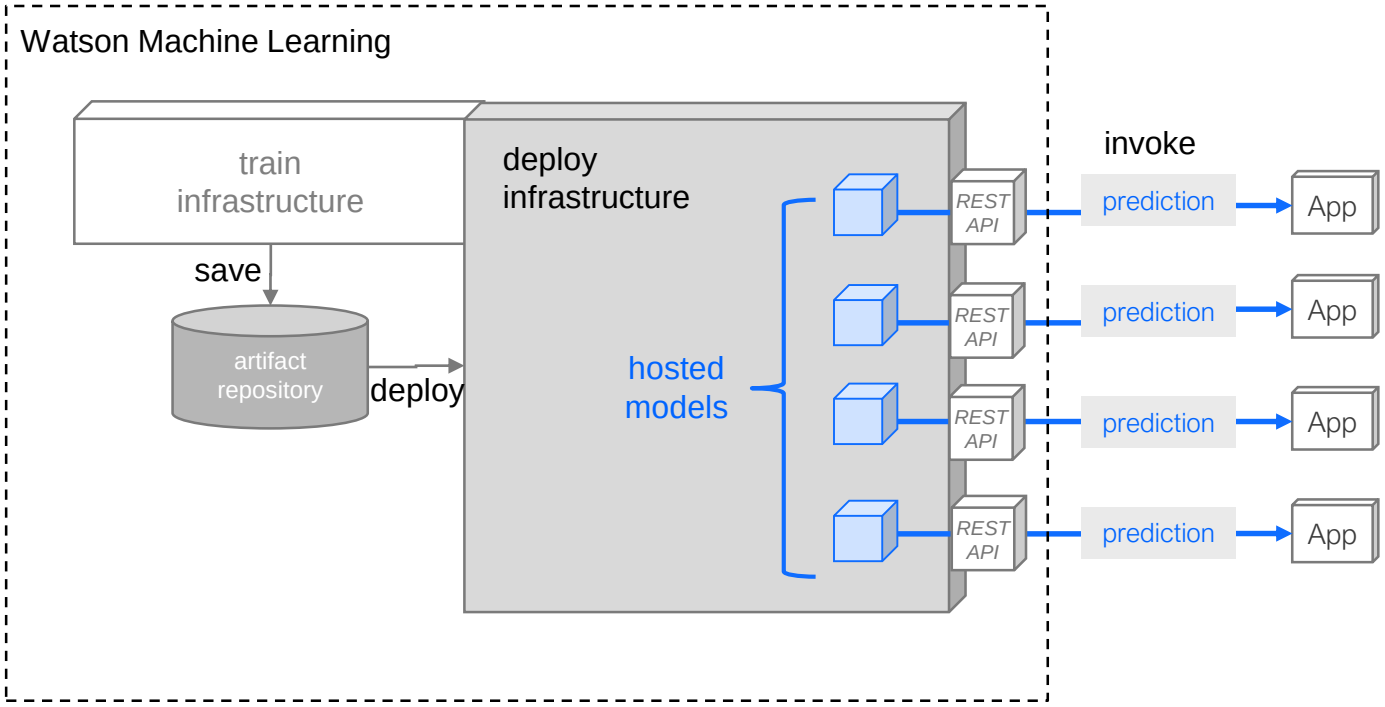
Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



Watson Studio Save and Deploy Trained Models

Save and Deploy Models with Watson Machine Learning



Watson Studio Save and Deploy Features

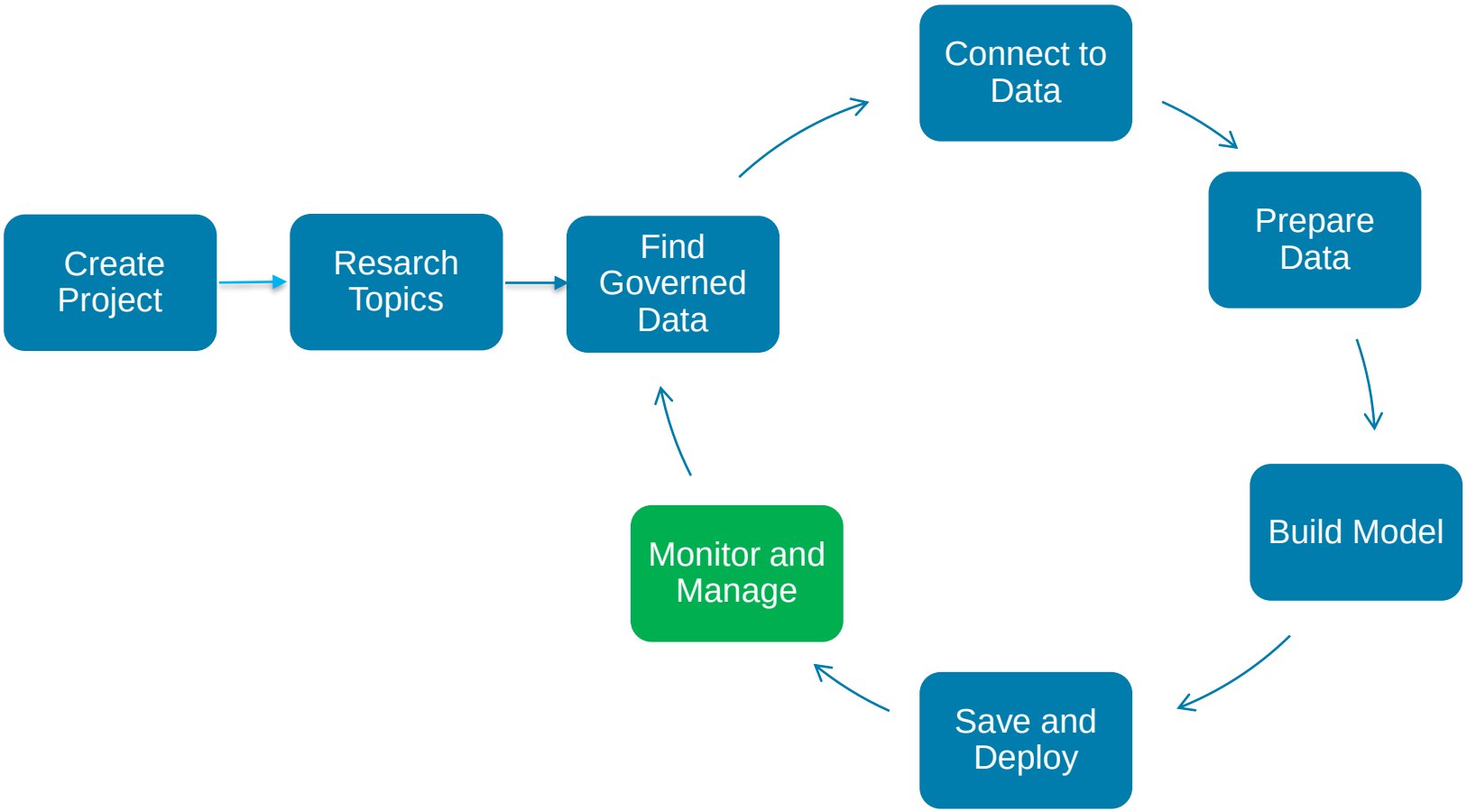
Save and Deploy Models with Watson Machine Learning

Save and Deploy

- Watson Machine Learning API to save/load models to/from repository
- Watson Machine Learning API to deploy saved models easily and have them scale automatically.
- Watson Machine Learning API to invoke deployed models

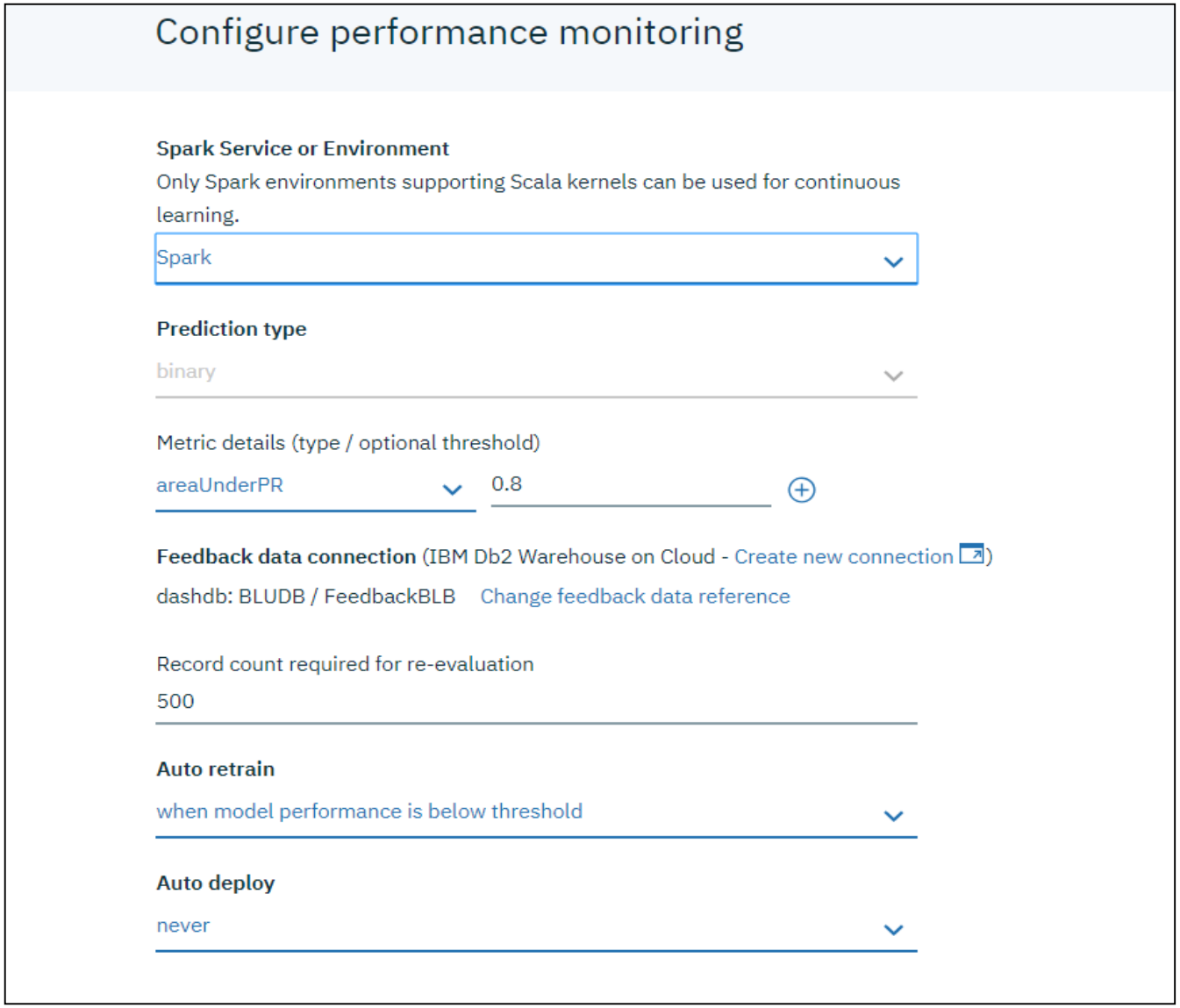
Watson Studio supports the Data Science Lifecycle

Build, train, deploy, and monitor at scale ML/DL workflows to infuse AI into the enterprise to drive innovation.



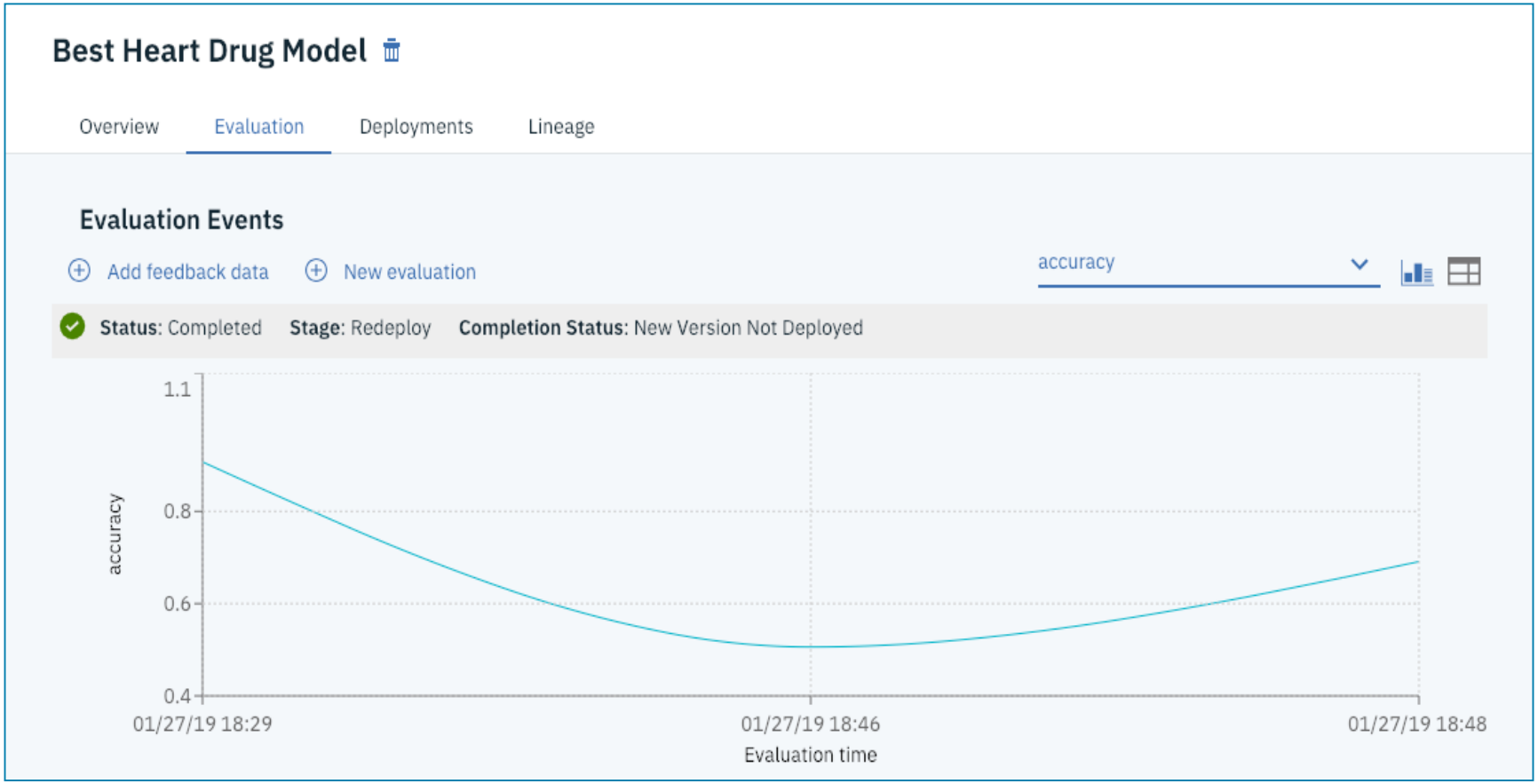
Watson Studio Monitor and Manage

Monitor and Manage



Watson Studio - Monitor and Manage

Monitor and Manage



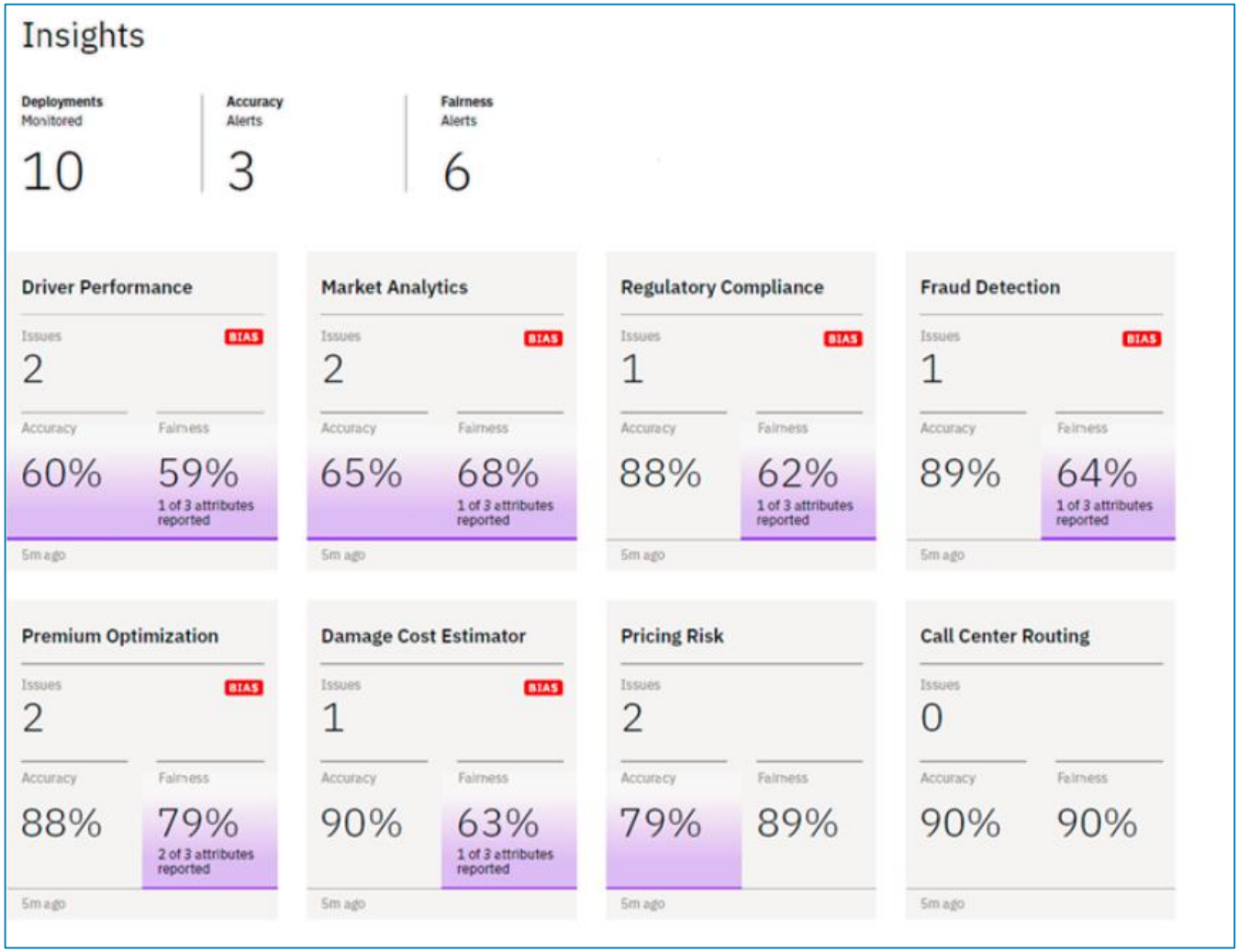
Watson Studio Monitor and Manage Features

Monitor and Manage

- Monitor the performance of the models in production and trigger automatic retraining and redeployment of models.

Watson OpenScale Monitor and Manage

Monitor and Manage



Watson Open Scale Monitor and Manage

Monitor and Manage

Trust and Transparency

- Intelligently delivers bias mitigation help
- Provides traceability & auditability of AI predictions made in production applications
- Tracks AI accuracy in applications
- Explains an outcome in business terms

Automation

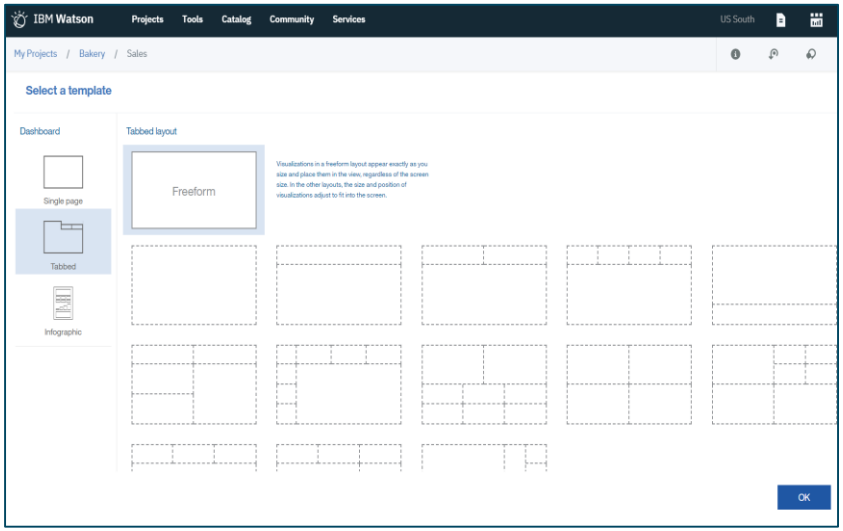
- Automatically detects and mitigates bias in model output, without affecting currently deployed model or outcomes
- NeuNetS (beta) – automatically generate Neural Networks

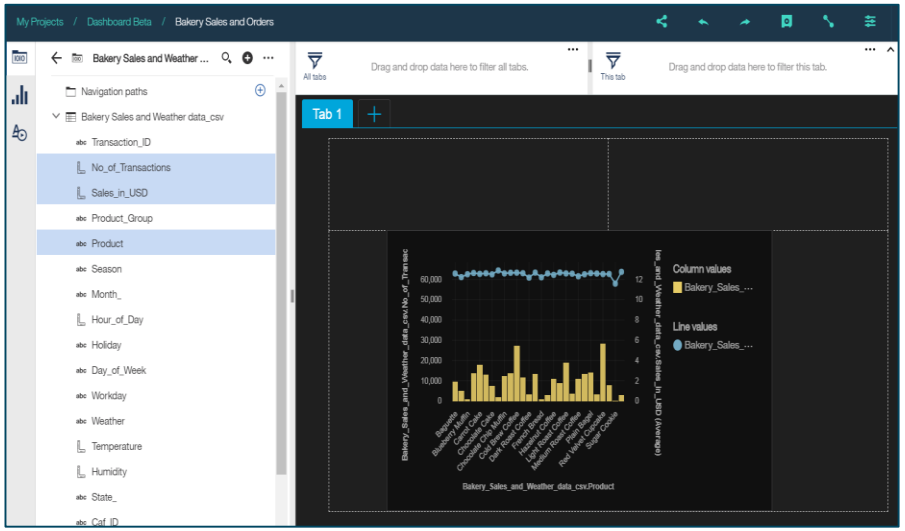
Open By Design

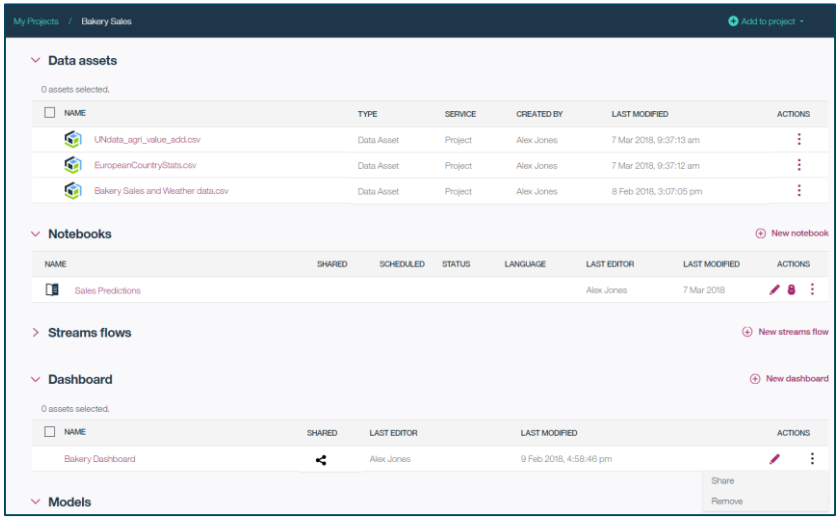
- Monitor models deployed on third party model server engines
- Deploy behind enterprise firewall or on IaaS provider

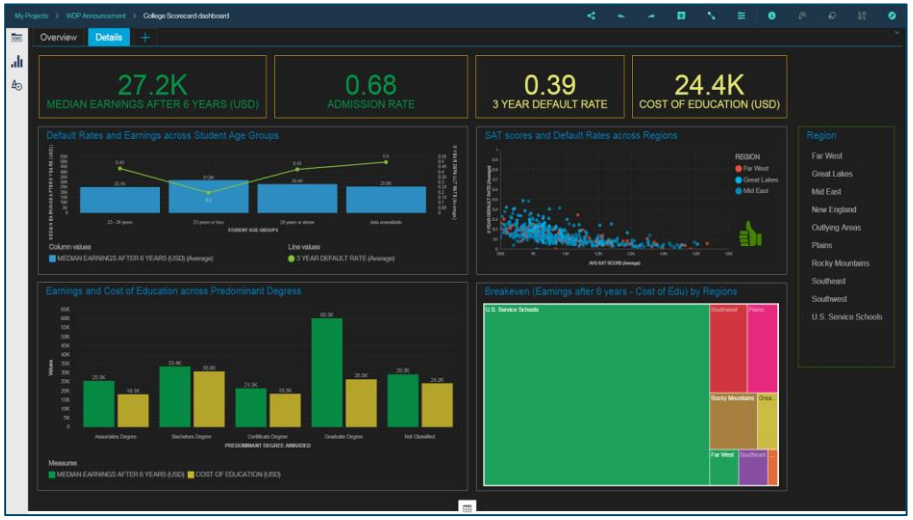
Watson Studio Dynamic Dashboards

Making insights available to all









Watson Studio Takeaways

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

Multiple Deployment Options

- Watson Studio on IBM Cloud – Managed offering
- Watson Studio Local – Private Cloud, Public Cloud-(IBM, Azure, AWS)
- Watson Studio Desktop

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

28

Labs

Lab-1 Sets up the environment which includes creating a Watson Studio project, creating object storage, and creating and associating services to the project (Machine Learning, Spark).

Lab-2 - This lab will use Jupyter Notebooks and the XGBoost library to apply machine learning to a classification problem in the healthcare profession. The Watson Machine Learning API will then be used to save and deploy the model.

Lab-3 - This lab will demonstrate Watson Machine Learning capabilities to simplify the building and deployment of machine learning models. The ability to monitor and adjust the deployed model will be demonstrated via the continuous learning capability of Watson Studio.

Lab-4 - This lab will explore Watson OpenScale’s bias detection, bias mitigation and algorithm explainability features.

Labs

Lab-5 This lab will feature the Watson Studio Neural Network modeler, and Experiment Assistant to build, train, and test a Convolutional Neural Network to classify handwritten digits.

Lab-6 - This lab will guide participants in using the Watson Studio SPSS Modeler capability to explore, prepare, and model passenger data from the Titanic. The SPSS Modeler is a drag and drop capability to build machine learning pipelines.

Lab-7 - Data Refinery Lab – This lab features the Data Refinery tool, a self-service data preparation facility.

Lab-8 – WML + DevOps Lab – This lab uses the Watson Machine Learning Model Builder capability to create a machine learning model based on the Titanic data set. The model will be deployed in the IBM Cloud, and an application will be built that uses the deployed machine learning model to predict survivability given passenger characteristics.

Lab Tips

- Labs are in www.github.com/bleonardb3/ML_POT_03-27 repository.
- Instructions for each Lab are in the README file in the respective Lab folder.
- Cloud development enables making frequent improvements in the user interface. We reviewed the lab instructions and made screen updates so they should be pretty faithful to the user interface. Small differences may occur but shouldn't get in the way of successfully completing the labs.
- Do not use Internet Explorer as the browser. For Mac users do not use Safari.
- For the Jupyter Notebook labs, you execute notebook cells by entering <Shift><Enter> when your cursor is in a code cell. Or you can click on the Run icon in the toolbar.
- All of the Labs should be done in the same project
- For Lab-2 make sure when you are creating the notebook that you switch the environment to the Python-Spark environment.

Thank You

Bernard Beekman
Executive I/T Architect
beekmanb@us.ibm.com

Michael Cronk
I/T Architect
michael.cronk@ibm.com