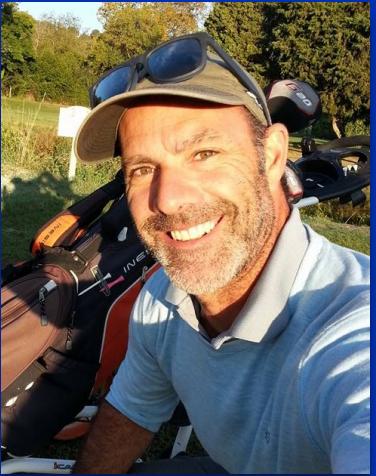


IBM Watson AI Platform

AI & Data Science



Emmanuel Génard, Data Scientist & Developer Advocate Europe
IBM Cloud
genard@fr.ibm.com |  @manuGenard |  <http://fr.linkedin.com/in/egenard>

Agenda

Day One - *November 21st*

- Introduction
- Flavors of Data Analytics
- An overview Artificial Intelligence
- Watson Studio tools
- ***Hands-on-Labs***

Day Two - *November 28th*

- Day one recap
- Prescriptive Analytics
- An overview of IBM Decision Optimization
- ***Hands-on-Labs***

http://ibm.biz/EDHEC_DataAnalytics

Different Flavors of Data Analytics

Descriptive Analytics

uses statistical models and forecasts techniques to understand the past and answer: **“What has happened?”**

Predictive Analytics

which uses data aggregation and data mining to provide insight and answer: **“What could happen?”**

Prescriptive Analytics

which use optimization and cognitive computing to advice on possible outcomes and answer: **“What should we do?”**

Different Flavors of Data Analytics

Descriptive Analytics

uses statistical models and forecasts techniques to understand the past and answer: **“What has happened?”**

Predictive Analytics

which uses data aggregation and data mining to provide insight and answer: **“What could happen?”**

Prescriptive Analytics

which use optimization and cognitive computing to advice on possible outcomes and answer: **“What should we do?”**

Different Flavors of Data Analytics

Descriptive Analytics

uses statistical models and forecasts techniques to understand the past and answer: **“What has happened?”**

Predictive Analytics

which uses data aggregation and data mining to provide insight and answer: **“What could happen?”**

Prescriptive Analytics

which use optimization and cognitive computing to advice on possible outcomes and answer: **“What should we do?”**

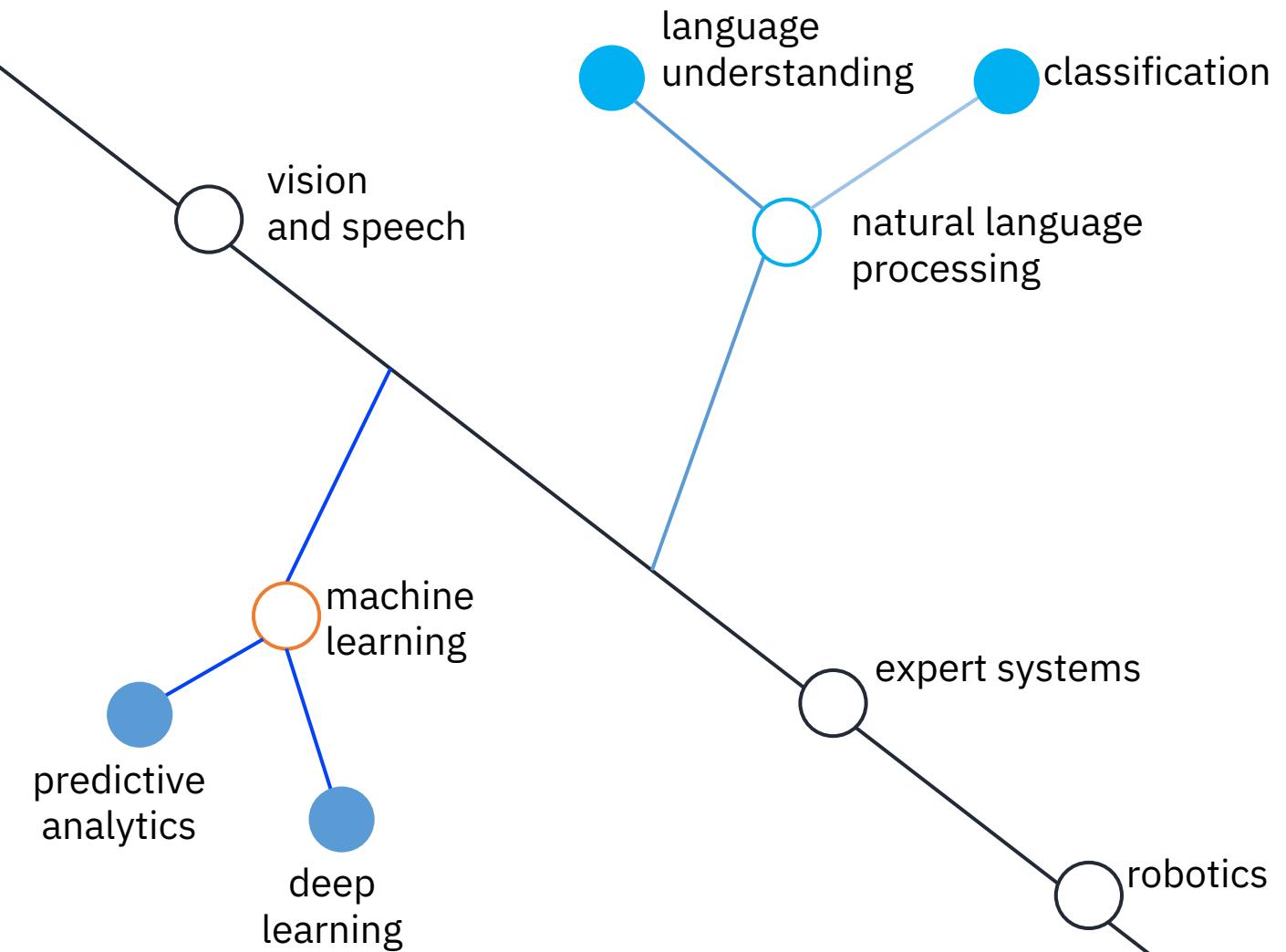
Can you think of a Prescriptive Analytics use case?



Artificial Intelligence

AI is the theory and development of computer systems to perform tasks that normally require human intelligence.

AI systems continuously learn using such techniques as machine learning and deep learning to solve problems in machine vision, speech and natural language.



Every Business and Enterprise is Embracing AI

Leaders everywhere
are monetizing data
and developing
strategies to embed
AI in business



Sales

Market basket analysis,
next-best offer,
customer churn,
propensity to buy, and
smart engagement.

Marketing

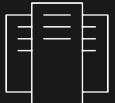
Discount targeting, email
optimization, and lifetime
client value, basket
recommendation
systems.

Human Resources

Medicare fraud, AI-
assisted diagnosis, and
drug demand
forecasting.

Supply Chain

Predictive maintenance,
process optimization,
and demand forecasting.



Energy & Utilities

Power usage prediction,
maintenance, and smart
grid management.

Finance

Customer segmentation,
credit risk, and credit
card fraud detection.

Security

Activity monitoring,
intrusion detection, and
log analysis.

Data & Technology

Dynamic pricing, call
center assistance,
tourism forecasting, and
self-driving cars.

Machine Learning is algorithm selection

Deep learning is neural network design

AI is systems architecture

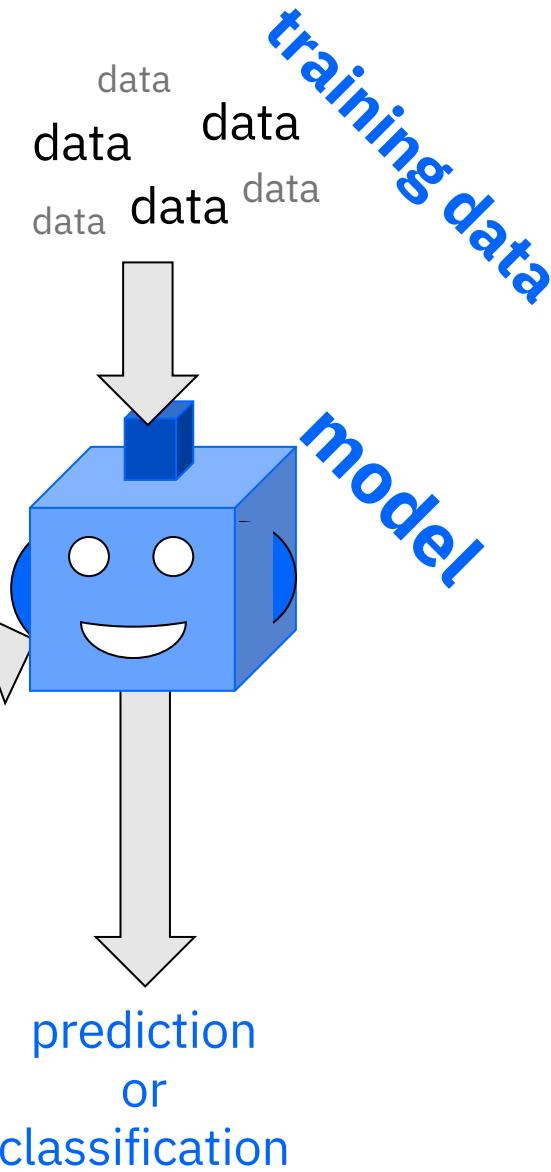
How does machine learning work?

Machine learning requires
TONS OF DATA

- ① A machine learning model is trained to recognize patterns in historical data

- ② The model is then shown new data and asked to predict or classify it.

- ③ If patterns in the new data match the training data then the model makes accurate predictions



AI Products & Services

AI Assistant

Integrate diverse conversation technology into your application.

Knowledge

Get insights through accelerated data optimization capabilities.

Vision

Identify and tag content then analyze and extract detailed information found in an image.

Speech

Convert text and speech with the ability to customize models.

Language

Analyze text and extract meta-data from unstructured content.

Empathy

Understand tone, personality, and emotional state.

Watson Assistant (formerly Conversation)

Lite • IBM

Watson Assistant a platform that allows developers and non-technical users to collaborate on building conversational AI-powered assistants.

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 being built, used, and

Compare Comply

IBM • Beta

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

Discovery

Lite • IBM

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

Teach Watson the language of your domain.

Language Translator

Lite • IBM

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

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 uses advanced natural language processing and machine learning techniques to create custom classification models. Users train

Natural Language Understanding

Lite • IBM

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

Personality Insights

Lite • IBM

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

Speech to Text

Lite • IBM

Low-latency, streaming transcription

Text to Speech

Lite • IBM

Synthesizes natural-sounding speech from text.

Tone Analyzer

Lite • IBM

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

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 custom

Watson Studio

Lite • IBM

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

PowerAI

Third Party

The accelerated deep learning platform for enterprise. Built on the IBM PowerAI platform, powered by Nimbix.

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

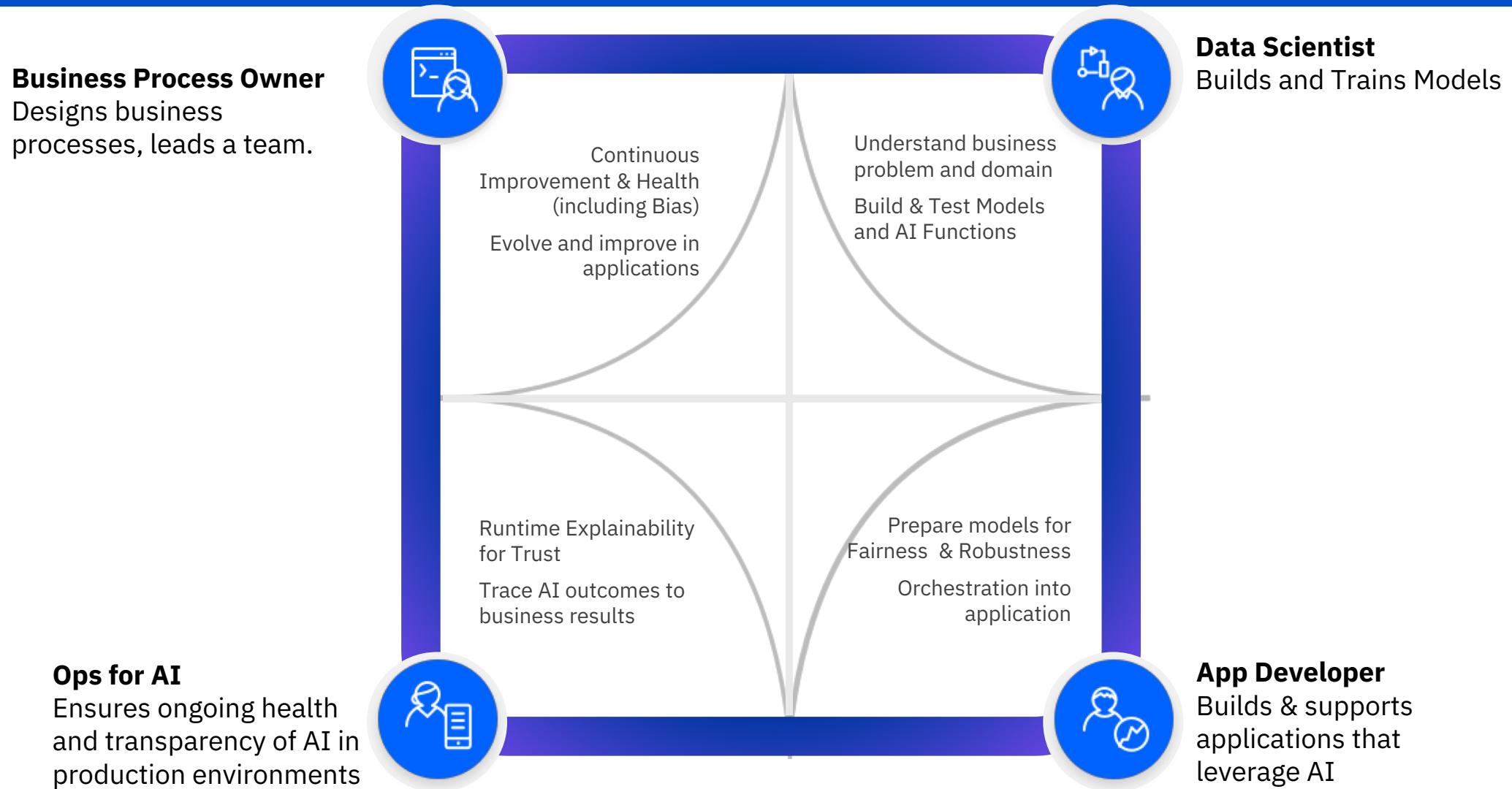
- Data Science skills are in low supply and high demand
- Nurturing new data professionals is challenging

Tools & Infrastructure

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

Watson AI Platform

Built for AI teams – enabling team productivity and collaboration



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

Supporting the end-to-end AI workflow

Connect & Access Data

Search and Find Relevant Data

Prepare Data for Analysis

Build and Train ML/DL Models

Deploy Models

Monitor, Analyze and Manage

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.

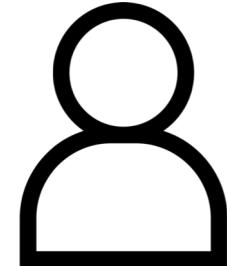
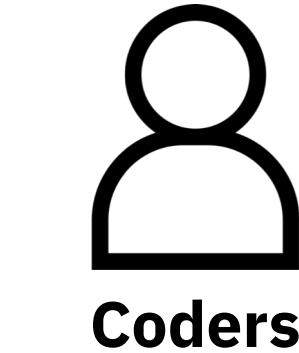
Democratize the creation of ML and DL models. Design your AI models **programmatically** or **visually** with the most popular **open source** and IBM ML/DL frameworks or leverage transfer learning on **pre-trained** models using **Watson tools** to adapt to your business domain. Train at scale on **GPUs**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

Two types of custom model builders



**Non-Coders
(Clickers)**



SPSS Modeler



Model Builder

Deep Learning



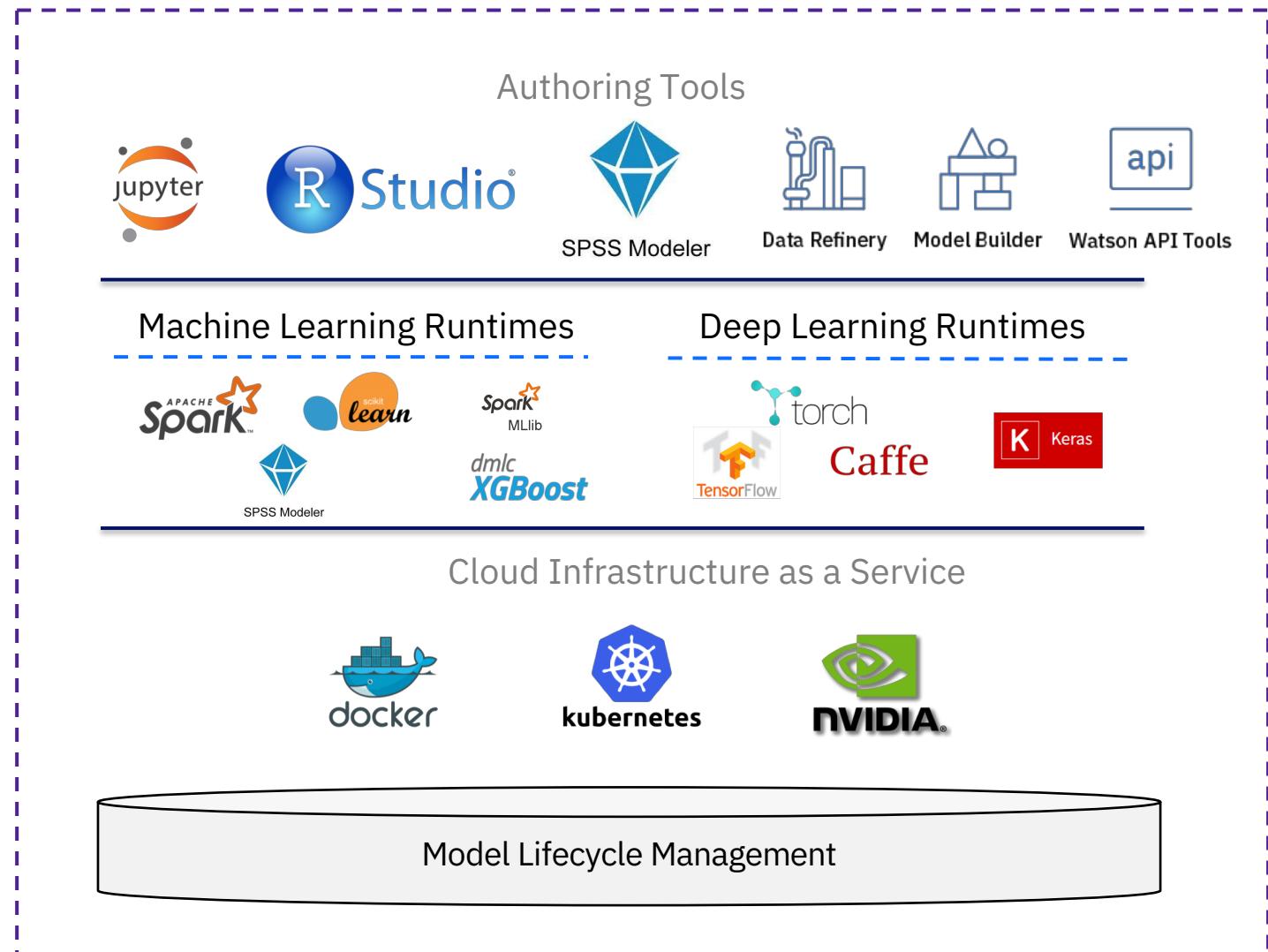
Caffe



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



Data Refinery

Making data fit for use

The screenshot shows the IBM Watson Data Preparation interface. At the top, there are navigation links for 'IBM Watson' (with a logo), 'Projects', 'Tools', 'Community', 'Services', and 'US South'. Below the header, the breadcrumb trail reads 'My Projects / Data Preparation / Customer_Orders_flow / Data Refinery'. On the right side, there are several icons for file operations like 'New', 'Open', 'Save', etc. The main area is titled '+ Operation' and contains the sub-tutorial 'Code an operation to cleanse and shape your data'. A table titled 'OPERATIONS' lists various data manipulation functions: 'arrange', 'count', 'distinct', 'filter', 'group_by', 'mutate', 'mutate_all', 'pull', 'push', 'select', 'summarise', and 'transform'. Each function has a brief description and a sample code snippet. To the right of the table, a sidebar titled '4 STEPS' provides a step-by-step guide:

- Data Source : Great Outd...**
- Remove duplicates**
Removed rows with duplicate values in STATE
- Sort ascending**
Sorted rows by CITY
- Sort descending**
Sorted rows by CUST_ORDER_NUMBER
- Sort ascending** JUST ADDED
Sorted rows by CITY

A large blue box at the bottom contains the text: "Self-service data refinement and cleaning".

IBM Watson Projects Tools Community Services US South 11:51

My Projects / Data Preparation / Customer_Orders_flow / Data Refinery

+ Operation Code an operation to cleanse and shape your data

Data Profile Visualizations Steps

| CUST_ORDER_NUMBER | COUNTRY | STATE | CITY |
|-------------------|----------------|-----------|-----------|
| String | String | String | String |
| FREQUENCY | FREQUENCY | FREQUENCY | FREQUENCY |
| 1371938 | United States | MA | Singapore |
| 172499 | China | CA | Chengdu |
| 174806 | Germany | IA | Brisbane |
| 100131 | Japan | Singapore | Auckland |
| 101146 | Brazil | PR-RJ | Beijing |
| 101937 | United Kingdom | NY | Sydney |
| 102004 | France | IL | Melbourne |
| 104913 | Canada | IN | Paris |
| 107726 | Australia | PR-SP | Brussels |
| 1190420 | Italy | SA | Guangzhou |

STATISTICS STATISTICS STATISTICS STATISTICS

Maximum length 6 Maximum length 14 Maximum length 9 Maximum length 25

Minimum length 1 Minimum length 100 Minimum length 2.3778 Mean length 8.1706

Mean length 768 Unique

Comprehensive profiling

The screenshot shows the IBM Watson Data Refinery interface. On the left, there's a sidebar with tabs for 'Input' (selected), 'Data', 'Profile', and 'Visualizations'. The 'Input' tab shows 'Columns: COUNTRY, PRODUCT_LINE' and 'Chart types: Treemap'. Below that is a 'Brunel syntax' box containing code:

```
tree x#PRODUCT_LINE size(#count) label(COUNTRY)  
tooltip(#all)
```

. The main area displays a treemap visualization where three categories—'Camping Equipment', 'Golf Equipment', and 'Mountaineering Equipment'—are represented as large blue rectangles. Each rectangle is further subdivided into smaller blue rectangles representing countries: Australia, Canada, Italy, Spain, and United Kingdom. A large blue callout box at the bottom contains the text 'Interactive visualization'.

IBM Watson Projects Tools Community Services US South 1:11

My Projects / Data Preparation / Customer Orders_flow Refine ▶ ⓘ ⓘ ⌂ ⌂

Summary

Source ⓘ Data flow Output

| | | |
|-----------------------------------|---------|----------------------------|
| Great Outdoor Customer Orders.csv | 3 Steps | Customer Orders_shaped.csv |
|-----------------------------------|---------|----------------------------|

Runs

History Schedule

| DATE | DAY | TIME | Summary |
|-------------|-----|---------|-------------------------|
| 6 Mar 2018 | Tue | 9:58 pm | |
| 7 Mar 2018 | Wed | 9:58 pm | START 6 Mar 2018 |
| 8 Mar 2018 | Thu | 9:58 pm | INTERVAL Every 1 day |
| 9 Mar 2018 | Fri | 9:58 pm | END 11 Mar 2018 9:58 pm |
| 10 Mar 2018 | Sat | 9:58 pm | |
| 11 Mar 2018 | Sun | 9:58 pm | |

Scheduling and monitoring

Watson Studio

The screenshot shows a Jupyter notebook environment within Watson Studio. The code cell contains Python code to generate a bar chart titled 'Tweets Country Distribution based on the User Profile'. The chart displays tweet counts for various countries. A comment from 'ARMAND RUIZ G...' is visible, saying 'Great work!'. The interface includes standard Jupyter notebook controls like Run, Cell, Kernel, and Help.

Not Trusted | Python 3.5

```
colors = ['gray'] + colors
lt.figure(figsize=(10,8))
lt.barh(y_pos, num_tweets, align='center', color=colors)
lt.yticks(y_pos, countries)
lt.xlabel('Number of Tweets')
lt.title('Tweets Country Distribution based on the User Profile')
lt.ylim(-1, len(y_pos))
lt.show()
```

Tweets Country Distribution based on the User Profile

GERMANY
MEXICO
CANADA
INDIA
JAPAN
SPAIN

Add a comment

ARMAND RUIZ G... 7:49 AM
Great work!

Open Source tools – Jupyter and RStudio

The screenshot shows the 'Environments' tab in Watson Studio. It provides a overview of active environment runtimes and allows users to define new environment definitions. A search bar at the top asks 'Which environment are you looking for?'. The 'Active environment runtimes' section indicates 'you currently have no active environment runtimes'. The 'Environment definitions' section lists a single entry: 'Default SPSS Modeler XS' with 'Wml_flow' as the tool, '2 vCPU and 8 GB RAM' as hardware configuration, and '2 Mar 2018' as the last modified date.

+ Add to project

Overview Assets Environments Bookmarks Deployments Collaborators Settings

Environments

Define the hardware size and software configuration for the runtime associated with DSX tools such as notebooks.

In this release you can use default environments or you can create custom environments with different hardware and software configurations. [Learn more](#).

Which environment are you looking for?

Active environment runtimes

| NAME | CAPACITY UNIT HOURS | OWNER | TOOL | HARDWARE CONFIGURATION | STARTED AT | ACTIONS |
|---------------------------------------------------|---------------------|-------|------|------------------------|------------|---------|
| you currently have no active environment runtimes | | | | | | |

Environment definitions

New environment definition

| NAME | TOOL | HARDWARE CONFIGURATION | LAST MODIFIED | ACTIONS |
|-------------------------|----------|------------------------|---------------|---------|
| Default SPSS Modeler XS | Wml_flow | 2 vCPU and 8 GB RAM | 2 Mar 2018 | ⋮ |

Elastic and customizable compute environments

The screenshot shows the Watson Visual Recognition interface for retraining a model. It features a 'Car recognition' project with a 'My Classes' tab displaying car categories like 'AMGeneralHummerS...', 'AcuraTSX...2012', and 'AcuraIntegraTypeR...'. Below this, a grid of car images is shown with labels such as 'AcuraTL...2012', 'AstonMartinV8Vanta...', 'Audi100Sedan1994', and 'AstonMartinV8Vanta...'. On the right, a sidebar lists uploaded ZIP files containing image datasets. A central panel shows a 'Train Model' button and a placeholder for dropping a .zip file.

Unclassified images: 0 Incomplete classes: 0

Train Model

Drop a .zip file of images here to add them to your project or browse to select a .zip file.

AMGeneralHummerS... 89 images

AcuraTSX...2012 81 images

AcuraIntegraTypeR... 89 images

Create a class

Search classes

49 classes

AcuraTL...2012 86 images

AstonMartinV8Vanta... 82 images

Audi100Sedan1994 81 images

AstonMartinV8Vanta... 90 images

Watson Visual Recognition – retrain Watson

The screenshot shows the 'ML flows' interface in Watson Studio. It displays a visual workflow for creating machine learning models. The flow starts with 'Image Data' input, followed by 'Conv 2d', 'ReLU', and 'Pooling 2d' layers. The output then branches into 'Flatten', 'Dense', 'Softmax', and 'Accuracy' layers. Finally, it concludes with 'Sigmoid Cross-E...' and 'SGD' layers. The interface includes standard workflow controls like Run, Cell, Kernel, and Help.

Single Convolution layer on M

```
graph LR; ID[Image Data] --> C1[Conv 2d]; C1 --> R1[ReLU]; R1 --> P1[Pooling 2d]; P1 --> F1[Flatten]; F1 --> D1[Dense]; D1 --> S1[Softmax]; S1 --> A1[Accuracy]; A1 --> SC[Sigmoid Cross-E...]; SC --> SG[SGD]
```

Watson Visual Recognition – retrain Watson

Create ML flows and design Neural Networks visually

Watson Knowledge Catalog

Unlock tribal knowledge and unleash your knowledge workers

Browse Assets Usage Statistics Access control Settings

What assets are you looking for?

Recently Added

| | | | | | |
|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| US Airlines Owner: Jay Limburn Added: 1 Mar 2018 13:04 Airline | Harry Rosen example Owner: Michael Tucker Added: 28 Feb 2018 17:30 Sales Forecast | Machine Learning using R Owner: Jay Limburn Added: 29 Feb 2018 14:55 notebooks... dxs | FDIC Failed Bank List Owner: Jay Limburn Added: 19 Feb 2018 11:25 banking | 2017 Small Business Banking Loan s Owner: Jay Limburn Added: 19 Feb 2018 13:36 banking loan | All US Banks Owner: Jay Limburn Added: 19 Feb 2018 13:36 fss banking |
|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|

Filter

Asset types

- Data Asset (56)
- Notebook (4)
- Connection (2)
- dashboard (2)

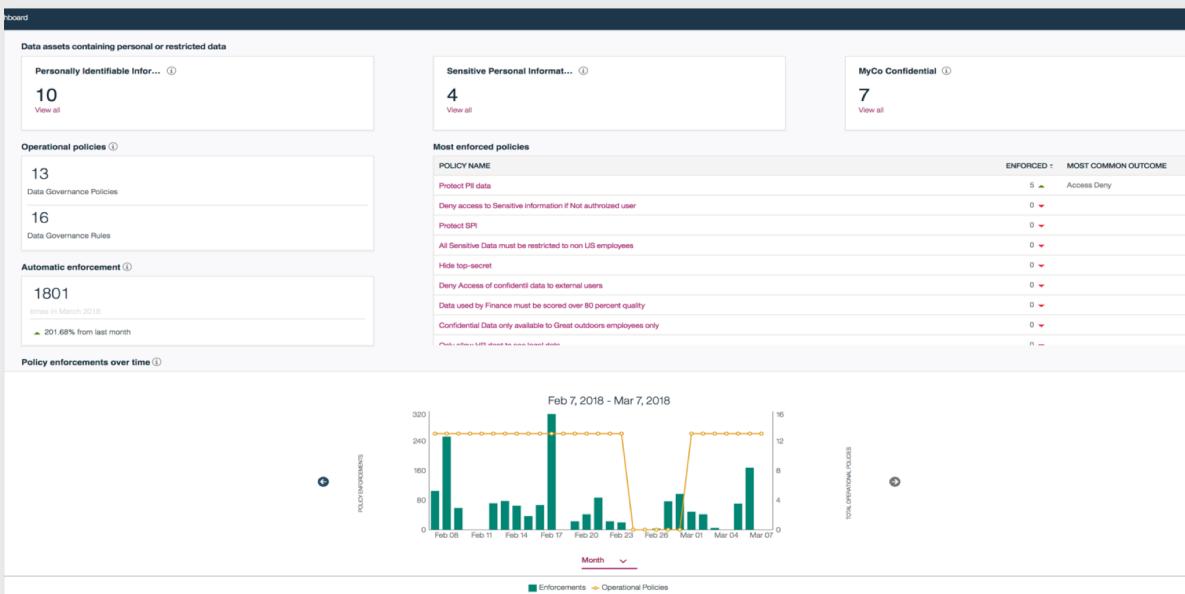
Tags

- discovered (24)
- SAMPLES (1)
- untagged (6)
- GOSALES (6)
- sales (5)
- banking (4)
- dxs (4)
- notebook (4)
- GOSALESDW (3)

Available Assets

Showing 64 of 64 assets

| NAME : | OWNER | TAGS | TYPE | DATE ADDED |
|-----------------------------------|-------------|-------------------------|------------|-------------|
| 2017 Small Business Banking Loans | Jay Limburn | banking loan | Data Asset | 19 Feb 2018 |
| ANCESTRY | Jay Limburn | discovered SAMPLES | Data Asset | 13 Dec 2017 |
| AWS Data Warehouse | Jay Limburn | | Connection | 13 Dec 2017 |
| All US Banking Branches | Jay Limburn | fss banking branches | Data Asset | 19 Feb 2018 |
| Banking Demographic Data | Jay Limburn | banking client customer | Data Asset | 19 Feb 2018 |
| BlobPower_T.csv | paul taylor | | Data Asset | 22 Jan 2018 |
| CONVERSION_RATE | Jay Limburn | discovered GOSALES | Data Asset | 13 Dec 2017 |
| COUNTRY | Jay Limburn | discovered GOSALES | Data Asset | 13 Dec 2017 |
| CUST_WIN_BACK | Jay Limburn | discovered SAMPLES | Data Asset | 13 Dec 2017 |
| Customer Orders Q3 | Jay Limburn | sales Warehouse | Data Asset | 11 Jan 2018 |



Browse Assets Change Statistics Access control Settings

Total assets ①

2
Connections
56
Data Assets
4
Notebooks
2
Other
64
Total

Deleted assets ①

< Showing MAR 2017 - MAR 2018 >

Number of Assets

12
9
6
3
0

APR JUN AUG OCT DEO FEB

MONTHS ▼

Deleted assets

Added assets ①

< Showing MAR 2017 - MAR 2018 >

Number of Assets

100
75
50
25
0

APR JUN AUG OCT DEO FEB

MONTHS ▼

Added assets

Assets accessed ①

March 2018

Number of Assets

200
150
100
50
0

1 2 3 4 5 6

Assets accessed Total assets

Business Glossary / MyCo Confidential

[Overview](#) [Related content](#)

MyCo Confidential

All confidential data for our entire knowledge set

Term details

Creator: jay@uk.ibm.com
Date created: 29 Sep 2017
Last editor: IBMid-2700028UJJ
Last modified: 7 Mar 2018

Owner
Jay Limburn

Tags
Compliance | Protection

Associated classifier or term
Confidential

Views over time

| Date | Number of Views |
|--------|-----------------|
| 9 Feb | 4 |
| 10 Feb | 2 |
| 11 Feb | 0 |
| 12 Feb | 1 |
| 13 Feb | 3 |
| 14 Feb | 2 |
| 15 Feb | 0 |
| 16 Feb | 1 |
| 17 Feb | 0 |
| 18 Feb | 0 |
| 19 Feb | 2 |
| 20 Feb | 4 |
| 21 Feb | 0 |
| 22 Feb | 0 |
| 23 Feb | 0 |
| 24 Feb | 0 |
| 25 Feb | 0 |
| 26 Feb | 0 |
| 27 Feb | 0 |
| 28 Feb | 0 |
| 1 Mar | 4 |
| 2 Mar | 0 |
| 3 Mar | 0 |
| 4 Mar | 1 |
| 5 Mar | 3 |

Description

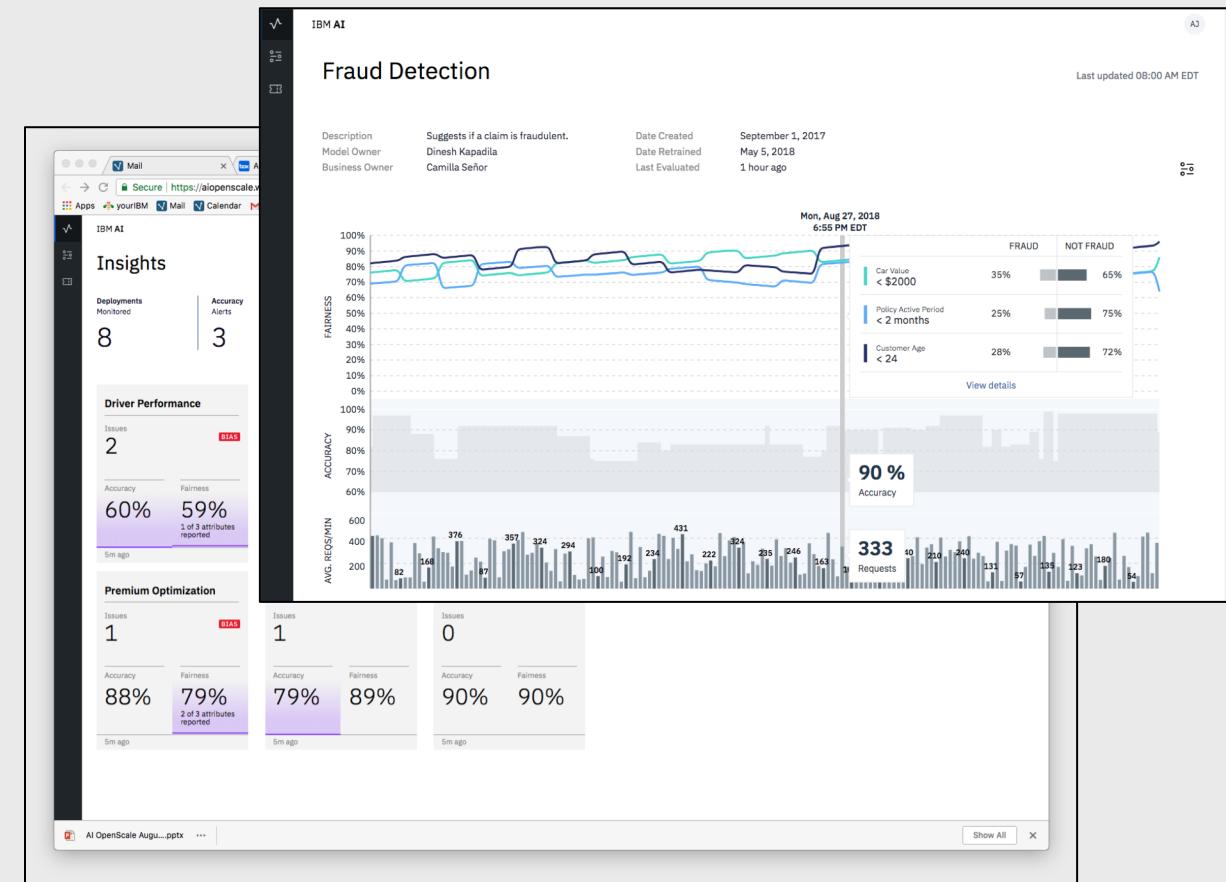
A term used by Great Outdoors to classify any information that should not be disclosed outside of the Great Outdoors company.

Introducing Trust and Transparency for AI

Infuse Visibility and Insights into your AI Solutions

The challenge business' face today is how they scale AI across the organization while making sure they can explain and trust the decisions & outcomes the technology is making

Trust and Transparency for AI are SW capabilities which address the AI “black box” problem. It provides organizations visibility into how AI assets are built, used, and scaled across organizations.



Trusted

Secure and Explainable Outcomes
Transparency into AI Infused Applications that are auditable and traceable for confident outcomes

Intelligent

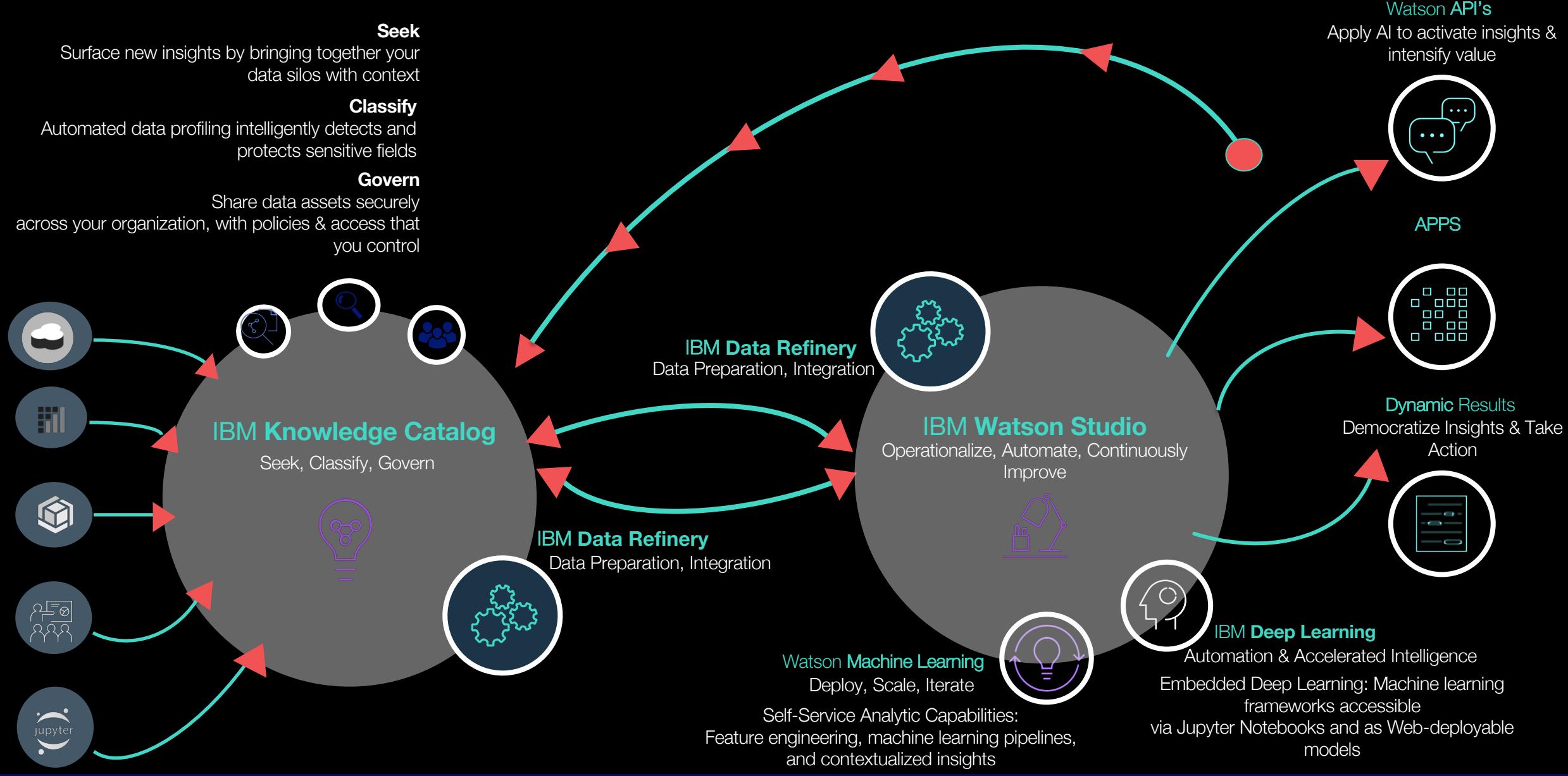
Fully managed, continuously evolving system
Dynamically & continuously improving to deliver superior business outcomes

Scalable

AI Infused Business Requires New Operating Models

Runs where your data and applications reside with visibility into how AI is used and performing at scale of the enterprise

Watson Data & AI for the Enterprise



A Toolkit for developers to get started with IBM Cloud and Watson today

Tech talks, Open source, Tutorials ...

- ibm.com/developerworks
- developer.ibm.com/code/
- medium.com/ibm-watson
- ibm.com/watson/
- console.bluemix.net/docs
- dataplatform.ibm.com
- cognitiveclass.ai

Events

- developer.ibm.com/code/community/events

Thank you !