avanade

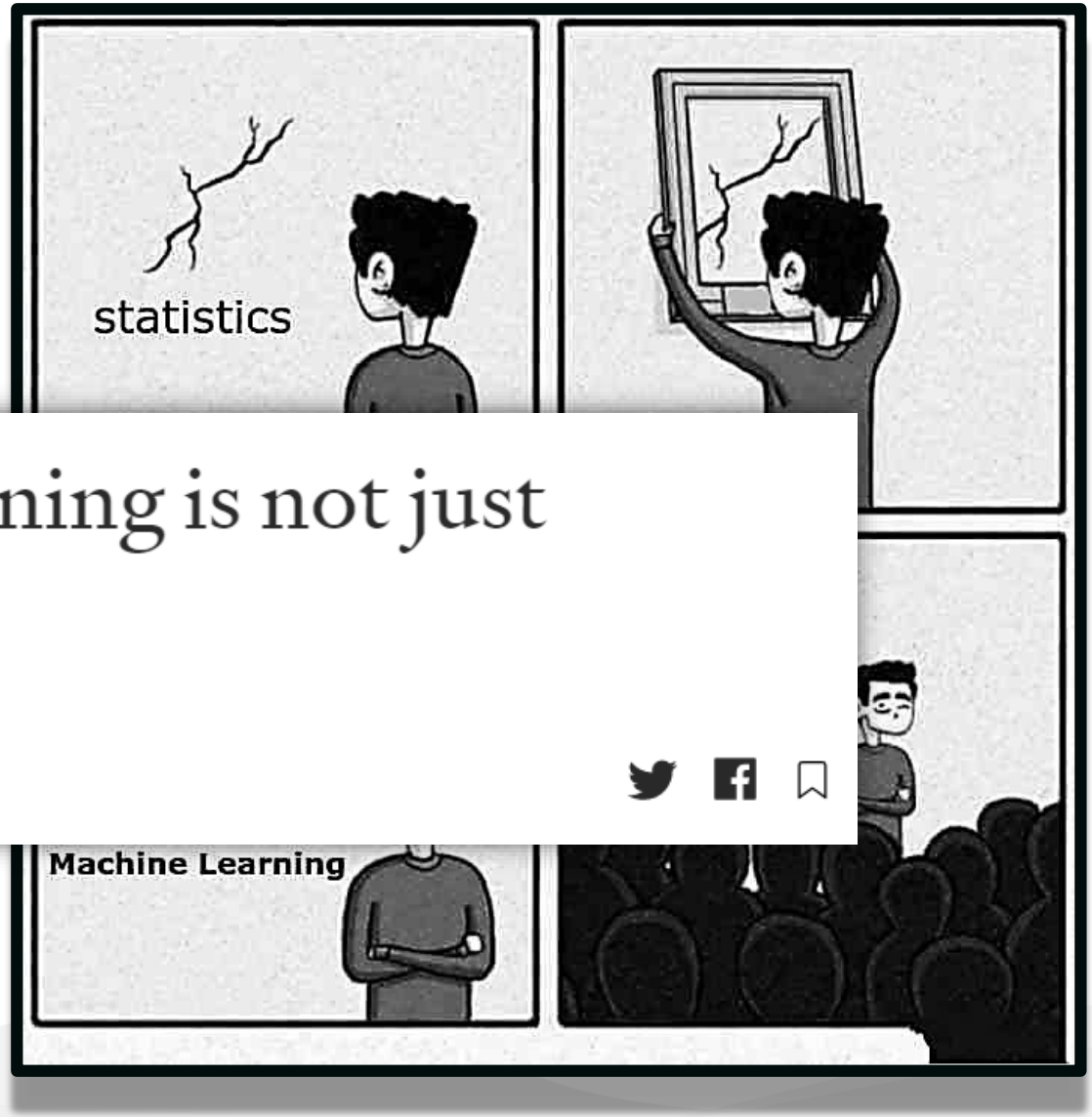
Business & Data Science: Challenges ahead

What companies need, in 2020, regarding Statistics, Data and Analytics?

uc3m

Universidad Carlos III de Madrid

A year ago …



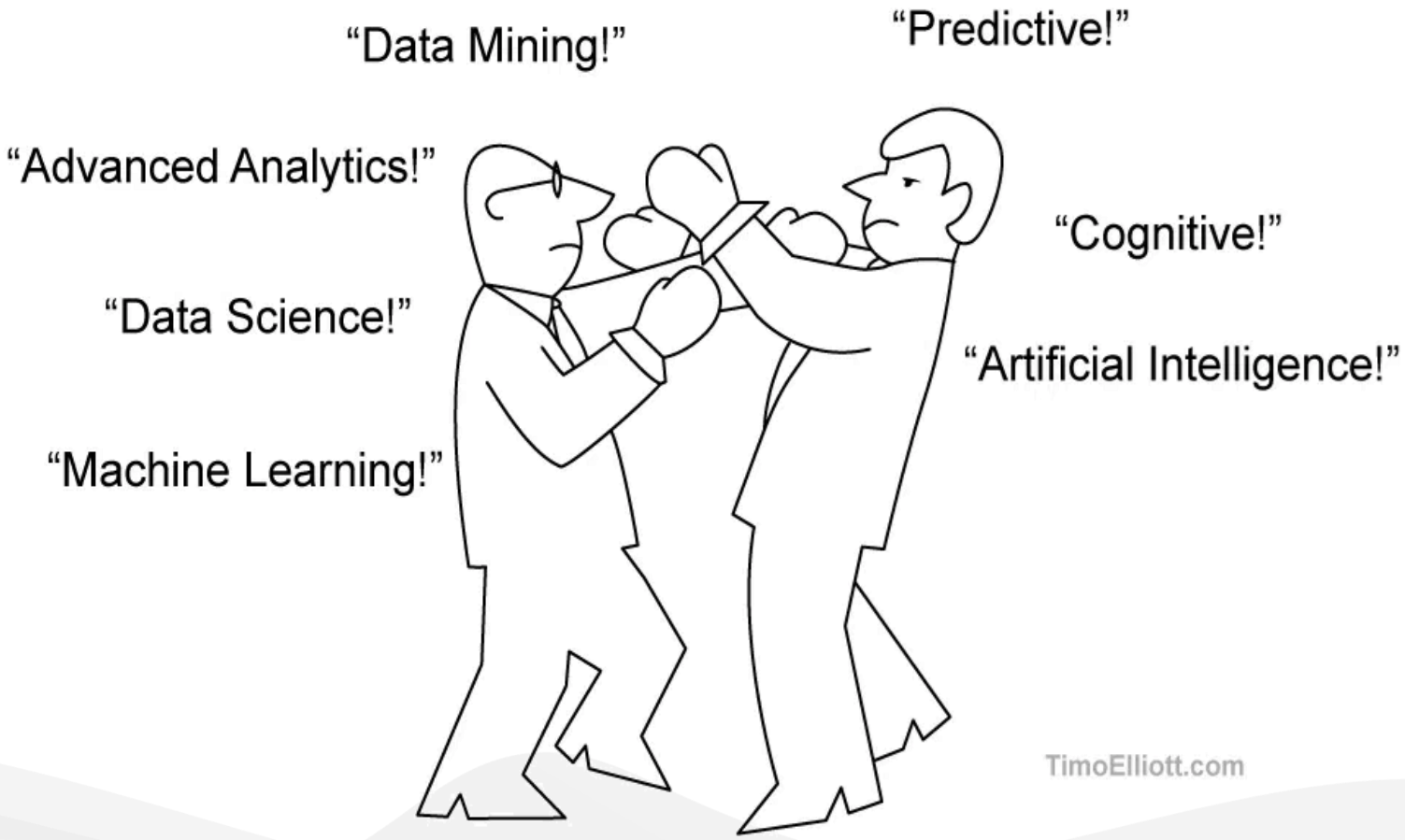
No, Machine Learning is not just glorified Statistics

Joe Davison Follow

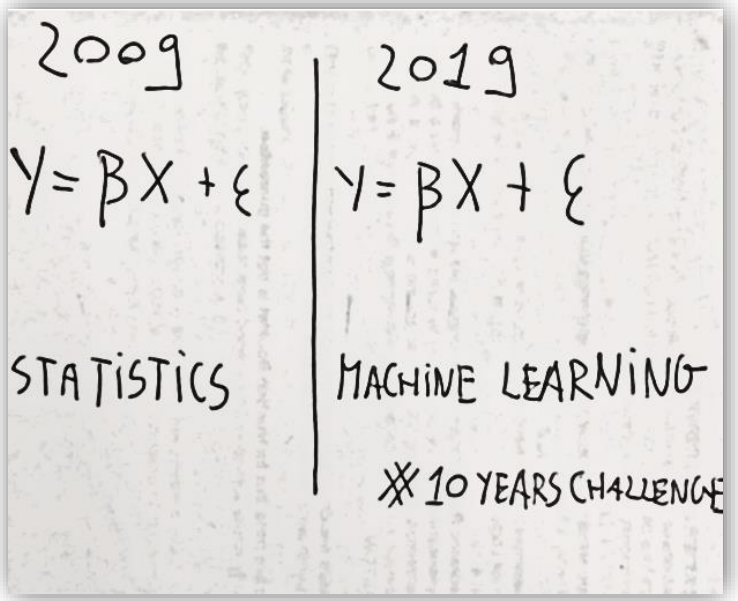
Jun 27, 2018 · 10 min read ★

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Unfortunately …



It's basically the same thing, with slightly different approaches !!!???



avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

But …

Forbes

5,488 views | Jan 19, 2020, 02:22pm

Roundup Of Machine Learning Forecasts And Market Estimates, 202

Louis Columbus Contributor
Enterprise Tech

- **One in ten enterprises now use ten or more AI applications; chatbots, process optimization and fraud analysis lead a recent**

- **83% of IT leaders say AI & ML is transforming customer engagement, and 69% say it is transforming their business.** 79% believe that AI will help their organization identify external and internal securi
Salesf
Resea

- **The global machine learning market is projected to grow from $7.3B in 2020 to $30.6B in 2024, attaining a CAGR of 43%.** AI-based proce
contri
Futur
2024.

- **IDC predicts spending on AI systems will reach $97.9B in 2023, more than two and one-half times the $37.5B that will be spent in 2019.** The compound annual growth rate (CAGR) for the 2018-2023 forecast period will be 28.4%. Spending on AI systems will be led by the retail and

Open jobs on LinkedIn requesting machine learning expertise in the U.S. further reflect its growing dominance in all businesses. There are 44,864 jobs in the U.S. today according to LinkedIn that list machine learning as a required skill, and 98,371 worldwide.

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Challenges ahead?

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

01

Who is using my models?

ginablaber
@ginablaber

The story of enterprise Machine Learning: “It took me 3 weeks to develop the model. It’s been >11 months, and it’s still not deployed.” @DineshNirmalIBM #StrataData #strataconf

Traducir Tweet
7:19 p. m. · 7 mar. 2018 · TweetDeck

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

02

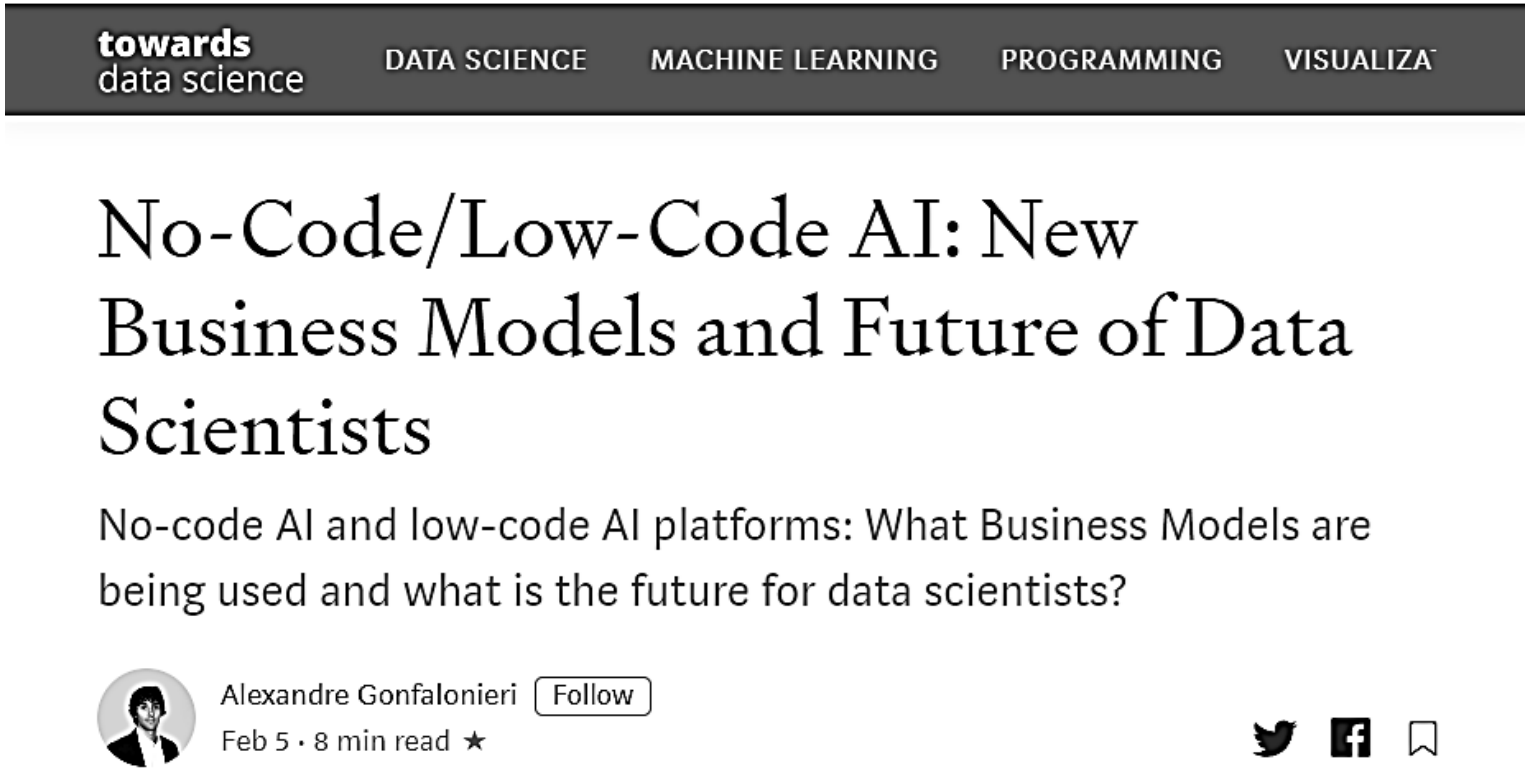
What is my real role?

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

03

Are Data Scientists essential?

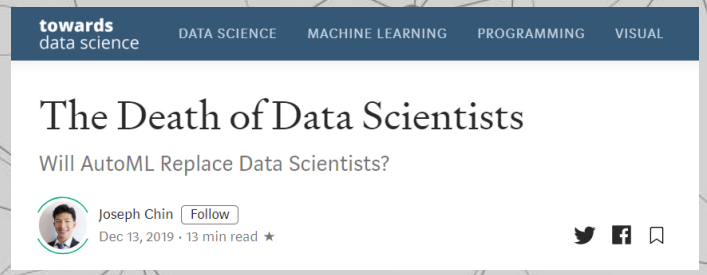


avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Death of Data Scientists

"Kaggle is to real-life machine learning as chess is to war. Intellectually challenging and great mental exercise, but you don't know, man! You weren't there!" – Lukas Vermeer



How can you build models that have an impact on the most vital areas of your brand, so that they bring value to your business?

avanade

Keywords

Industrialization

MLOps

"Let's put in production … automate and accelerate the machine learning lifecycle"

Democratization

Citizen DS

"Enable the people closest to business challenges to resolve them using intelligent apps"

Productivity

Auto ML

"… empowers customers, with or without data science expertise, to identify an end-to-end machine learning pipeline for any problem, achieving higher accuracy while spending far less of their time.

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

MLOps

Machine Learning Operations

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

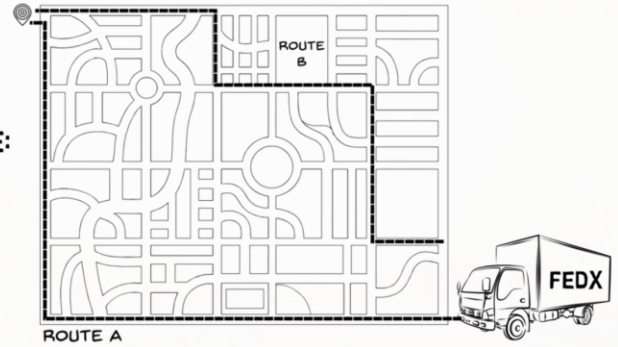
BUSINESS PROBLEM

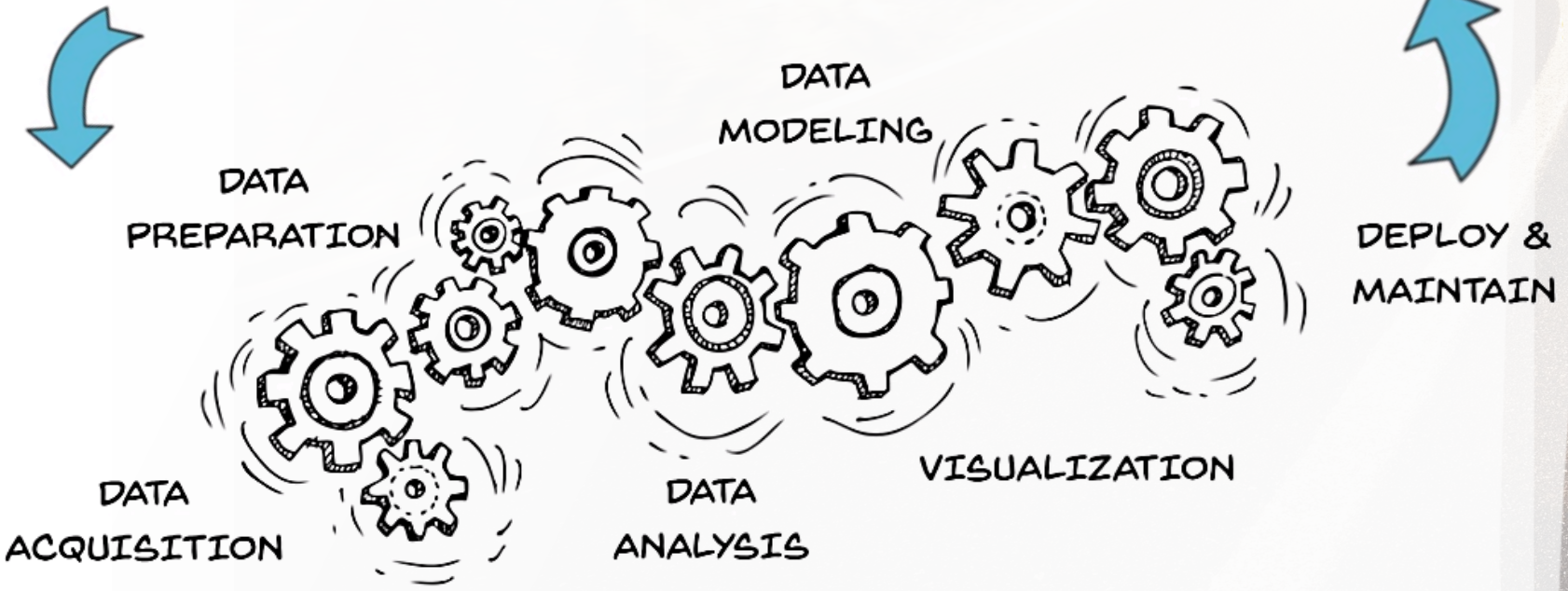
WHY?....WHY?....WHY?....

DISCOVERED THE BEST TIME AND ROUTES TO SHIP

BEST TIME:

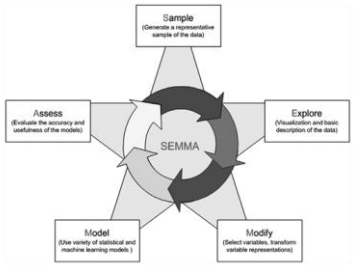
12:00PM



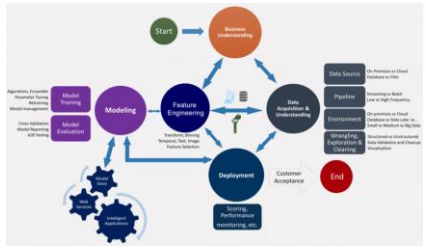


CRISP-DM

SEMMA



TEAM DS PROCESS



avanade

THE DATA SCIENCE PROCESS

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Demand Forecasting

Predictive Maintenance

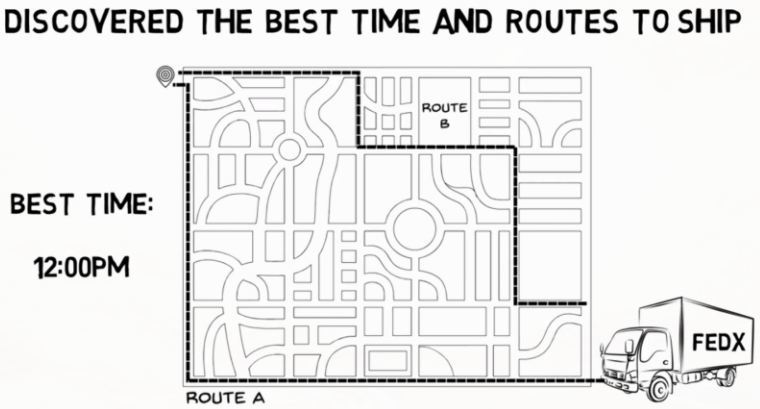
Delay Predictions

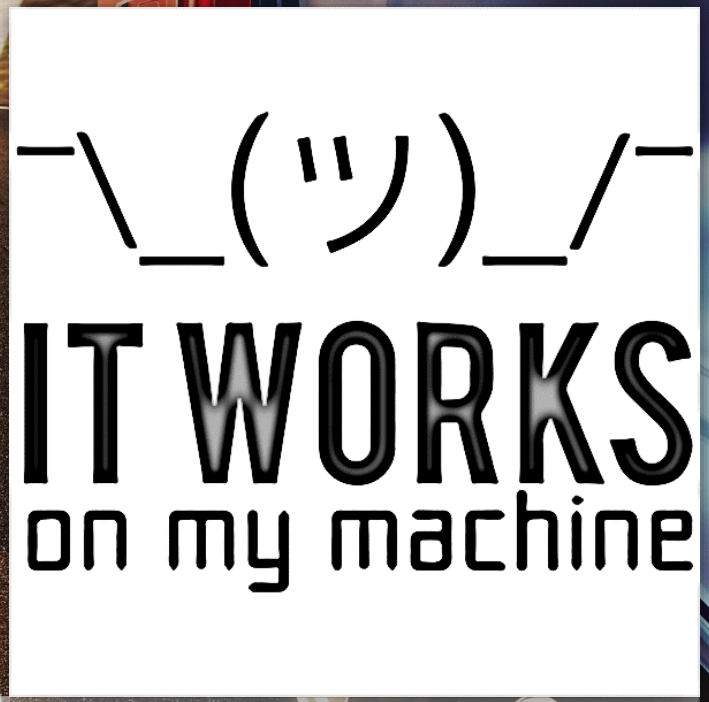
Mails Classification

Inventory Management

Staff Scheduling

IoT - Waves of data!





Dataset Dependency | Simple to complex pipelines | Scalability |
Monitoring and Evaluation | Real-time scoring | Process and Collaboration

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

"There's a huge difference between building a Jupyter notebook model in the lab and deploying a production system that generates business value."

Andrew Ng. Artificial Intelligence
Stanford University. Coursera & deeplearning.ai co-founder

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

MLOps: to put an AI application into production

AI/ML teams know how to develop models that can transform a business

Software development teams know how to deliver business applications and cloud services

TDSP

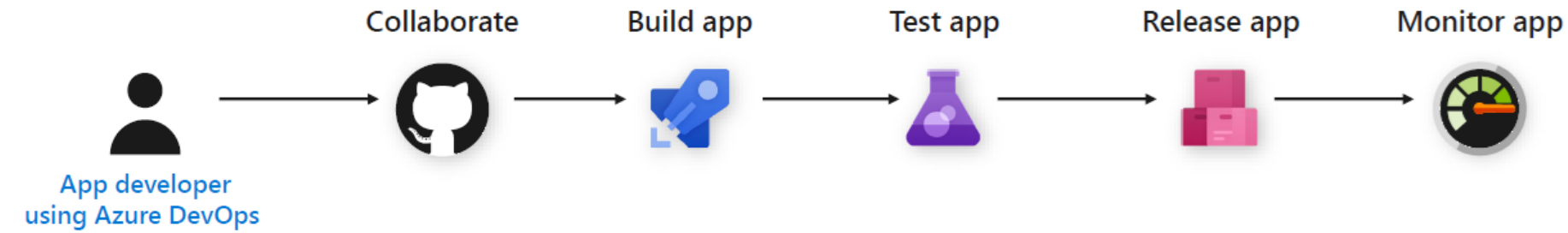
DevOps

> *"The difficult part about integrating AI or ML into an application is not the technology, or the math, or the science or the algorithms. The challenge is getting the model deployed into a production environment and keeping it operational and supportable."*
>
> Kate Baroni Software Architect, Microsoft Azure

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

MLOps with Azure Machine Learning

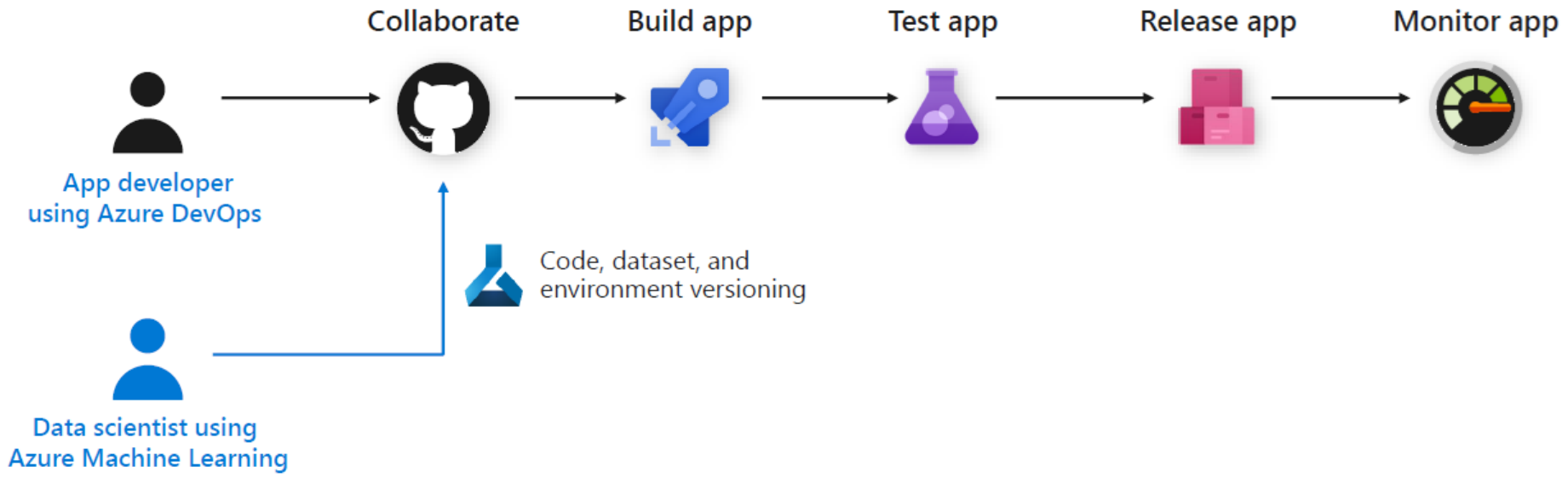


- ☐ Model reproducibility
- ☐ Model validation
- ☐ Model deployment
- ☐ Model retraining

avanade

© 2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

MLOps with Azure Machine Learning



- [x] Model reproducibility
- [ ] Model validation
- [ ] Model deployment
- [ ] Model retraining

Step 1 of 4

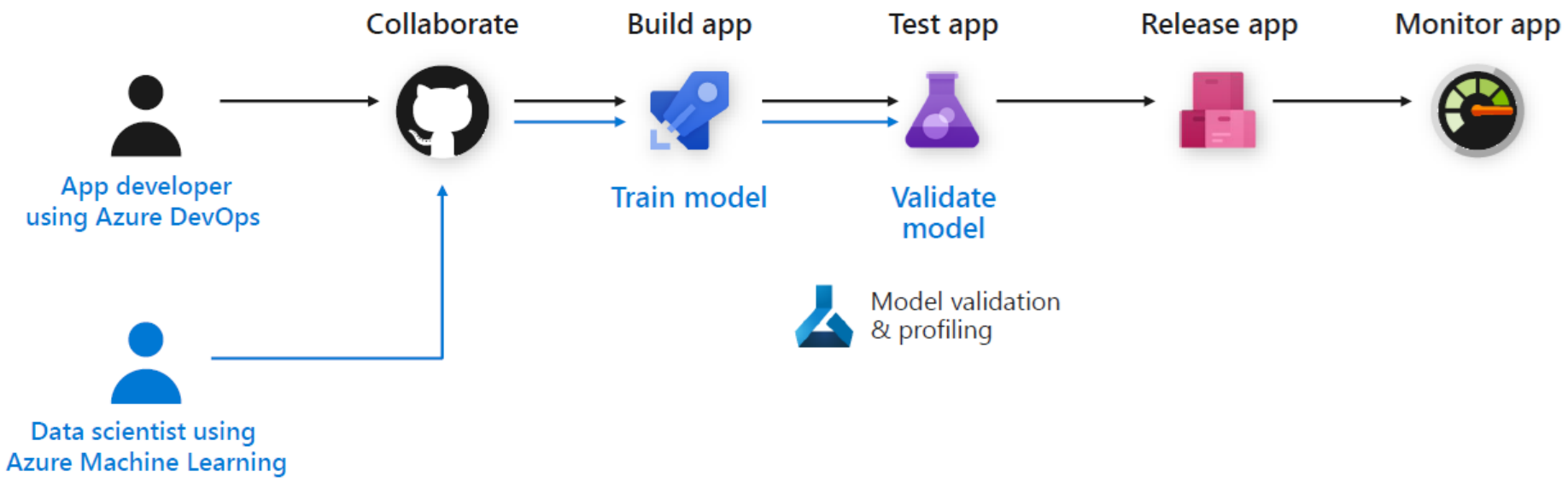
Build and train reproducible models

Turn your training process into a reproducible pipeline using machine learning pipelines to stitch together all the steps, from data preparation to model evaluation.

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

MLOps with Azure Machine Learning



- [x] Model reproducibility
- [x] Model validation
- [ ] Model deployment
- [ ] Model retraining

Step 2 of 4

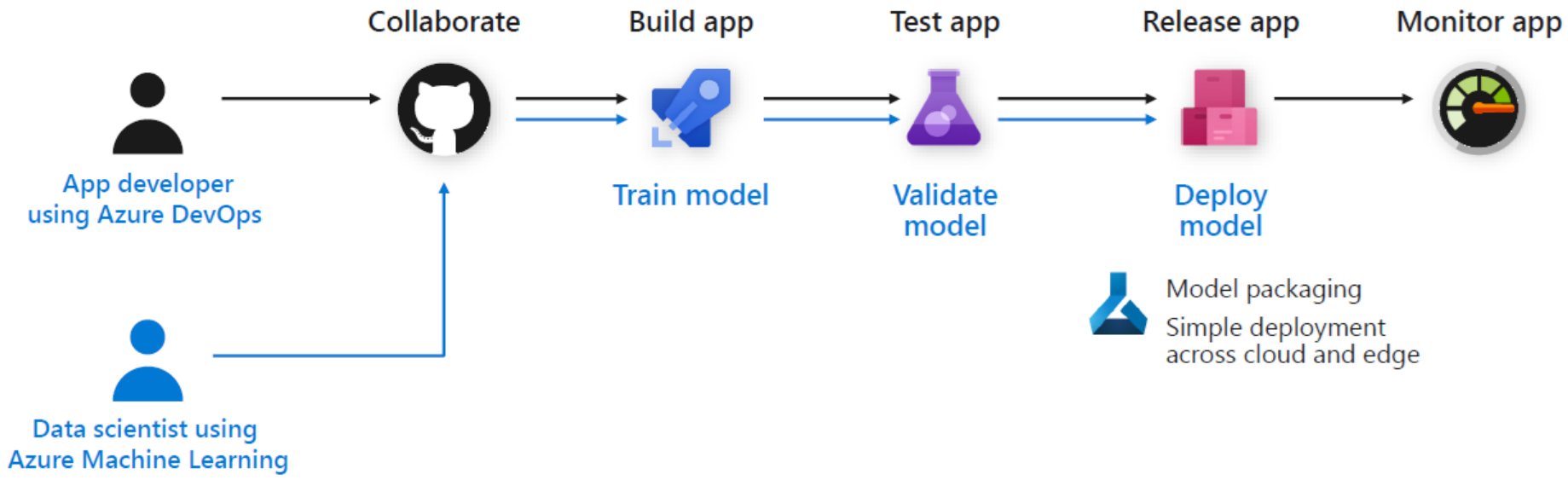
Validate and Profile models

Register the model

avanade

© 2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

MLOps with Azure Machine Learning



- [x] Model reproducibility
- [x] Model validation
- [x] Model deployment
- [ ] Model retraining

Step 3 of 4

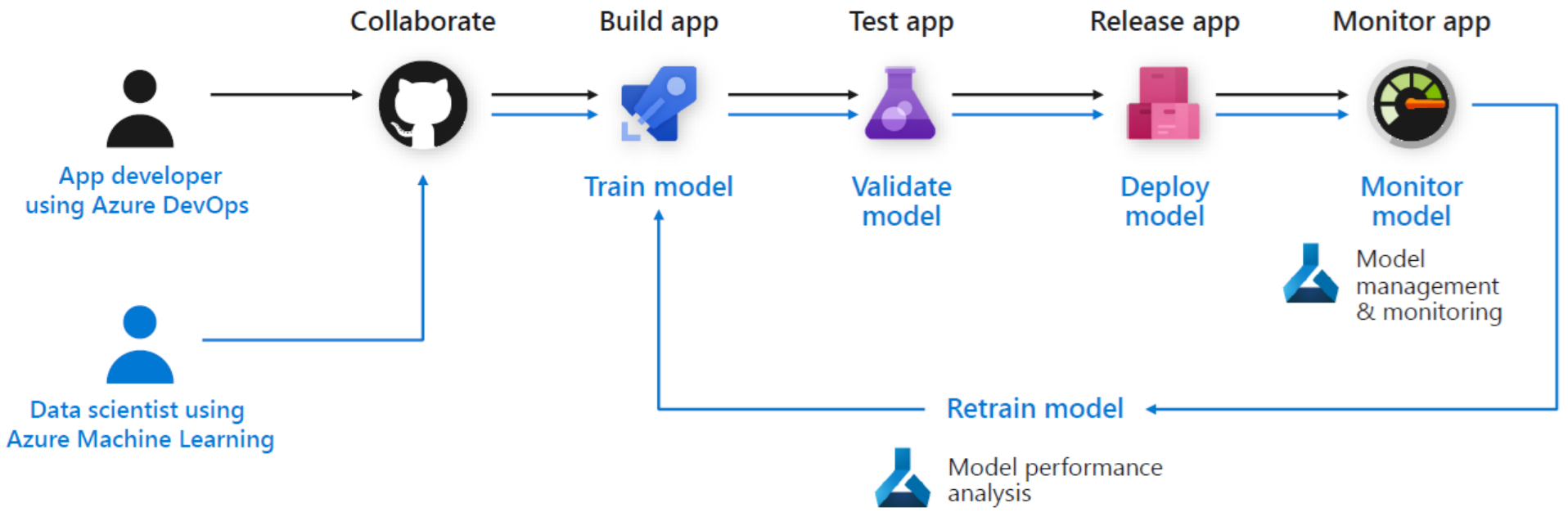
Package and deploy models

Package the model into a container image and then deploy it. Use profiling to determine the ideal CPU and memory settings, and to validate models.

avanade

© 2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

MLOps with Azure Machine Learning



Model reproducibility | Model validation | Model deployment | Model retraining

Step 4 of 4

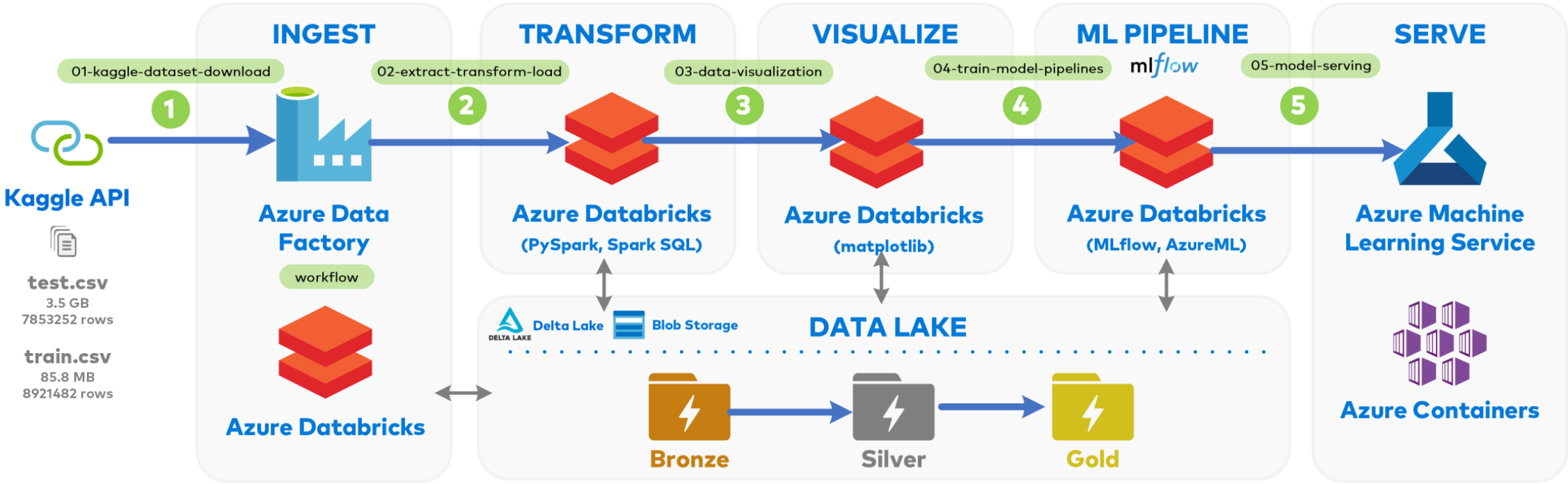
Automate workflows, monitor, and manage

Automate the end-to-end machine learning lifecycle with Azure Machine Learning and GitHub to frequently update models, test new models, and continuously roll out new machine learning models alongside your other applications and services.

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Understanding MLOps with Azure Databricks



avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Citizen Data Scientist

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

*"… a **person** who creates or generates **models** that uses **advanced** diagnostic **analytics** or predictive and prescriptive capabilities, but whose primary job function is **outside the field of statistics and analytics**."*

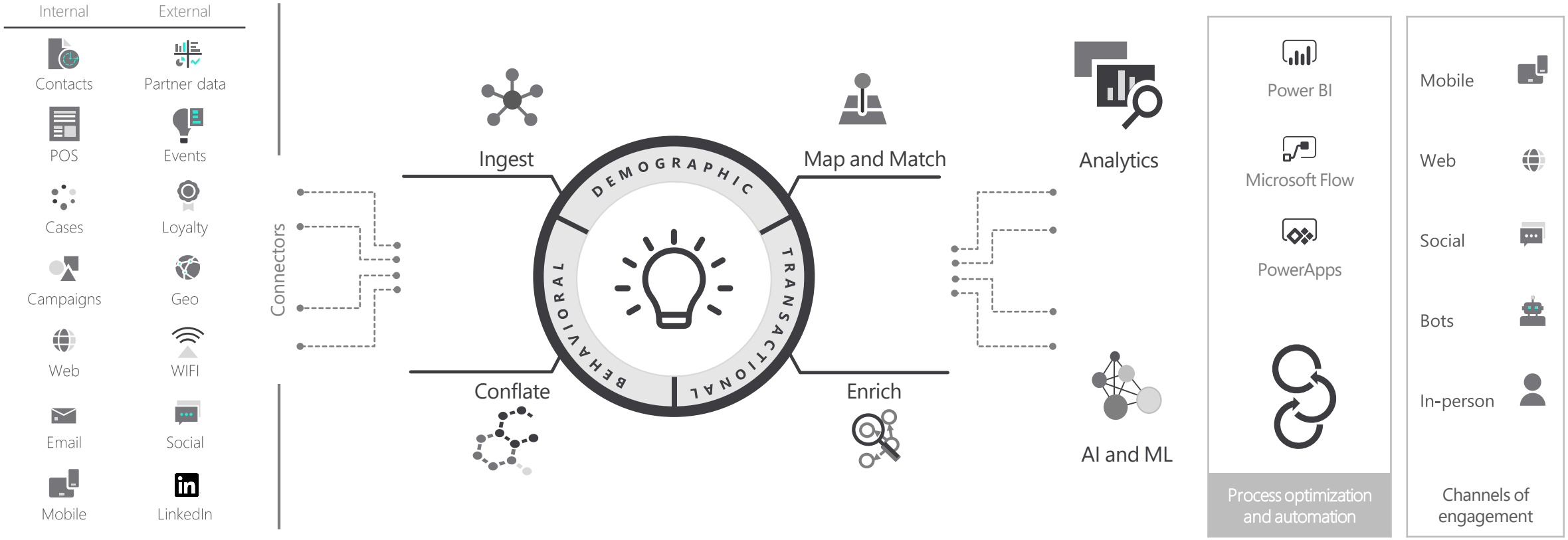
Garner Research. 09 December 2016. "Citizen Data Science Augments Data Discovery and Simplifies Data Science"

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Customer Insights

Overview



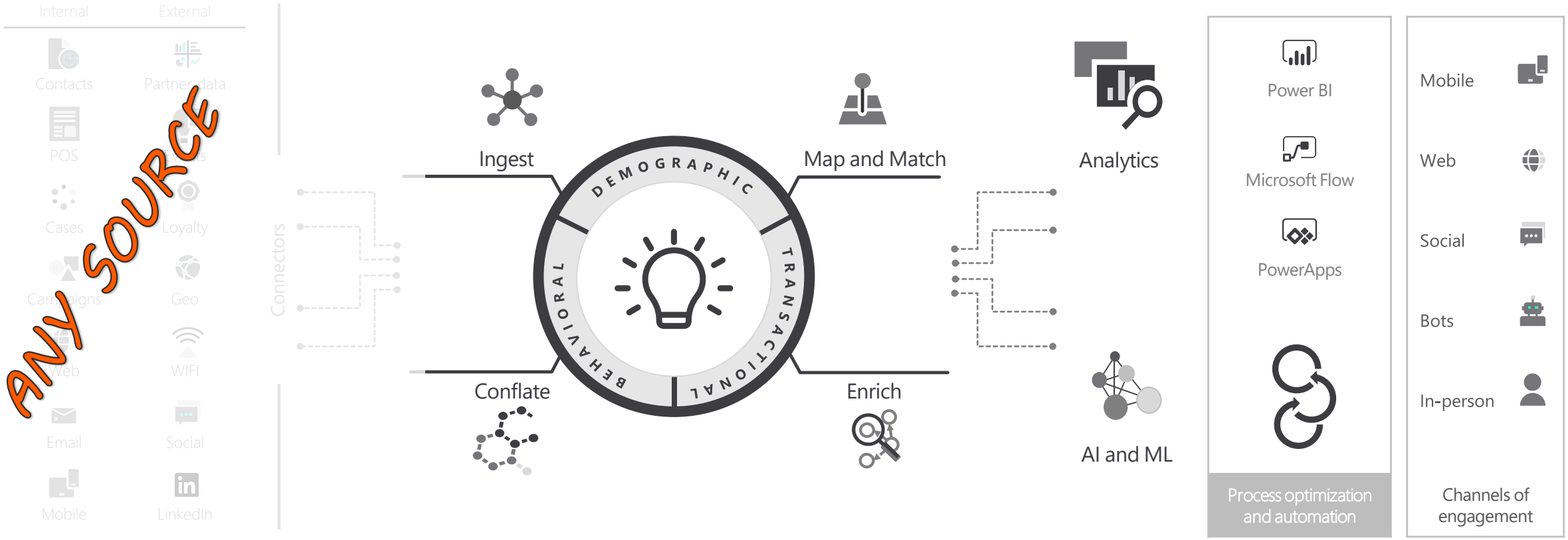
Data — Unification — Insights — Action

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Customer Insights

Overview



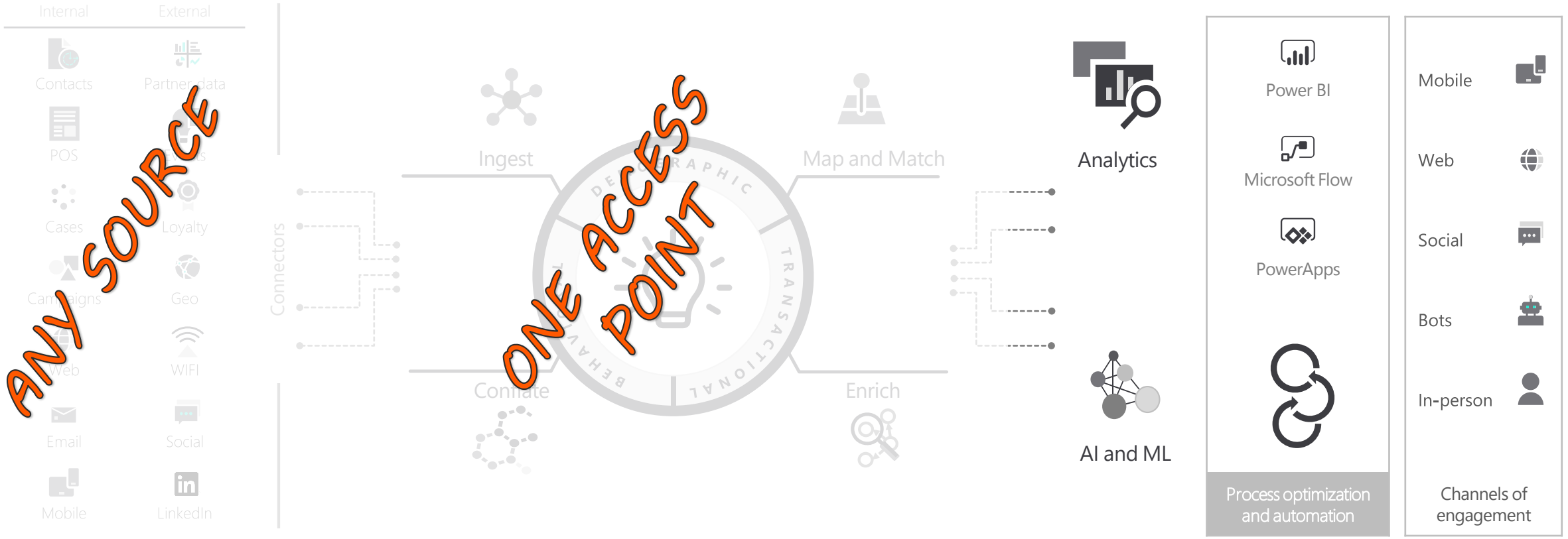
Data — Unification — Insights — Action

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Customer Insights

Overview



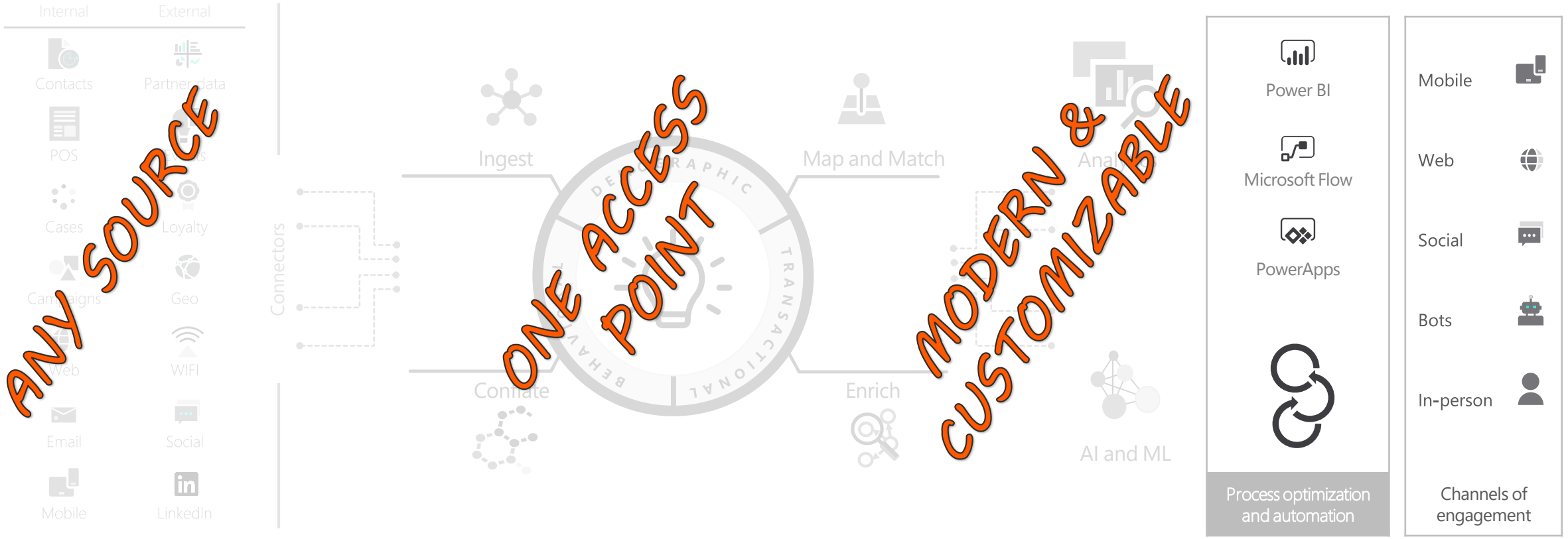
Data — Unification — Insights — Action

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Customer Insights

Overview



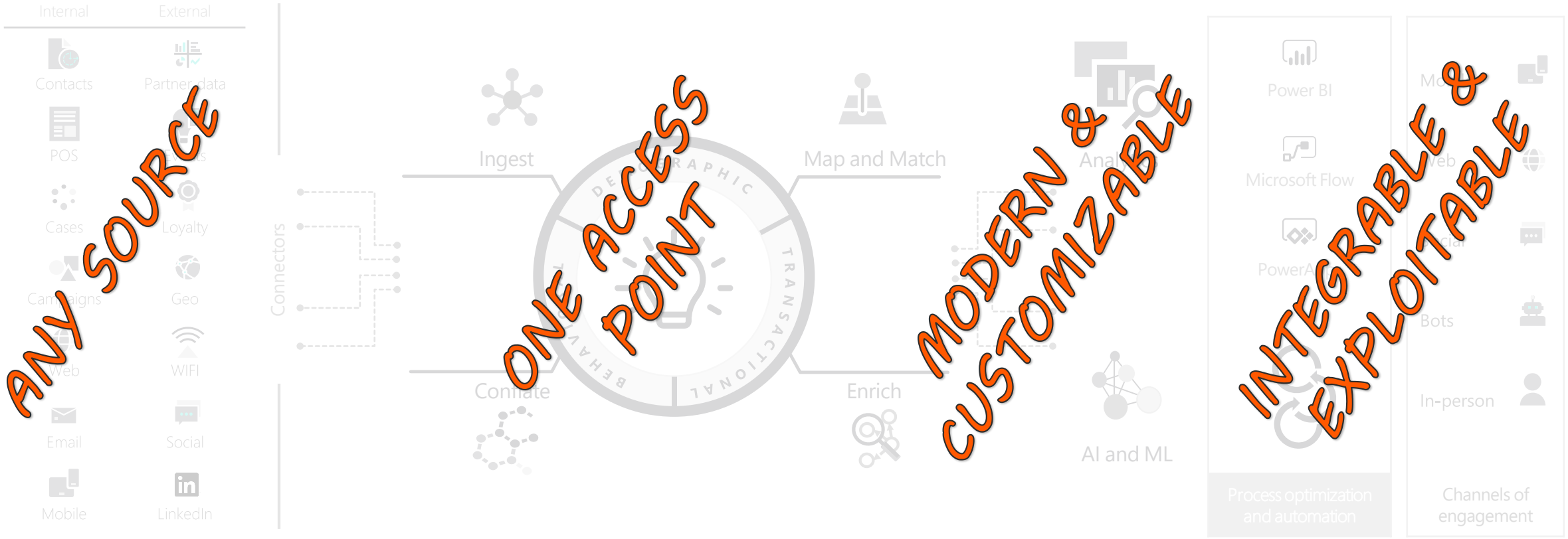
Data — Unification — Insights — Action

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Customer Insights

Overview



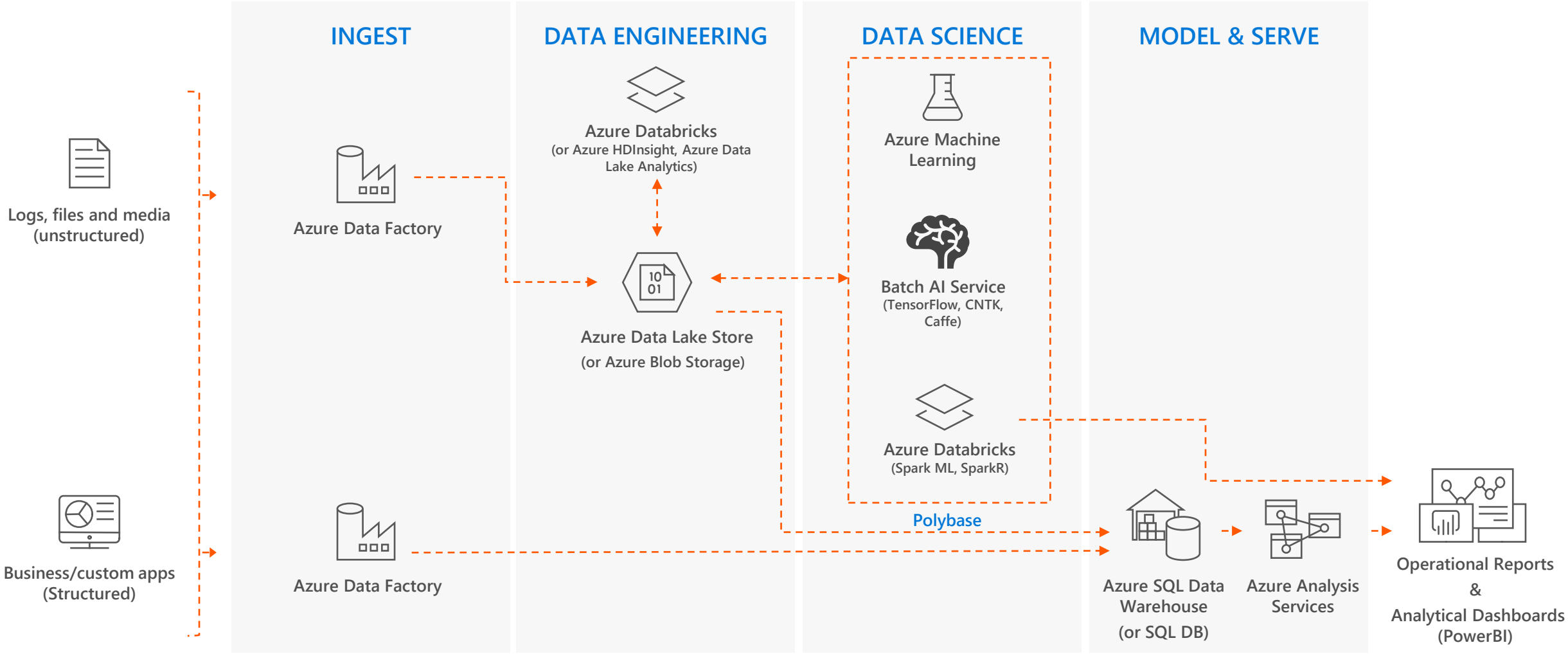
Data ---------- Unification ---------- Insights ---------- Action

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

The Advanced Analytics & AI Platform

for data of all shapes and sizes



avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

avanade

Microsoft Dynamics 365
Customer Insights

Microsoft

Empower every organization to unify and understand its customer data & usage of their products/services to derive insights that power personalized experiences and processes.

Let's see …

A demo of Customer Insights

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

AutoML

Automated Machine Learning

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

*"Train **high-quality** custom machine learning **models** with **minimal effort and machine learning expertise**."*

Google. Cloud AutoML
https://cloud.google.com/automl

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

What is AutoML?

Automated machine learning, also referred to as automated ML, is the process of **automating the time consuming, iterative tasks of machine learning model development**. It allows data scientists, analysts, and developers to build ML models with high scale, efficiency, and productivity all while sustaining model quality. Automated ML is based on a breakthrough from our Microsoft Research division.

Google Cloud

H2O Driverless AI
Award-winning Automatic Machine Learning Platform

DataRobot

TPOT

data iku

Microsoft

Automated ML
Automatically train and tune a model using a target metric.

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

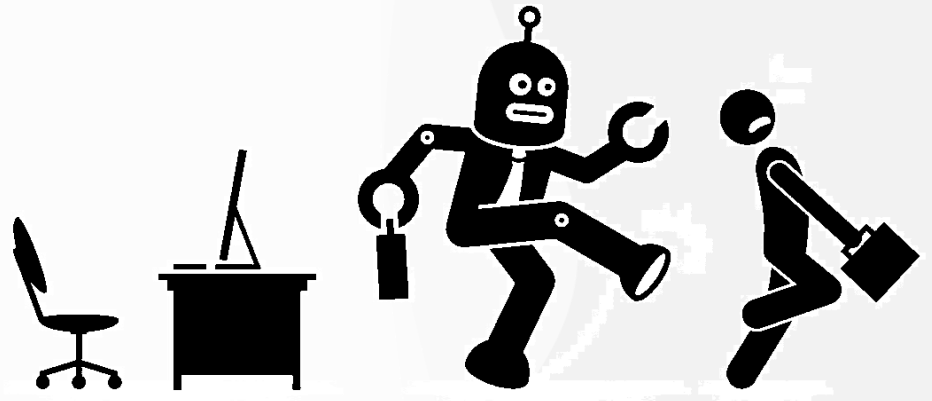
Why AutoML?

"... machine learning is only fully understood by a small group of advanced engineers and statisticians called data scientists. Due to the fact that they are some of the most sought after professionals at the moment, only the biggest companies and organizations can afford to attract them.

But Auto ML is changing all that. **It makes it possible for organizations and companies of all sizes to access the benefits of machine learning without having to establish their own resident team of experts.**"

RapidMiner

Automated ML **democratizes** the machine learning model development process, and **empowers** its users, **no matter their data science expertise**, to identify an end-to-end machine learning pipeline **for any problem.**



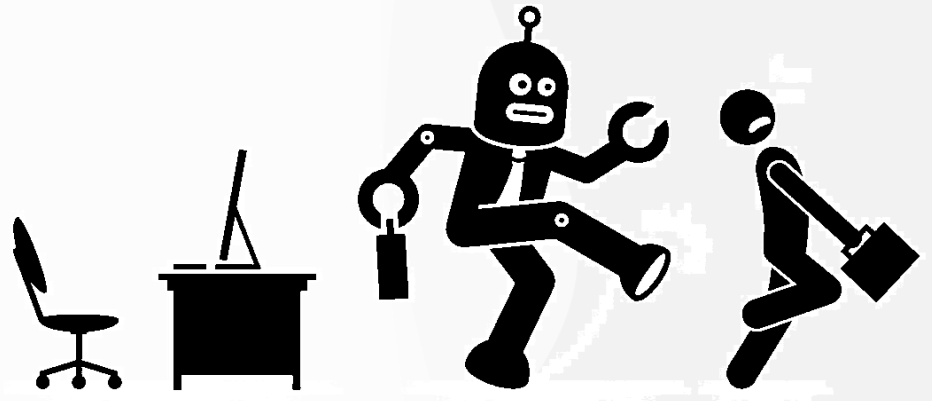
avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Why AutoML?

- AI democratization
- Costs Comparison
- Fill the gap between "supply" and "demand" in the Data Science market
- Save time of Data Science team
- Outperforms average Data Scientist: increase model quality

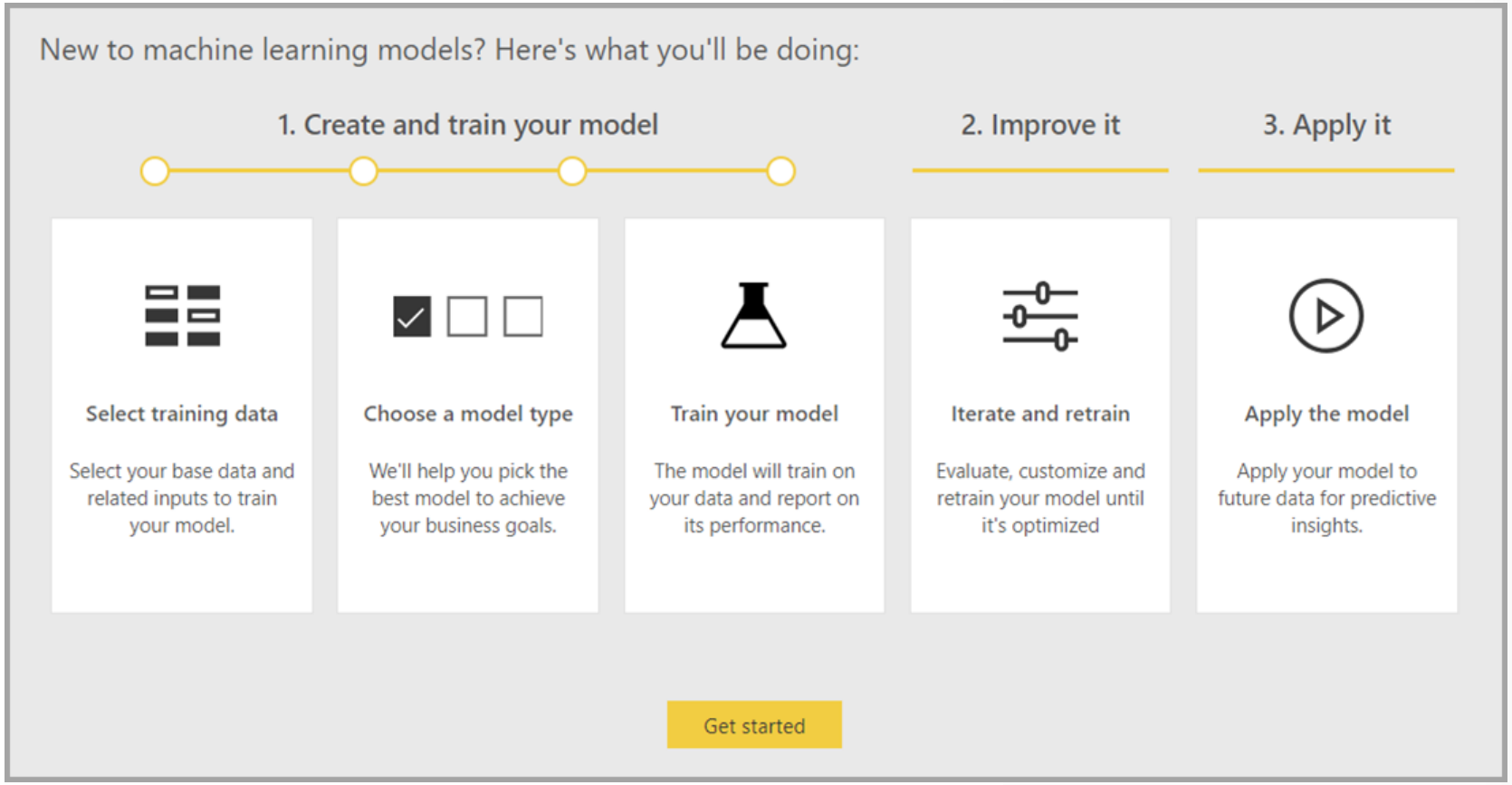
Automated ML **democratizes** the machine learning model development process, and **empowers** its users, no matter their data science expertise, to identify an end-to-end machine learning pipeline **for any problem.**



avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Automated Machine Learning in Power BI



AutoML in Power BI enables data analysts to **use dataflows to build machine learning models with a simplified experience**, using just Power BI skills. Most of the data science behind the creation of the ML models is automated by Power BI. It has guardrails to ensure that the model produced has good quality and provides visibility into the process used to create your ML model.

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Our Mission

Solve *Real* Problems using Data & *Ideal* Solutions

avanade

©2019 Avanade Inc. All Rights Reserved. <Highly Confidential>

Our Mission

Solve *Real* Problems using *Real* Data *Science*

Technology

avanade

©2019 Avanade Inc. All Rights Reserved. <Confidential>

Leon Eyrich Jessen
@jessenleon

Domain specific knowledge is crucial in #dataScience, this is my new favourite picture to illustrate why! You may be an absolute #rstats or #python ninja, but be sure to team up with a domain specialist if you find yourself in unknown territory 💻📊📈📉👍

Cutting tennis balls in half lets you store 2 more balls, saving space

8:42 AM · Feb 19, 2020 · Twitter Web App

g tennis
in half
ou store
re balls,
g space

We change things
for the better