Introduction to Watson Studio

Power of Data. Simplicity of Design. Speed of Innovation

Federal CTO A3 Center

Bernie Beekman Executive Architect

Michael Cronk Architect

Prithvi Rao

Healthcare Executive Architect

IBM

The AI Ladder

A prescriptive approach to accelerating the journey to AI

AI

INFUSE – decision making ,bias detection, explainability, decision audits

INFUSE – Operationalize AI with trust and transparency

ANALYZE – Build, test, deploy and manage AI models

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

You need a platform for data management
You need a platform for unified data governance & integration
You need a platform for data science and business analytics

You can't do AI without Machine Learning
You can't do Machine Learning without Analytics
And you can't do Analytics without the right data architecture!

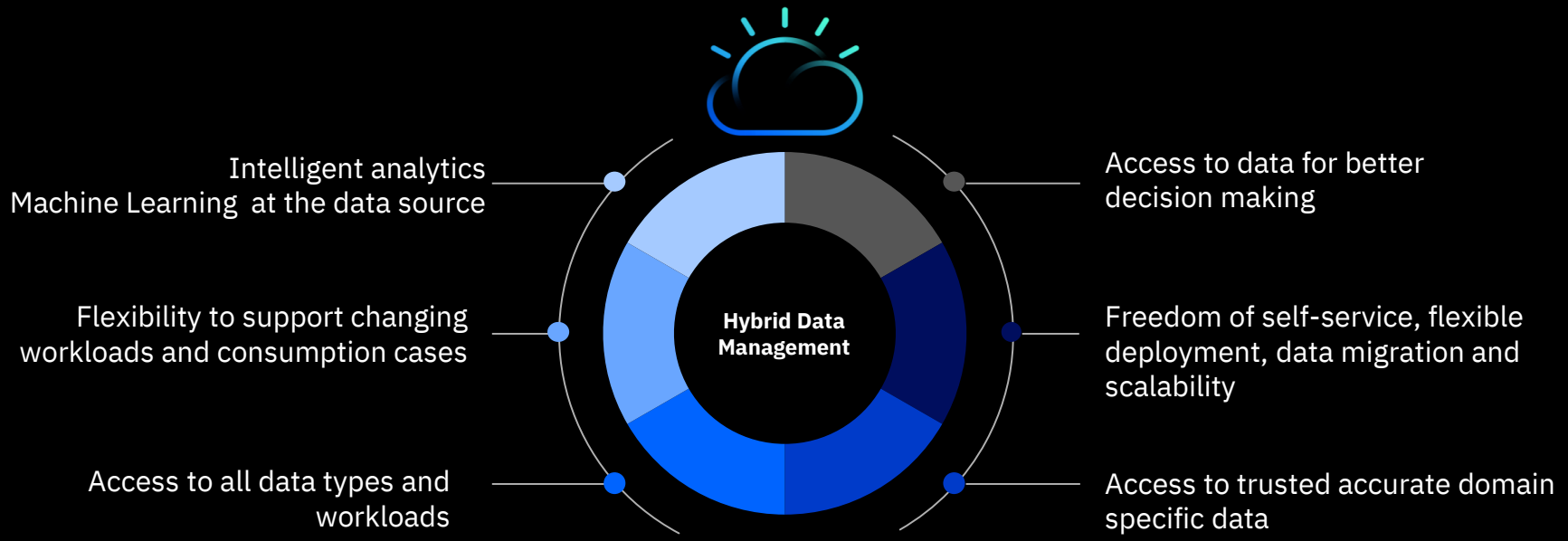
Collect Data

Make data simple and accessible

All data should be discoverable, well documented, trusted and obtainable

Challenges:

- Data held captive in rigid platform choices – segregated & splintered views
- Lack of access to trusted, consistent data across the enterprise
- Data resides in higher volumes, in greater variety and across the enterprise

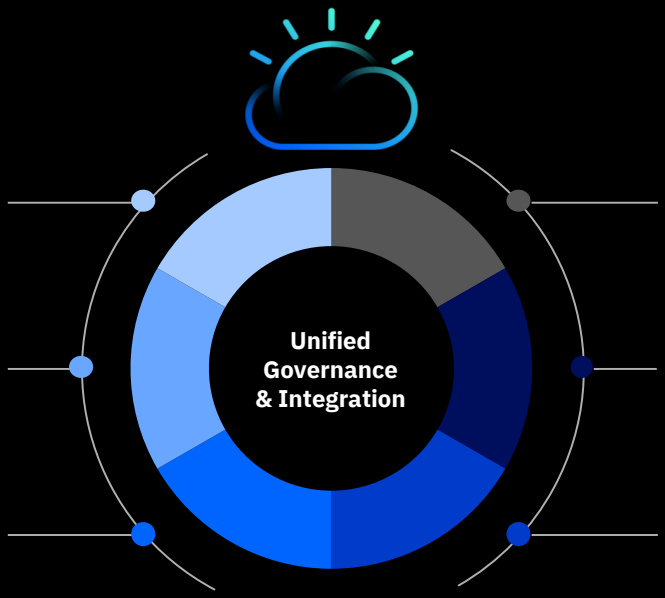


Organize Data

Create a trusted, business-ready analytics foundation

Data is trustworthy, complete and consistent

Challenges:
- Lack of data quality and data governance
- Lack of data consistency and completeness
- Lack of trust in the data



Enable policy driven visibility, detection and reporting

Manage data with built in protection and compliance

Profile, cleanse, integrate and catalog all types of data

Manage data at all stages of the information lifecycle – from creation to deletion

Leverage single source of truth to reduce planning cycles

Provide access to trusted accurate domain specific data

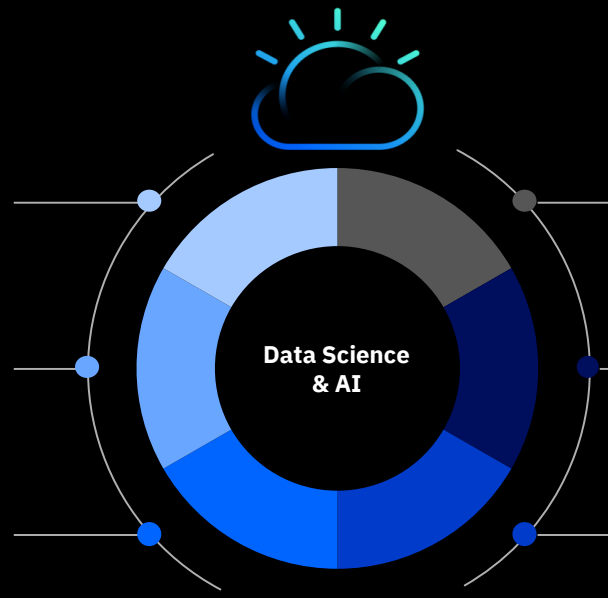
Analyze Data

Scale insights on demand with AI everywhere

Leverage AI and Data Science Platform to leverage the trusted data sets, build, manage, train and deploy ML models

Challenges:

- Point solutions create complexity in integration, maintenance and support
- Lack of a unified analytics platform often results in poor ROI
- Inability to leverage the data to drive better business outcomes with AI & ML



Data Science Platform to make analyzing data and building AI models easier and more accessible

Infuse AI

Operationalize AI with trust and transparency

Ability to trust AI deployed models in production. Ensure fairness, bias mitigation and provide explainability for AI decisions/predictions

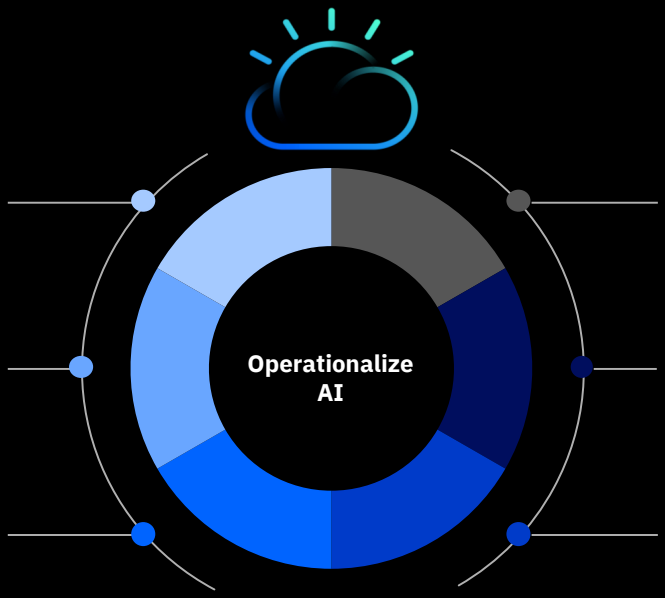
Challenges:

- Do I trust the performance of AI enabled applications?
- If data used to train the models has unfair bias, resulting recommendations aren't transparent & trusted
- Does the AI model provide any explanation for the recommendation made? Lack of explainability

Automate fairness and issue detection

Provide full transparency and provide explainability for the recommendations made

Detect and proactively mitigate bias to ensure fair model performance



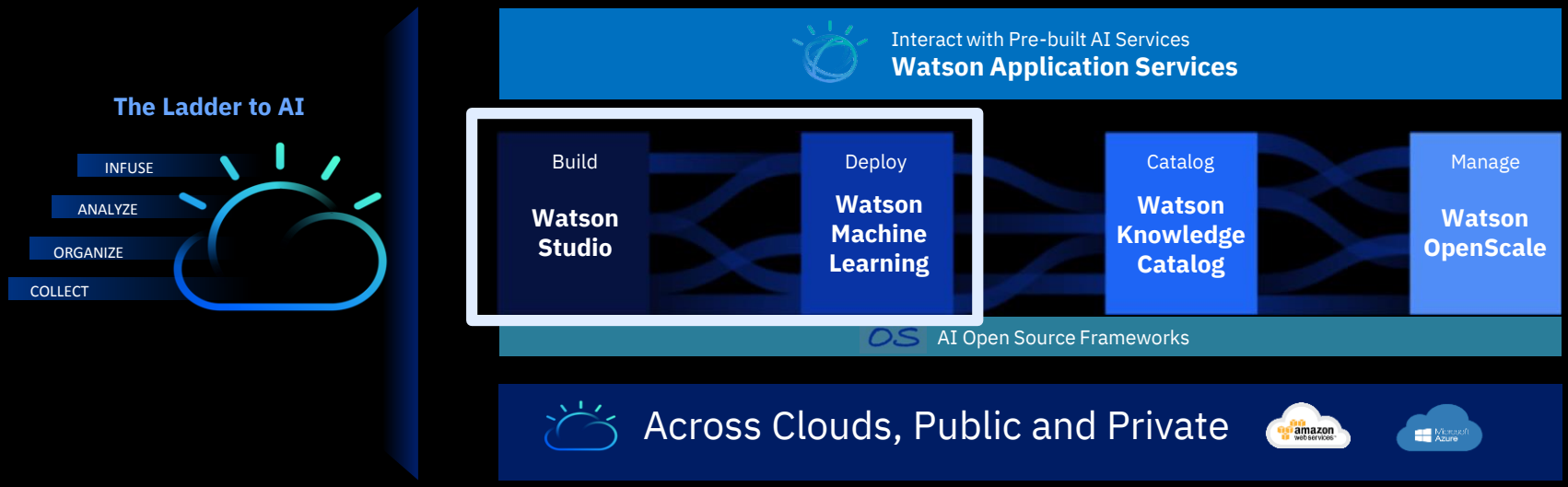
Trace data lineage and training used to build models

Automate the design and deployment of neural networks

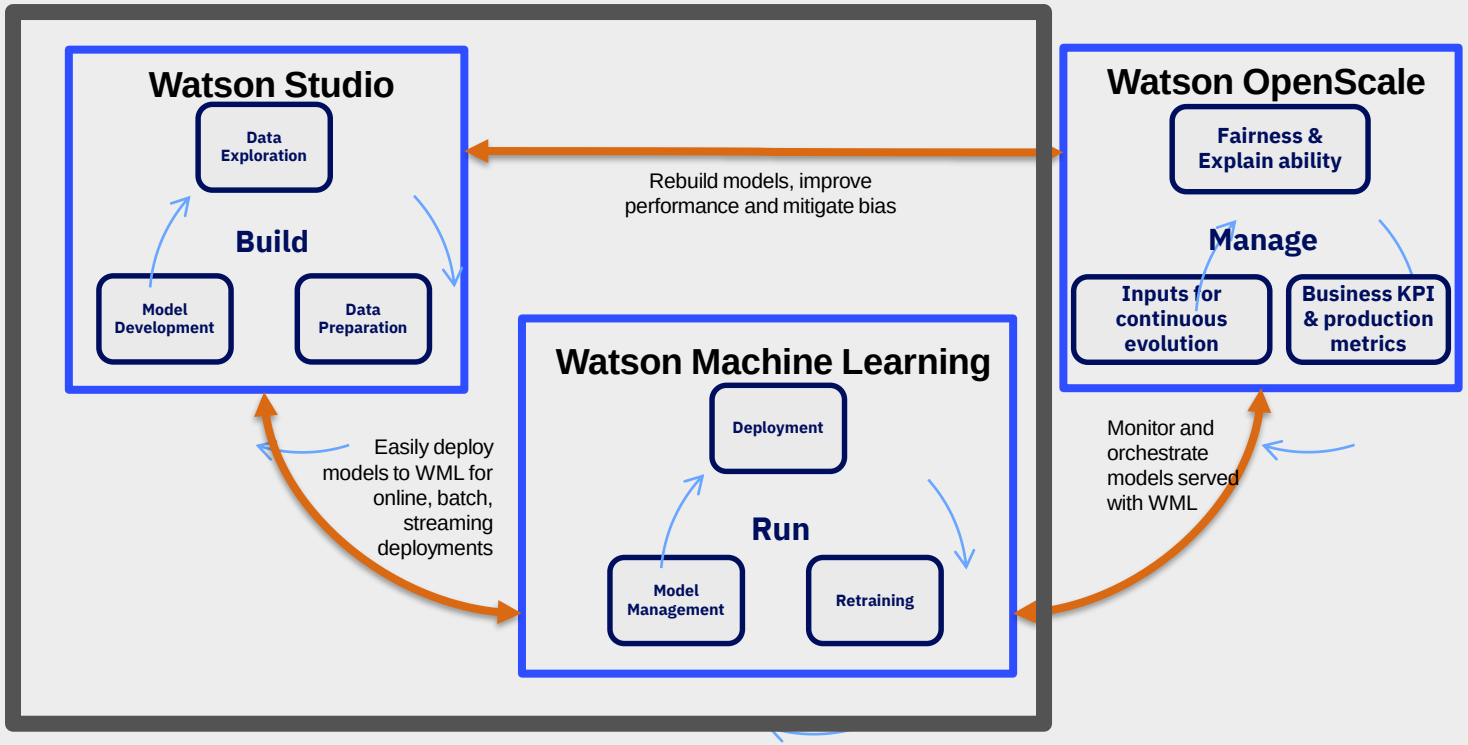
Automate design, monitoring and continuous improvement of models to reduce burden on teams

IBM AI Portfolio

Everything you need for Enterprise AI, on any cloud

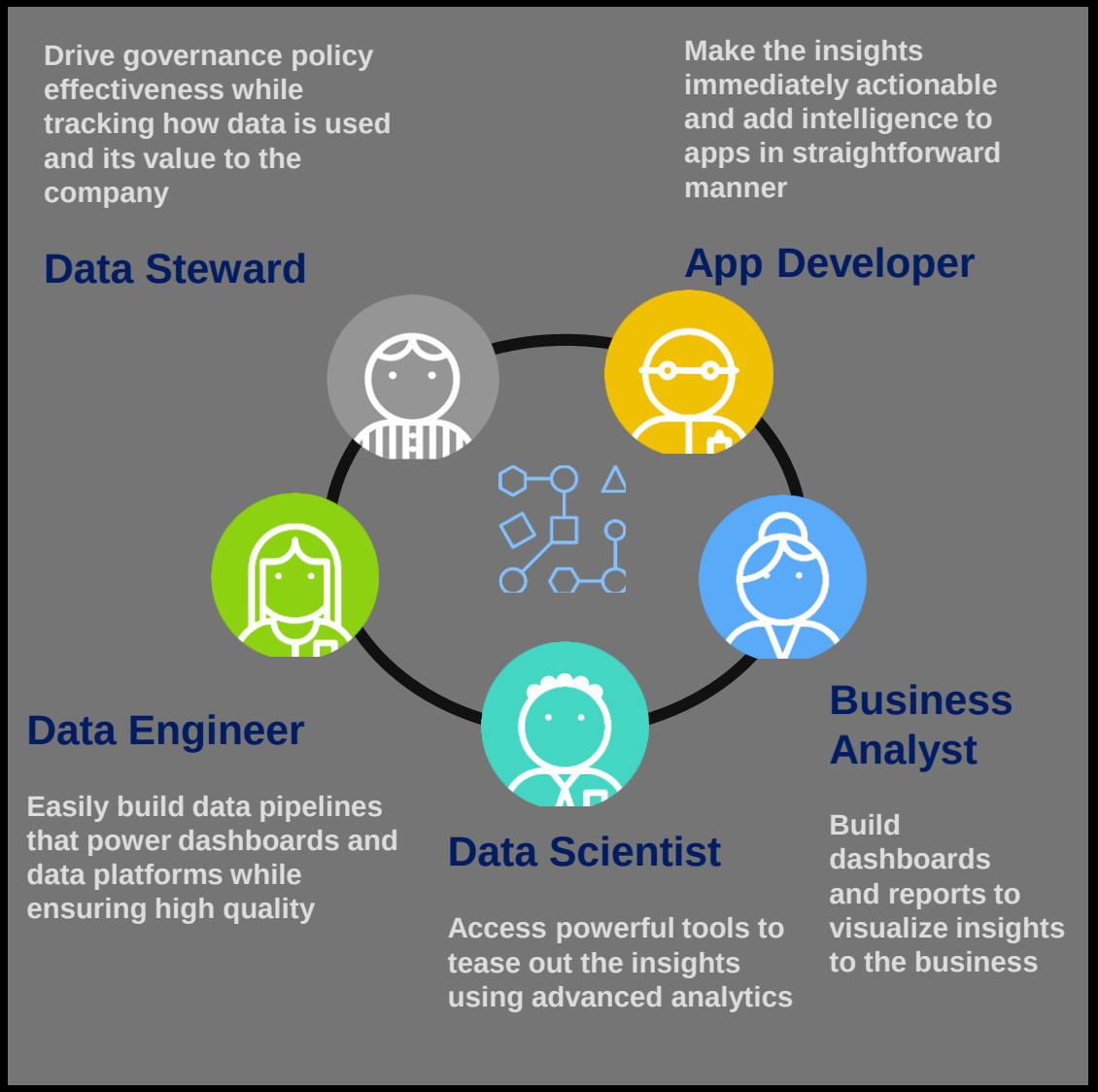


Accelerate your data science lifecycle from discovery to production

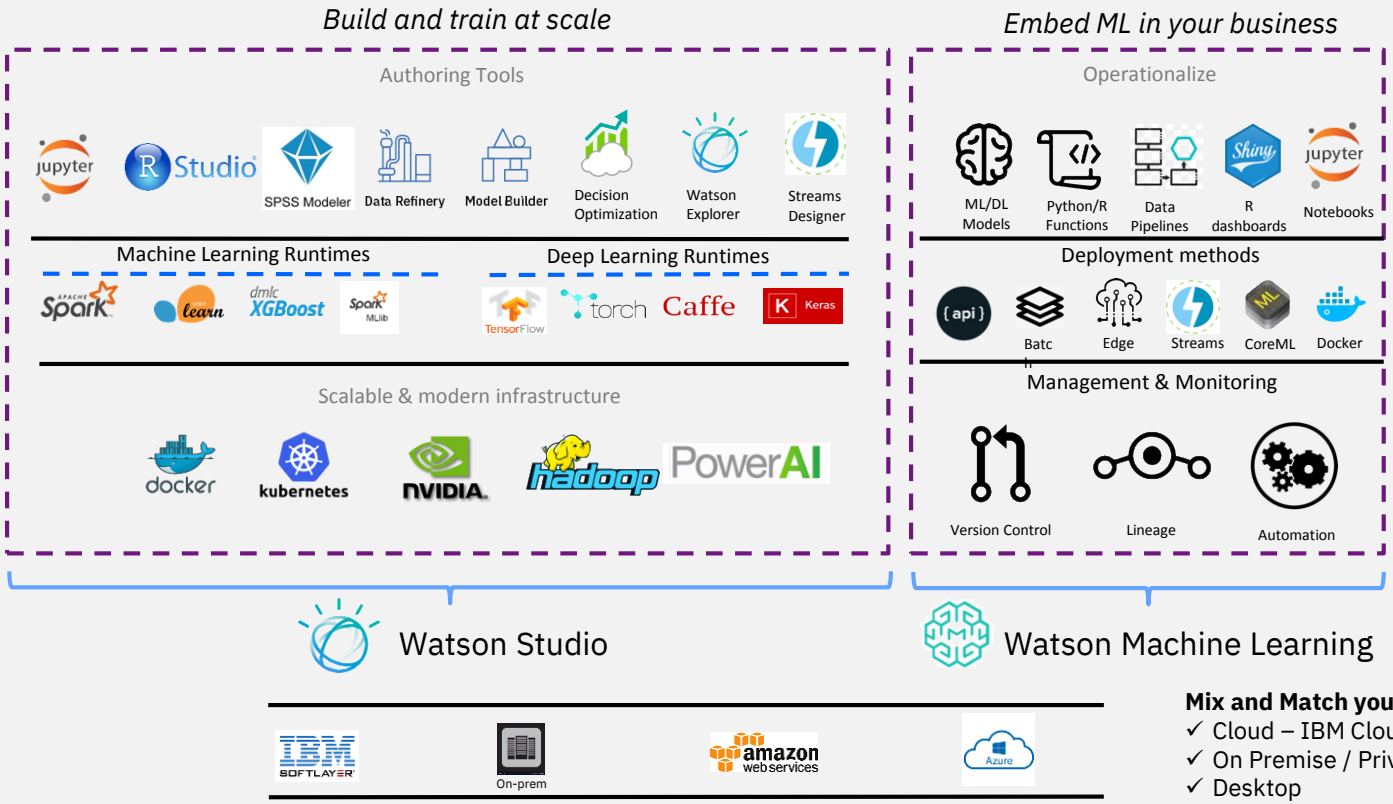


IBM Watson Studio Platform

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



Watson Studio and Watson Machine Learning inject AI firepower into your business

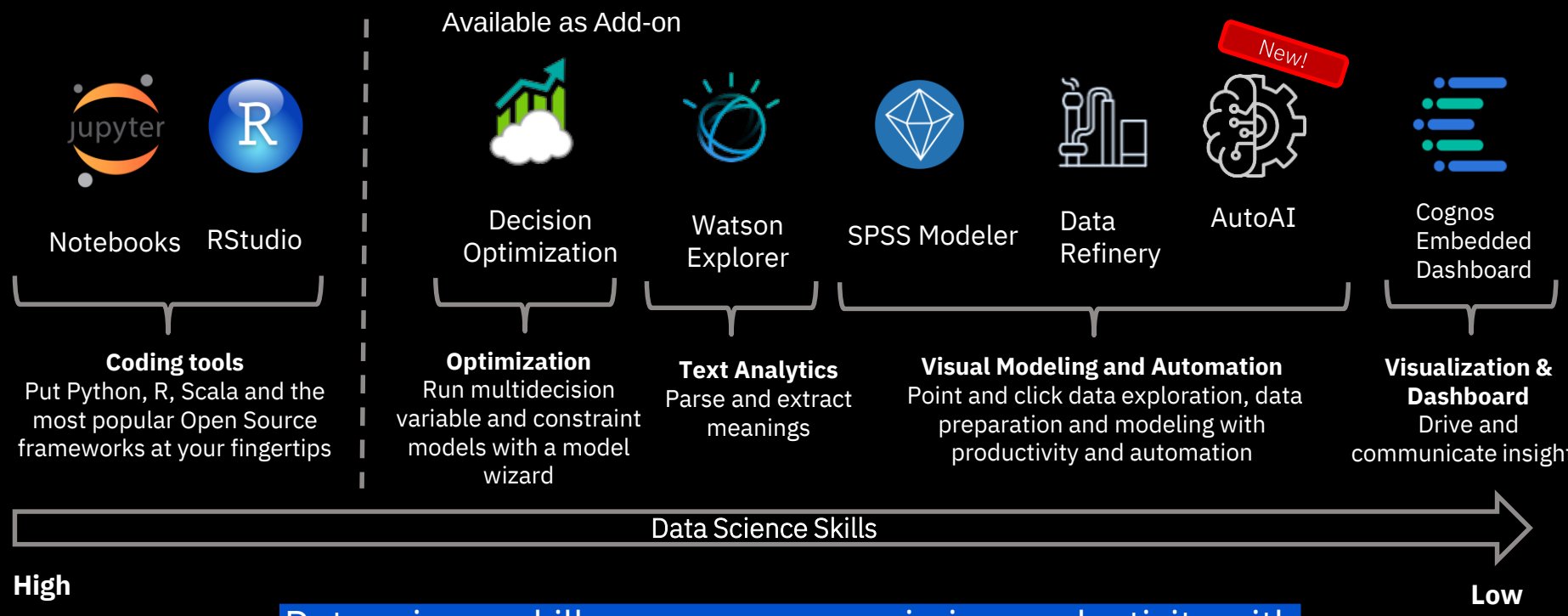


Mix and Match your deployment

- ✓ Cloud – IBM Cloud, Azure, AWS
- ✓ On Premise / Private Data center
- ✓ Desktop

Watson Studio and Watson Machine Learning Add-ons

More powerful and flexible tools built for teams

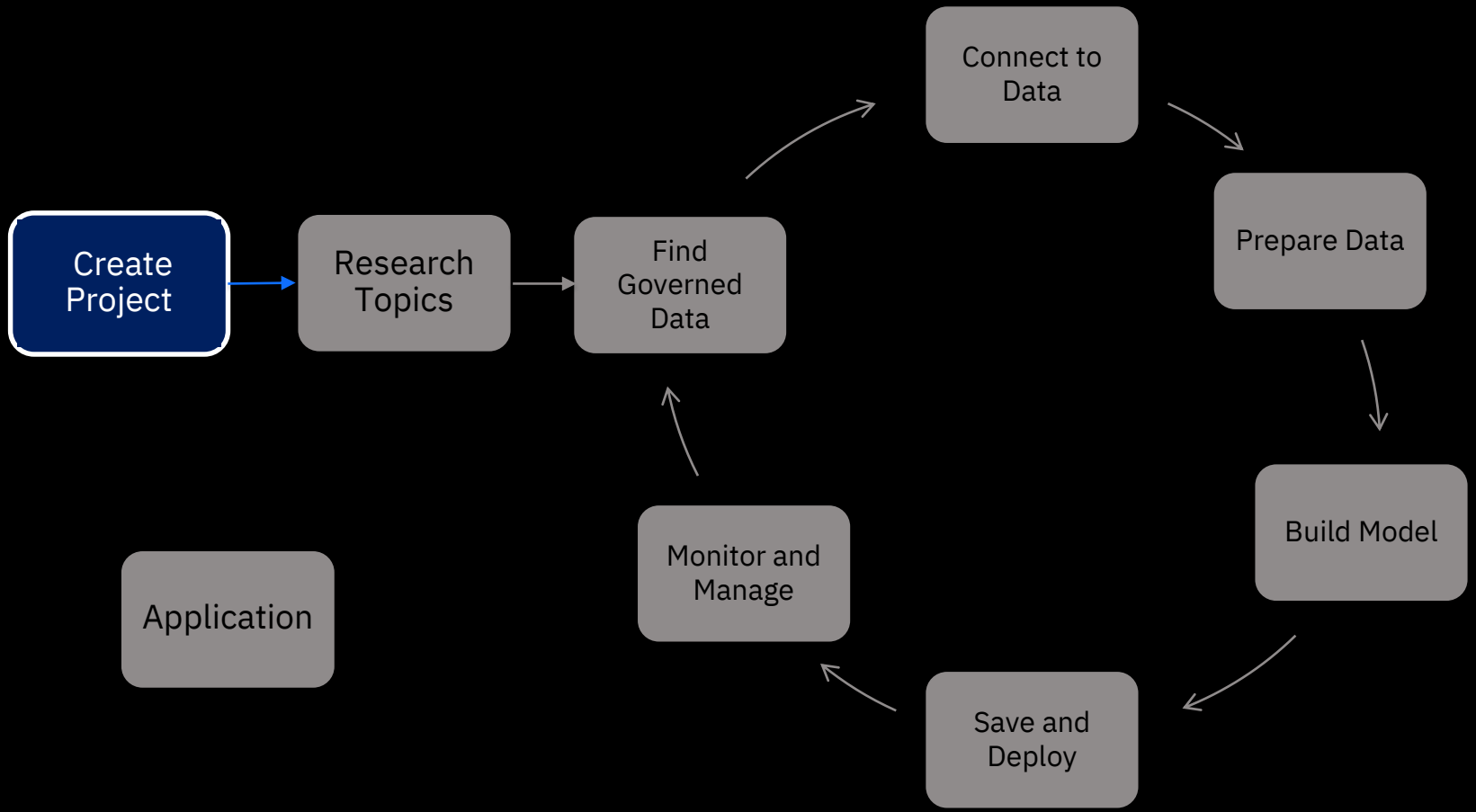


Data science skills are scarce – maximize productivity with IBM Watson Studio

IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Studio Project Features

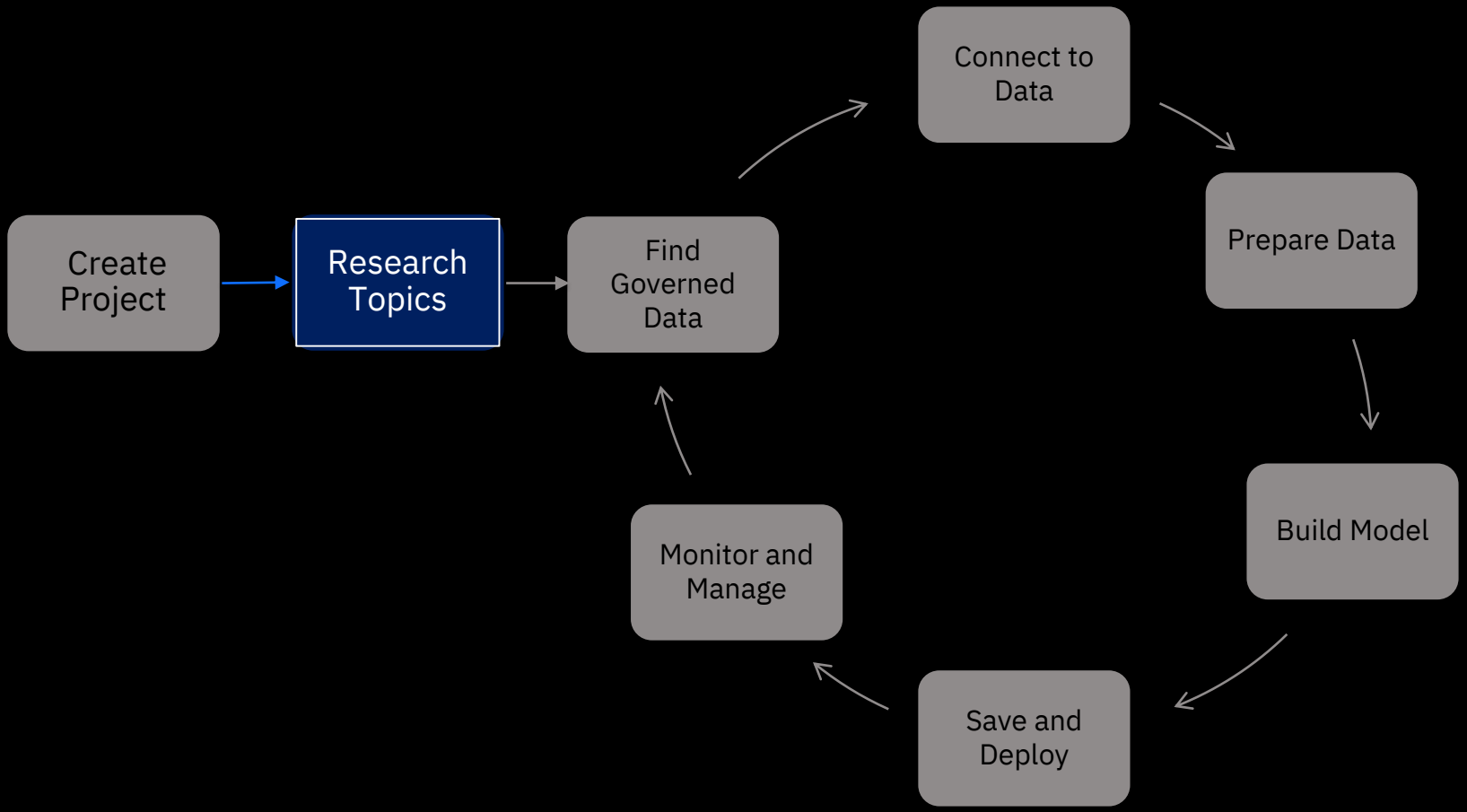
Making Data Science a Team Sport

- Organize resources to achieve a particular data analysis goal
- Support role-based collaboration (Admin, Editor, Viewer)
- Share Assets from all IDEs across one Watson Studio project: notebooks, data sources, flows, models, etc.
- Export/Import Projects

IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Studio Project Features

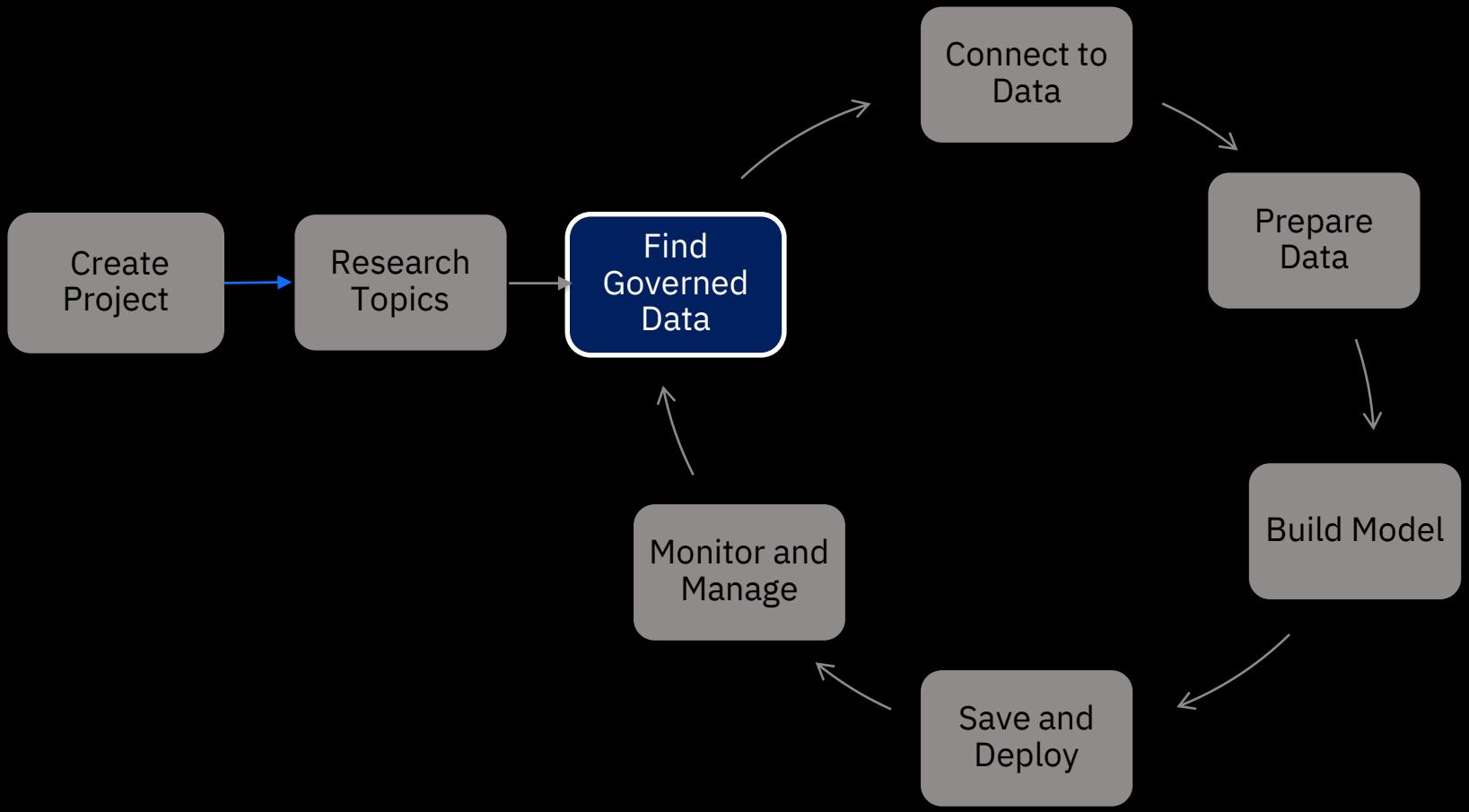
Built in learning to get started

- Enable Community Card Feature that include curated articles, tutorials, notebooks, data sets, and papers
- Enable Bookmark in Projects
- Copy notebooks or Data Sets into projects
- Continuously update IBM’s managed services

IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Knowledge Catalog Features

Unlock tribal knowledge and unleash knowledge workers

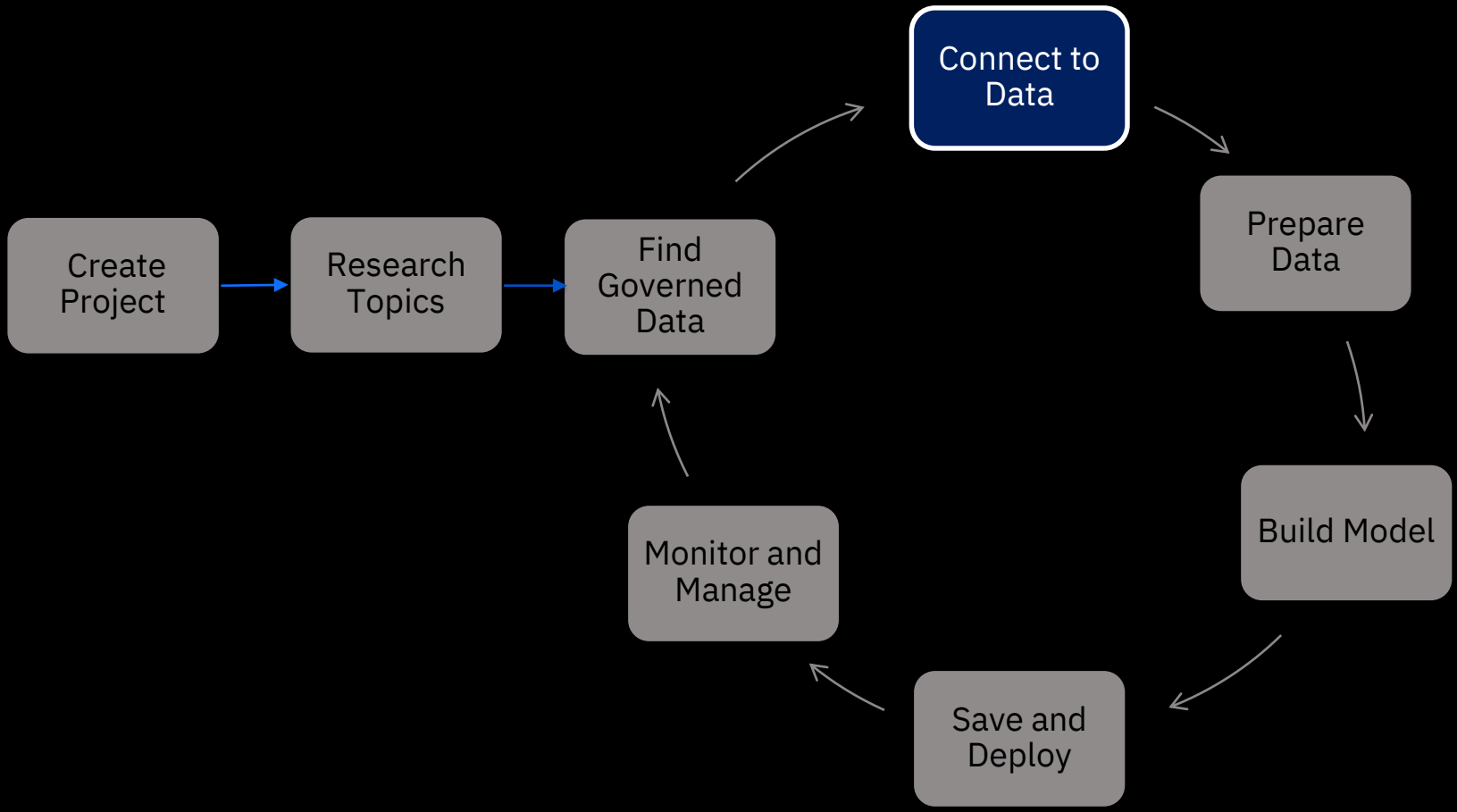
- **Find** data (structured, unstructured) and AI assets (e.g., ML/DL models, notebooks, Watson Data Kits) in the **Knowledge Catalog** with intelligent search and giving the right access to the right users.

- Discover assets, profiling, classification
- Author policies and rules
- Enforce policies and rules
- Provide asset usage statistics

IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Studio Connection Features

Unlock tribal knowledge and unleash knowledge workers

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

IBM Watson Studio Platform

Watson Studio Connection Features

Unlock tribal knowledge and unleash knowledge workers

Connect to Data

New connection

IBM services

- BigInsights HDFS
- Cloud Object Storage
- Cloud Object Storage (infrastructure)
- Cloudant
- Cognos Analytics
- Compose for MySQL
- Compose for PostgreSQL
- Db2
- Db2 Big SQL
- Db2 for i
- Db2 for z/OS
- Db2 Hosted
- Db2 on Cloud
- Db2 Warehouse
- Informix
- Object Storage OpenStack Swift (infrastructure)
- PureData for Analytics
- Watson Analytics

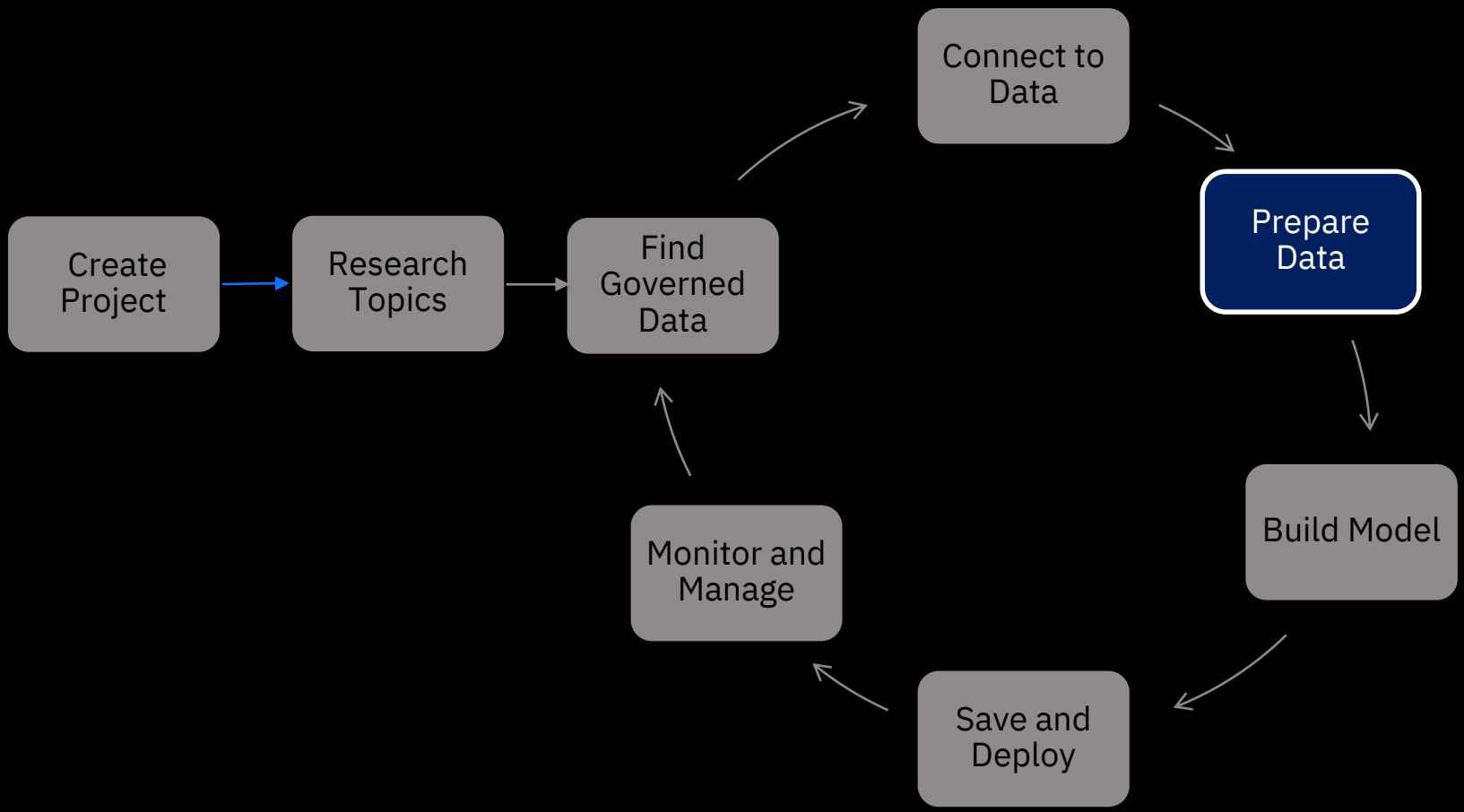
Third-party services

- Amazon Redshift
- Amazon S3
- Apache Hive
- Cloudera Impala
- Dropbox
- FTP
- Google BigQuery
- Google Cloud Storage
- Hortonworks HDFS
- Looker
- Microsoft Azure Data Lake Store
- Microsoft Azure SQL Database
- Microsoft SQL Server
- MySQL
- Oracle
- Pivotal Greenplum
- PostgreSQL
- Salesforce.com
- Sybase
- Sybase IQ
- Tableau
- Teradata

IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Studio Data Refinery Features

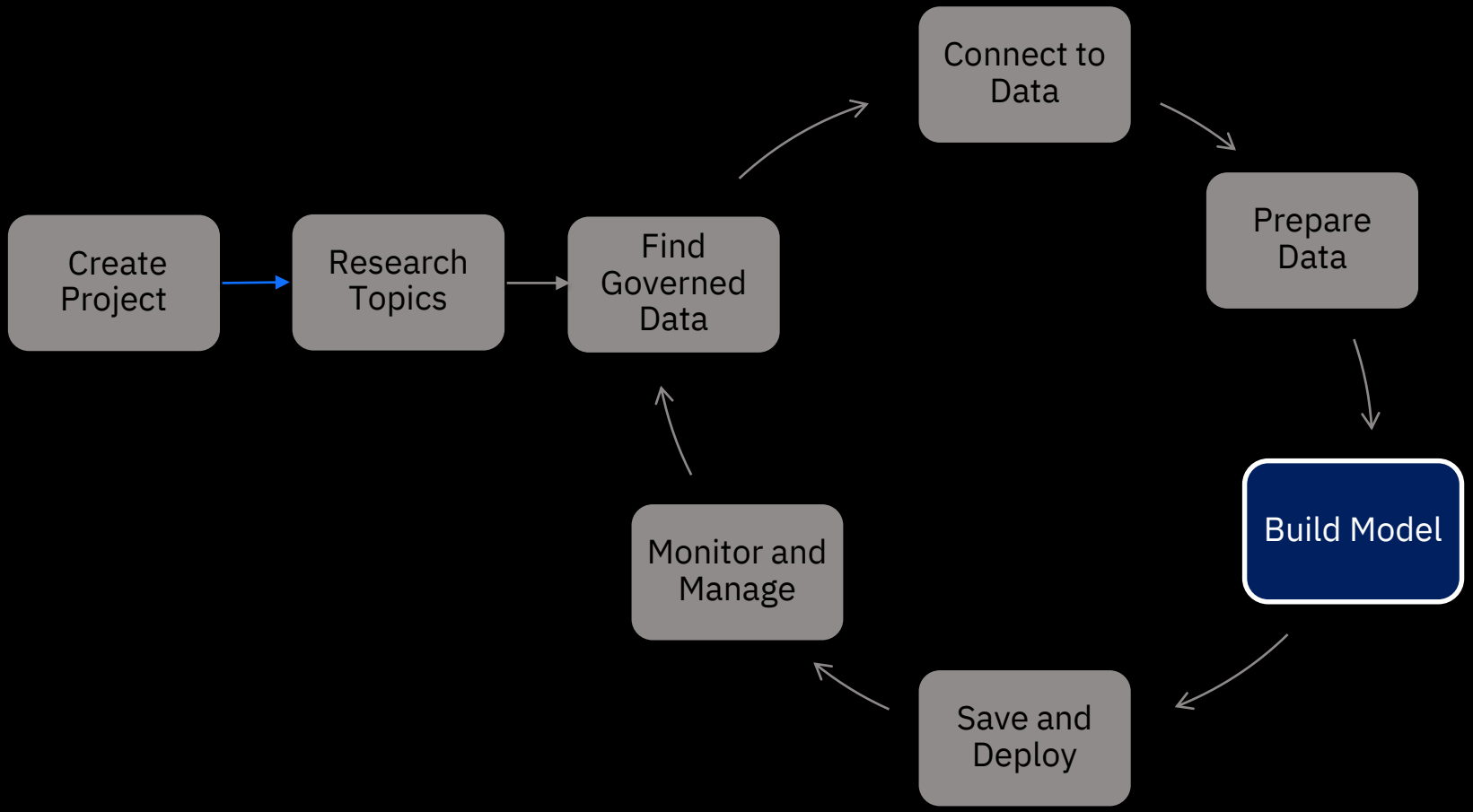
Making data fit for work

- Use the Data Refinery tool to profile, visualize, and shape data.
- Create data preparation pipelines via point and click capability on subset of data
  - ✓ Cleanse the data: fixing or removing data that is incorrect, incomplete, improperly formatted, or duplicated
  - ✓ Shape the data: customize data by filtering, sorting, combining, or removing columns, and performing operations
- Run the pipeline on all the data
  - Manually (on demand)
  - Automated (scheduled)

IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Studio Model Building Features

The best of Open Source and IBM Watson Tools to create state of the art data products

Open Source Tools

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

IBM Tools

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

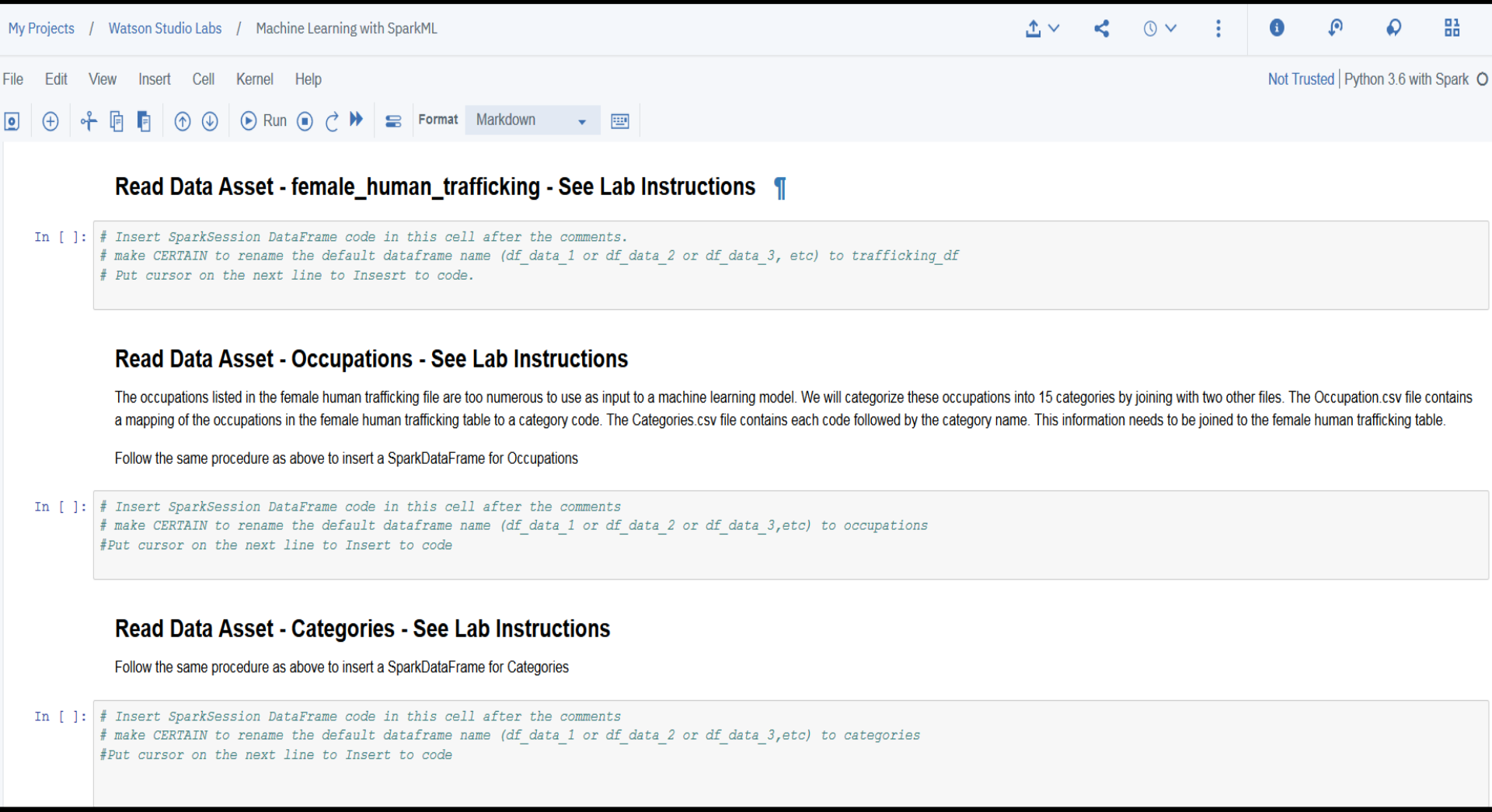
Train at scale on **GPUs** and **distributed** compute

** in hands-on labs

Build Model

IBM Watson Studio Platform

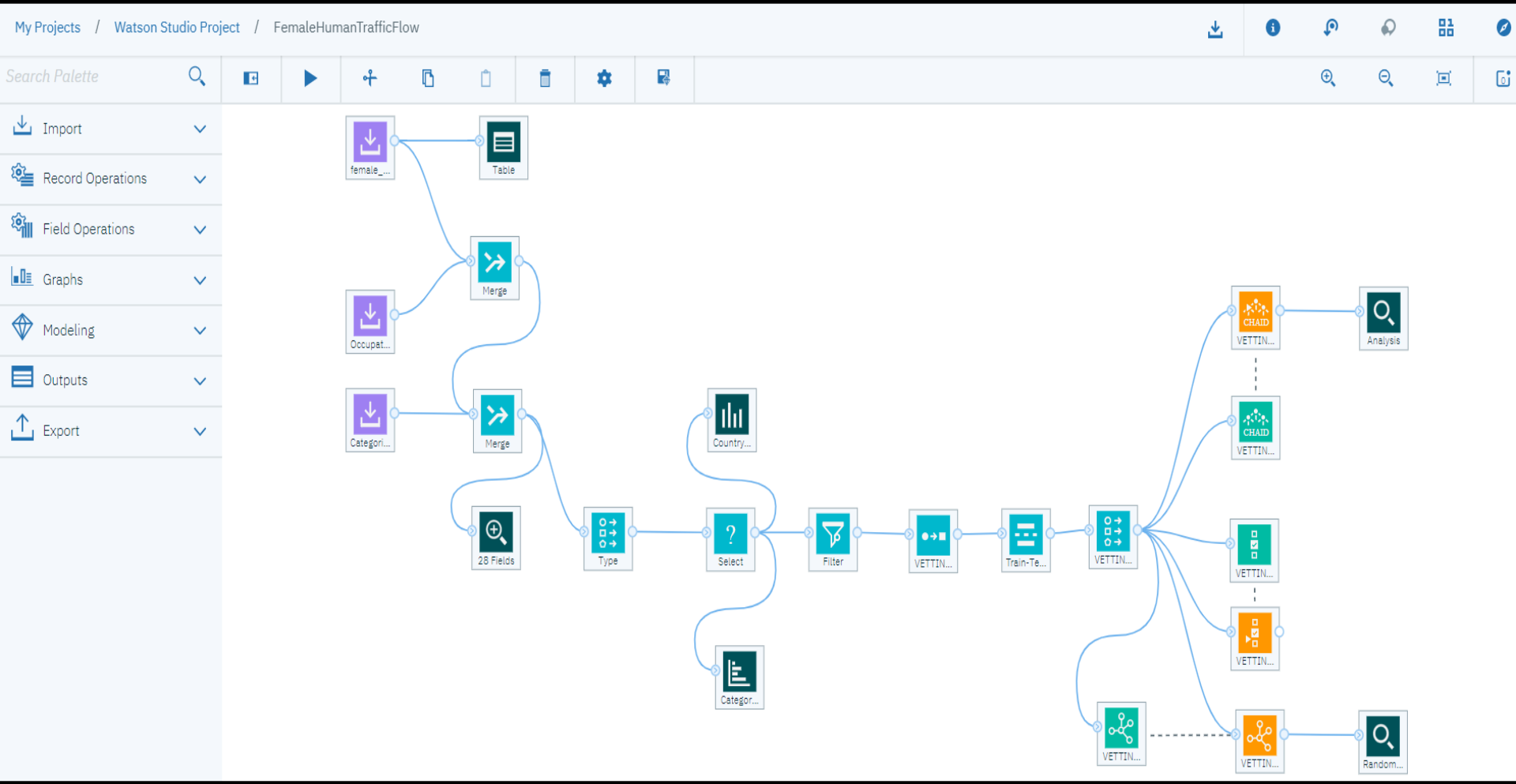
Watson Studio Jupyter Notebook



Build Model

IBM Watson Studio Platform

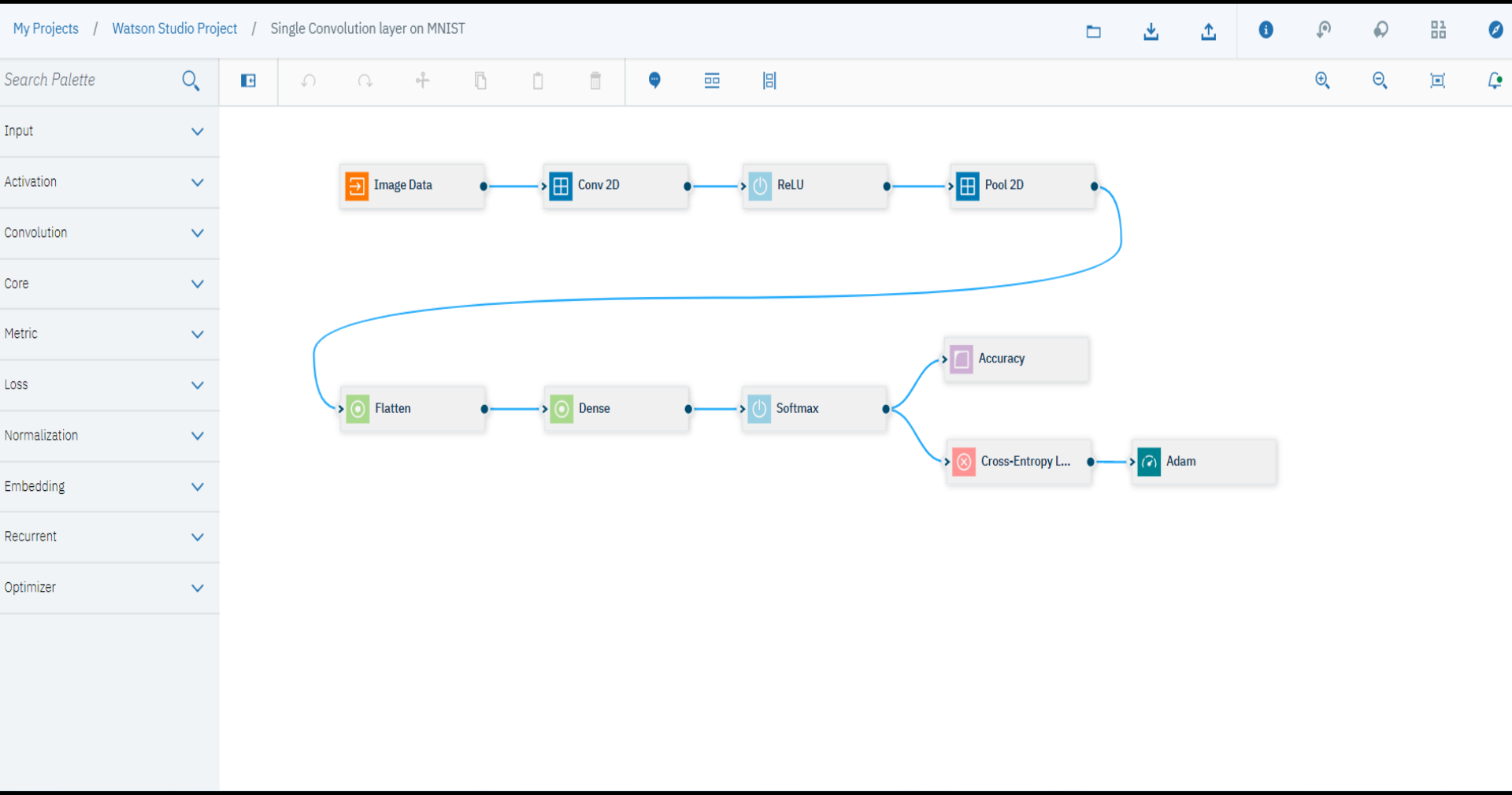
Watson Studio SPSS Modeler



Build Model

IBM Watson Studio Platform

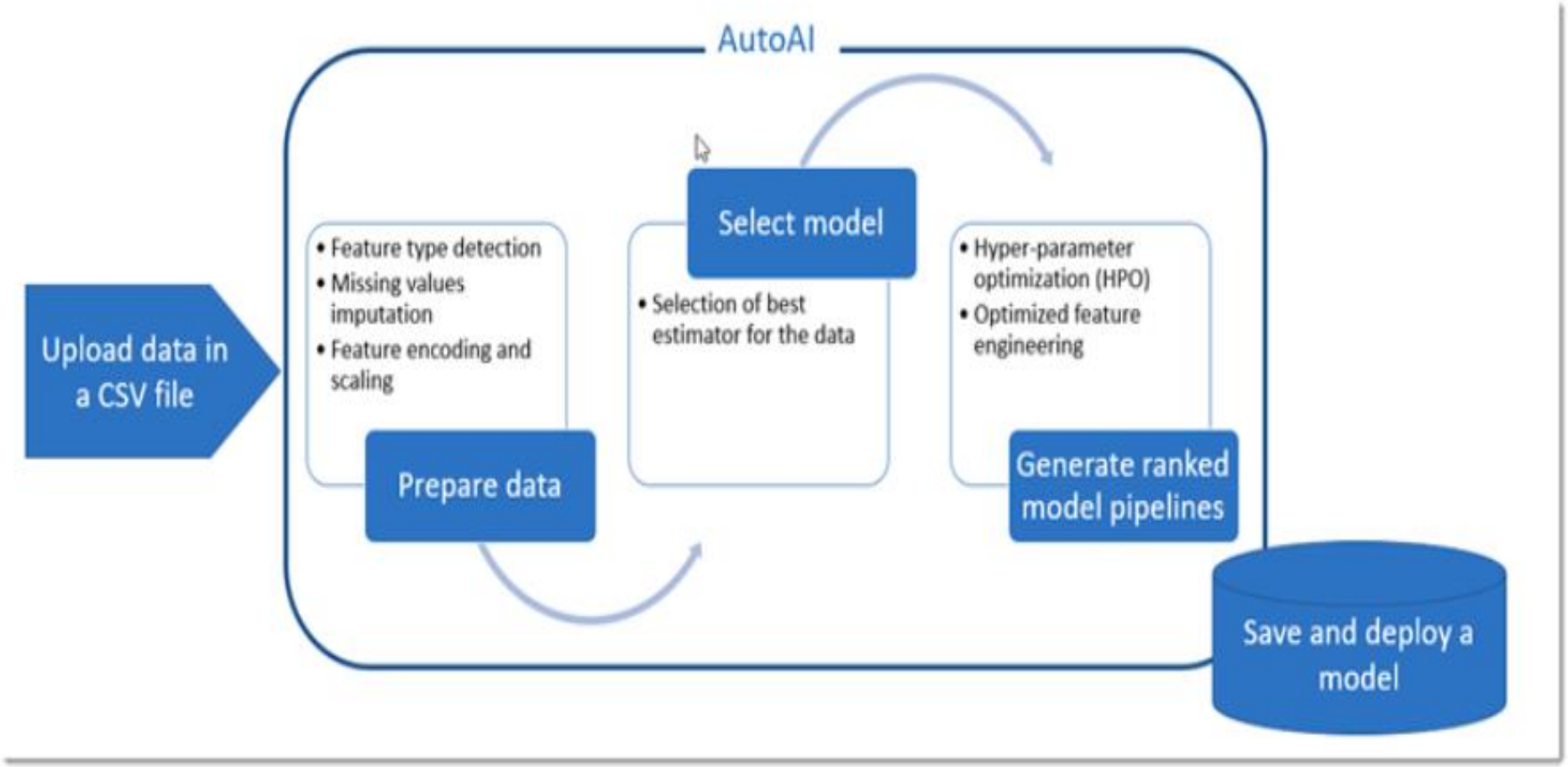
Watson Studio Neural Network Modeler



Build Model

IBM Watson Studio Platform

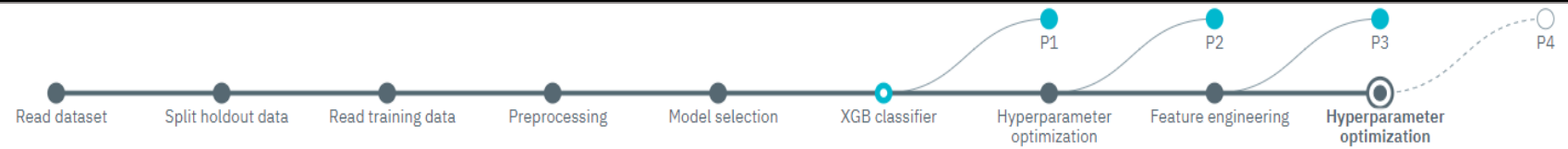
Watson Studio Auto AI



Build Model

IBM Watson Studio Platform

Watson Studio Auto AI



Pipeline leaderboard

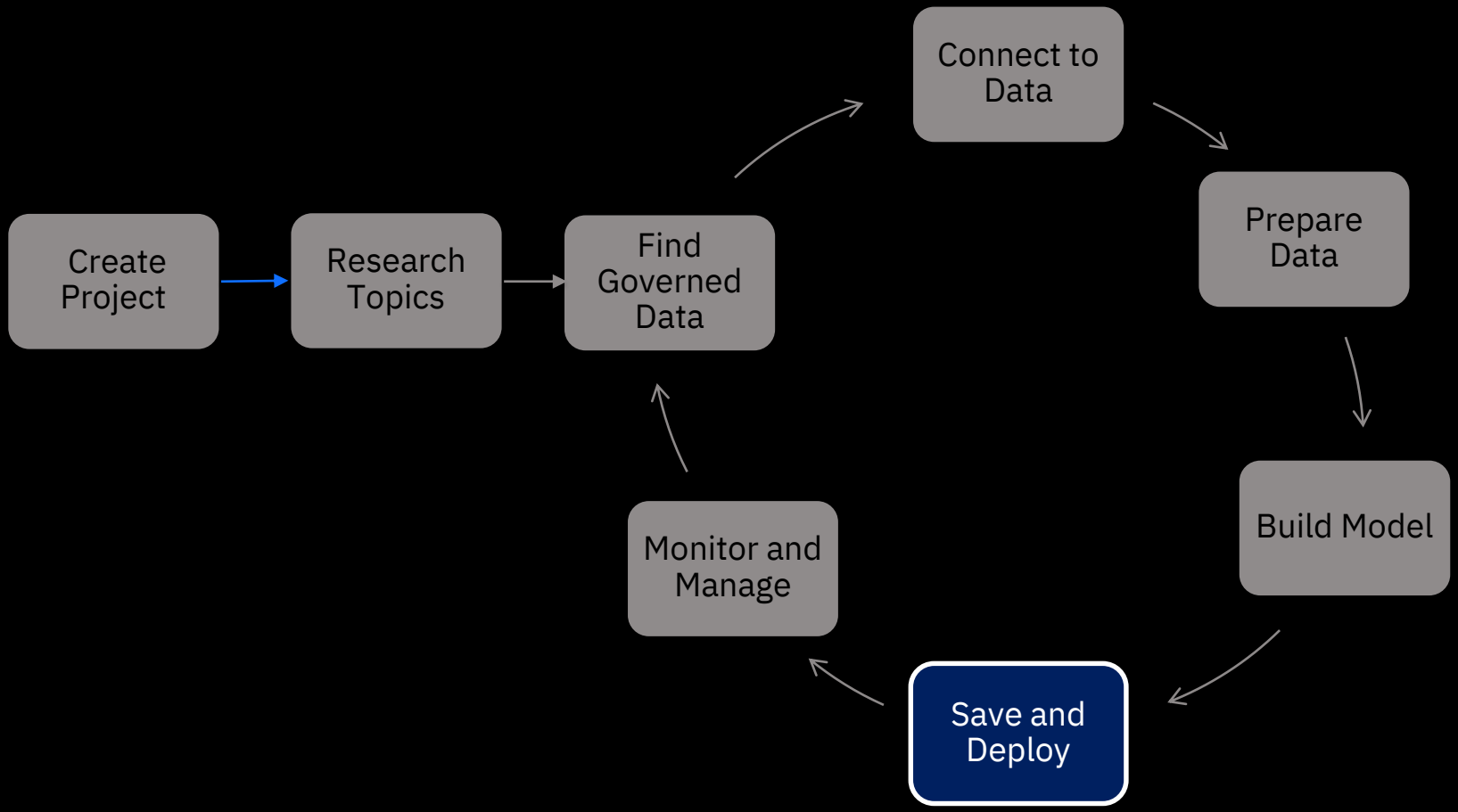
Compare pipelines Ranking based on: Accuracy

	RANK	ACCURACY	PIPELINE INFORMATION		
>	1	0.897	P3 - XGB classifier estimator Transformers (8): Preprocessing > Standard scaler > Univariate feature selection > Sine > Univariate feature selection > Tangent > ...	View details	Save as model
>	2	0.884	P1 - XGB classifier estimator Transformers (2): Preprocessing > XGB classifier estimator	View details	Save as model
>	3	0.884	P2 - XGB classifier estimator Transformers (2): Preprocessing > XGB classifier estimator	View details	Save as model

IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Studio Save and Deploy Features

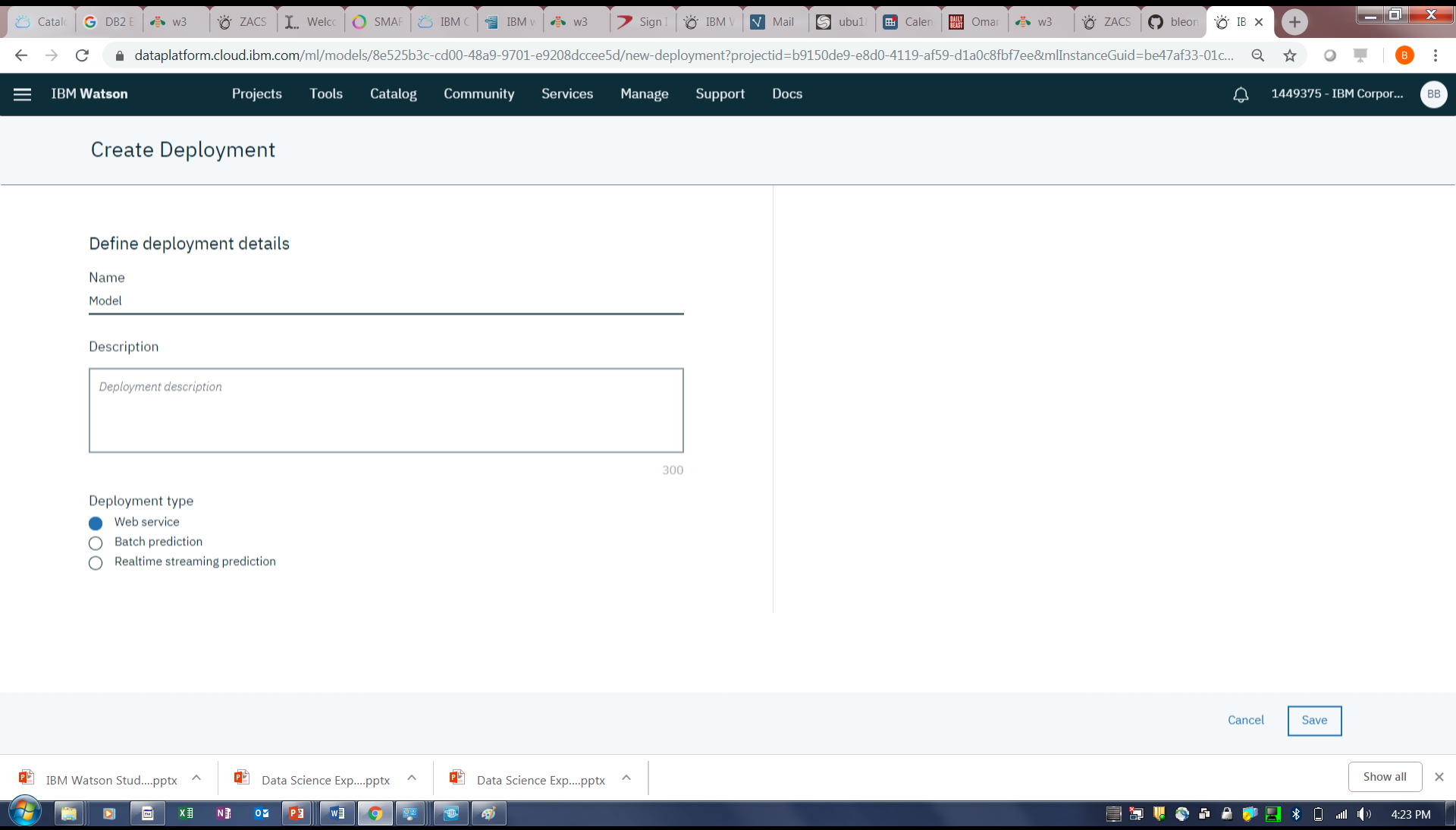
Save and deploy models with Machine Learning in Watson Studio Platform

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

Save & Deploy

IBM Watson Studio Platform

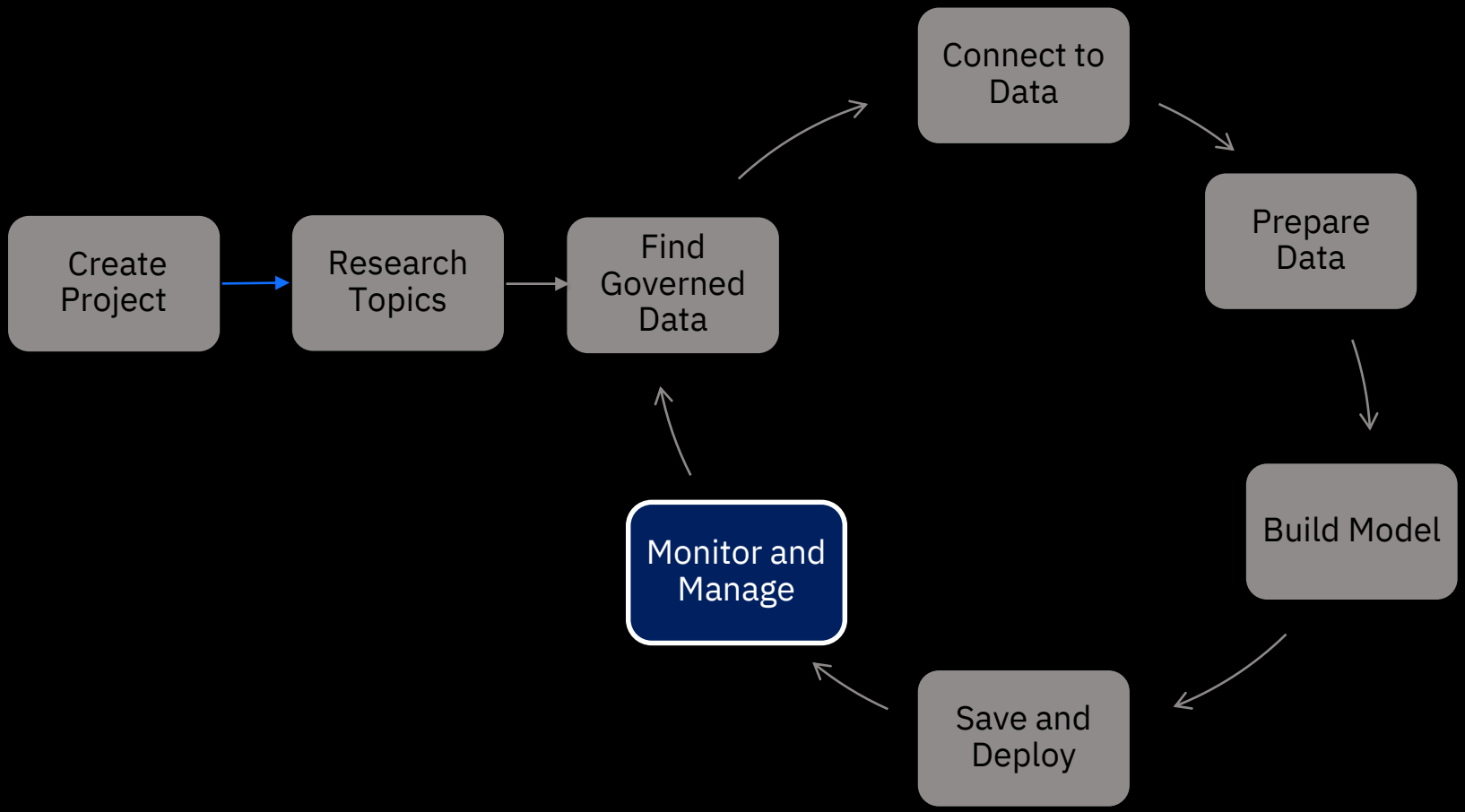
Watson Studio Save and Deploy Features



IBM Watson Studio Platform

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.



IBM Watson Studio Platform

Watson Studio Monitor and Manage Features

Save and deploy models with Machine Learning in Watson Studio Platform

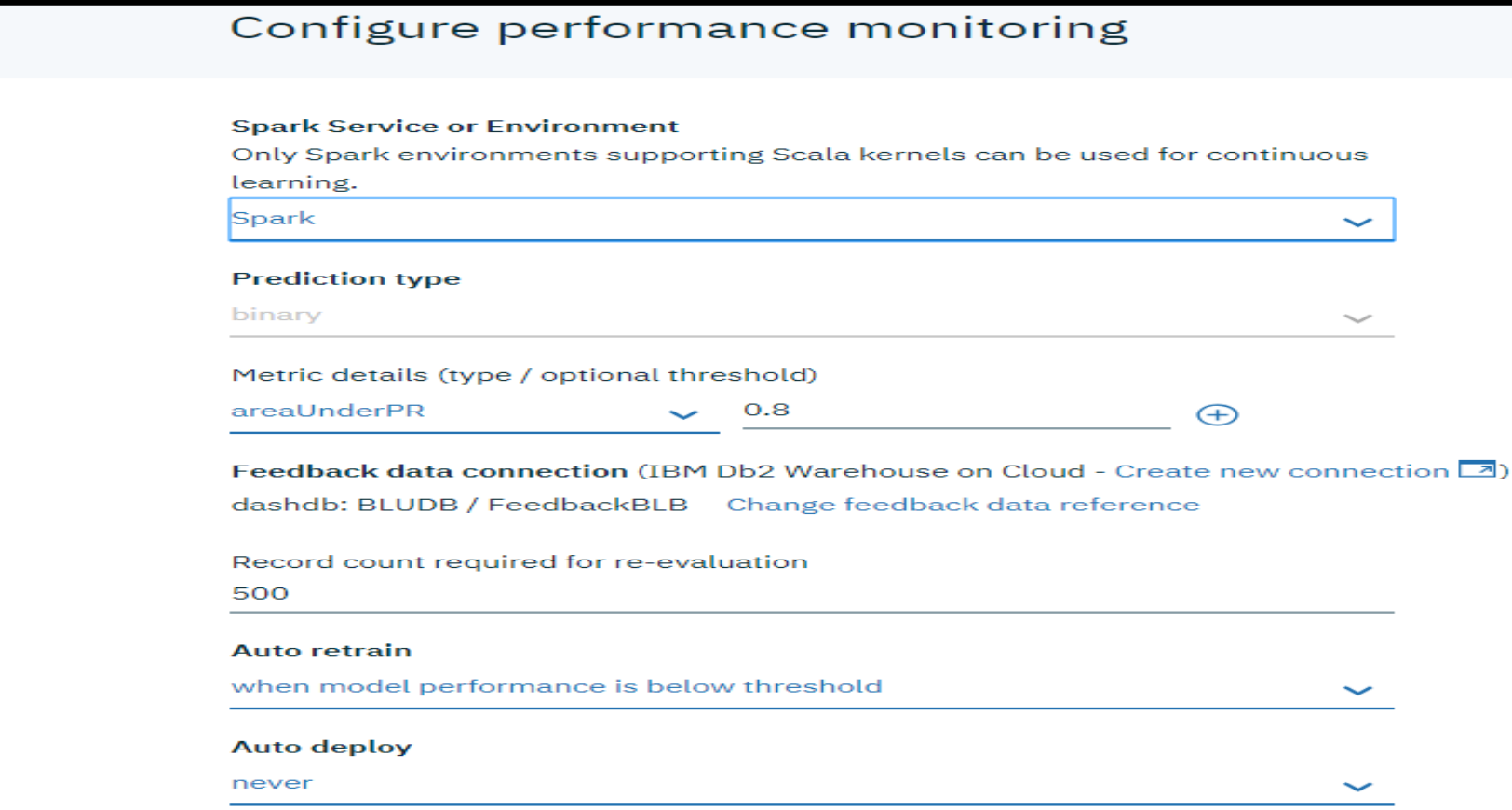
- Monitor the performance of the models in production and trigger automatic retraining and redeployment of models.

Monitor & Manage

IBM Watson Studio Platform

Watson Studio Monitor and Manage Features

Monitor and Manage with Watson Studio Platform



Trust & Transparency

IBM Watson Studio Platform

Watson Studio Open Scale

Build trust, transparency and explain ability with IBM Open Scale

Trust and Transparency

- Intelligently deliver bias mitigation
- Provide 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

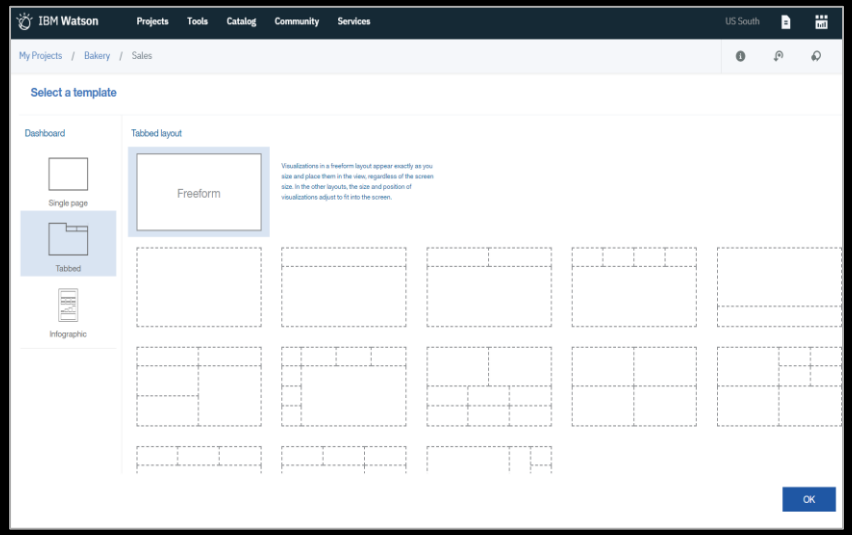
- Monitor models deployed on third party model server engines
- Deploy behind enterprise firewall or on IaaS provider

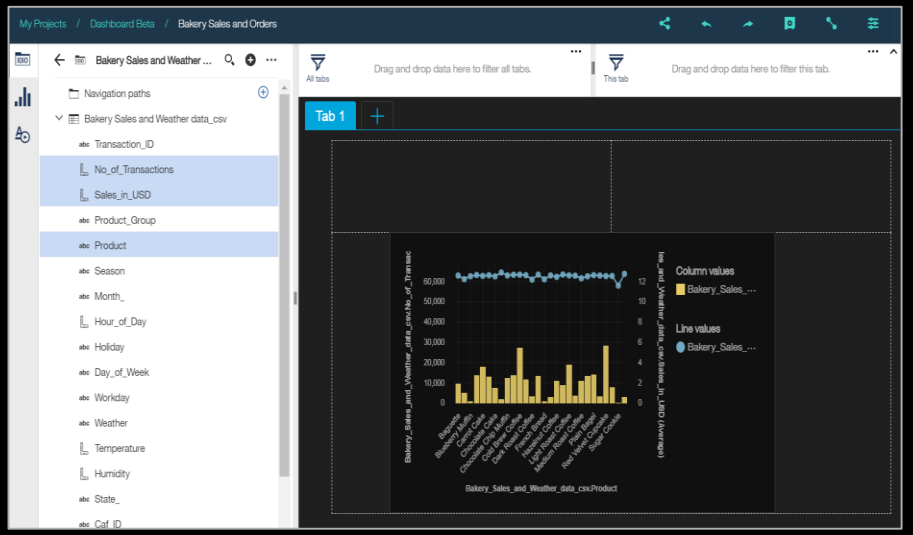
* https://arxiv.org/abs/1901.06261

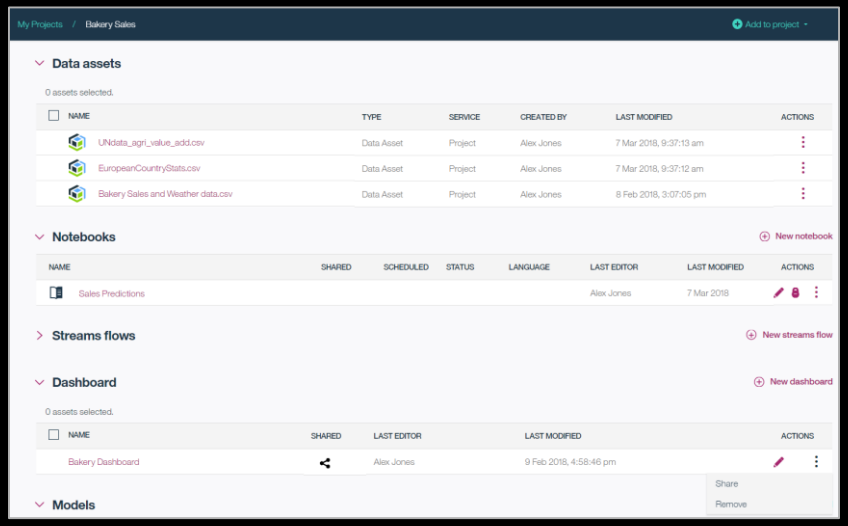
IBM Watson Studio Platform

Watson Studio Dynamic Dashboard Features

Making insights available to all







IBM Watson Studio Platform

Watson Studio Dynamic Take Away

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 – IBM Cloud Pak for Data OpenShift Container Platform, 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