

Using big data, the cloud, and AI to enable intelligence at scale

Danielle Dean, Wee Hyong Tok
Principal Data Scientist Lead
Cloud AI
Microsoft

@danielleodean | @weehyong



Overview

Big Data, Cloud and AI Use Cases

What technology/service should I use to get started?

Going deeper – custom solutions

Practical tips for putting it together

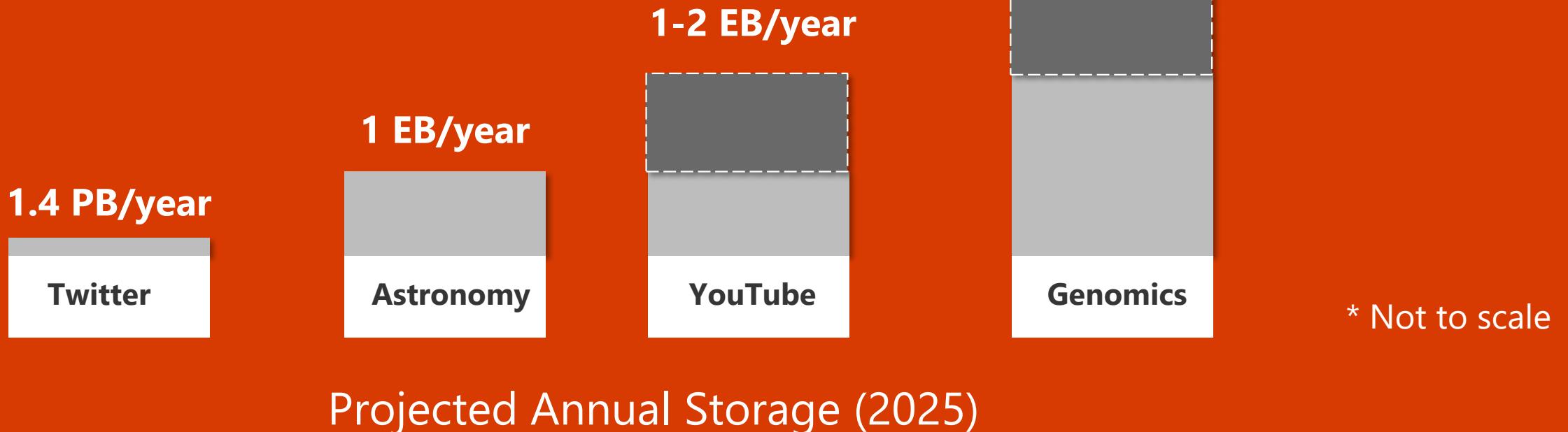


Why does AI matter to the
this Community?

"AI is going to disrupt every single business app – whether an industry vertical like banking, retail and health care, or a horizontal business process like sales, marketing and customer support"

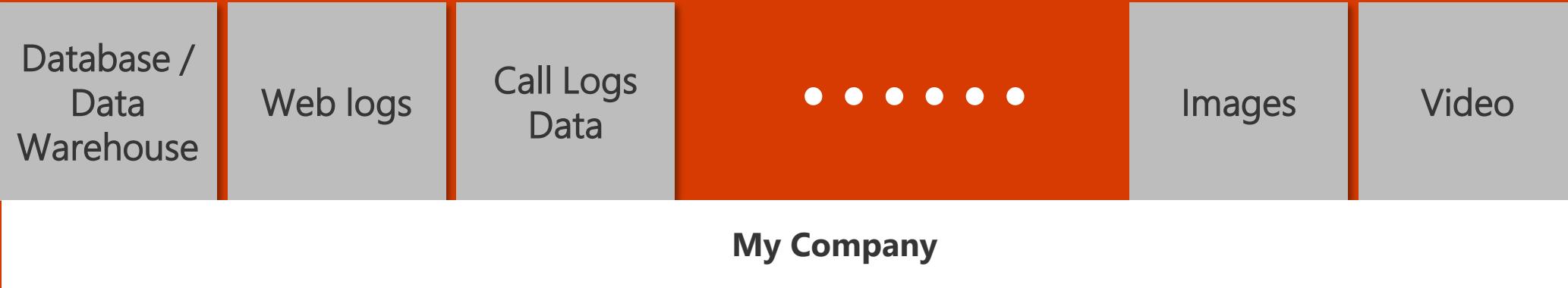
*- Harry Shum, Microsoft Executive VP,
AI and Research*

Major Generators of Big Data



Src: "Big Data: Astronomical or Genomical?", PLOS | Biology, July 2015

Organization's Data



Data

Intelligent
Cloud

Deep Learning

Big Data, Cloud and AI

How is it being used?

Big Data, Cloud and AI - Use Cases



Lowe's - Customers Style/Preference Matching For Kitchen-Remodeling



ESmart – Connected Drone and Power line inspection



Jabil – PCB Boards Defect Inspection

Jabil

Paul Bunting
Architect

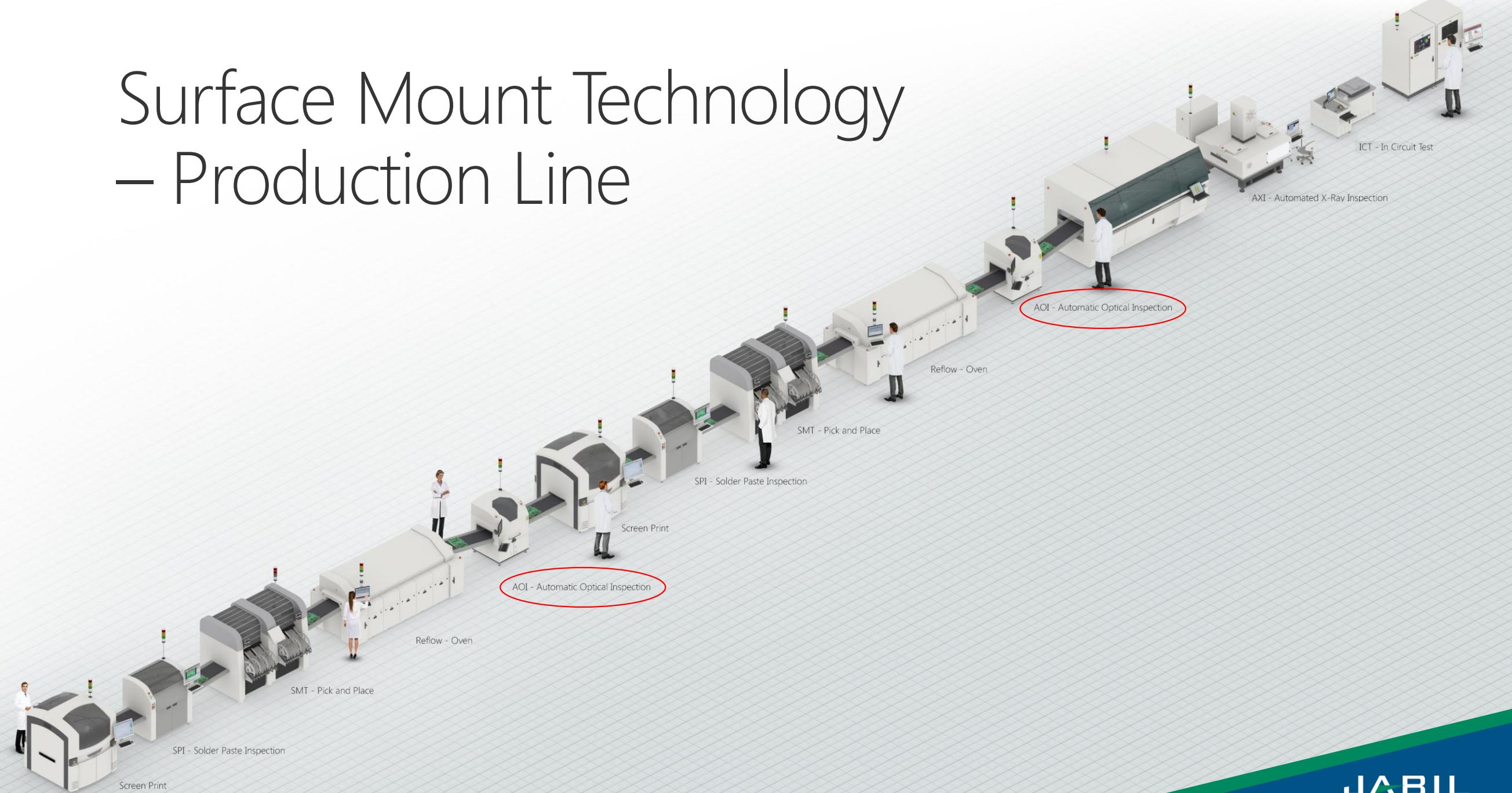
Jabil AOI Deep Learning Image Analysis and Predictive Hub



Paul Bunting – Architect

Jabil's advanced predictive hub pipeline and user interface overview.

Surface Mount Technology – Production Line

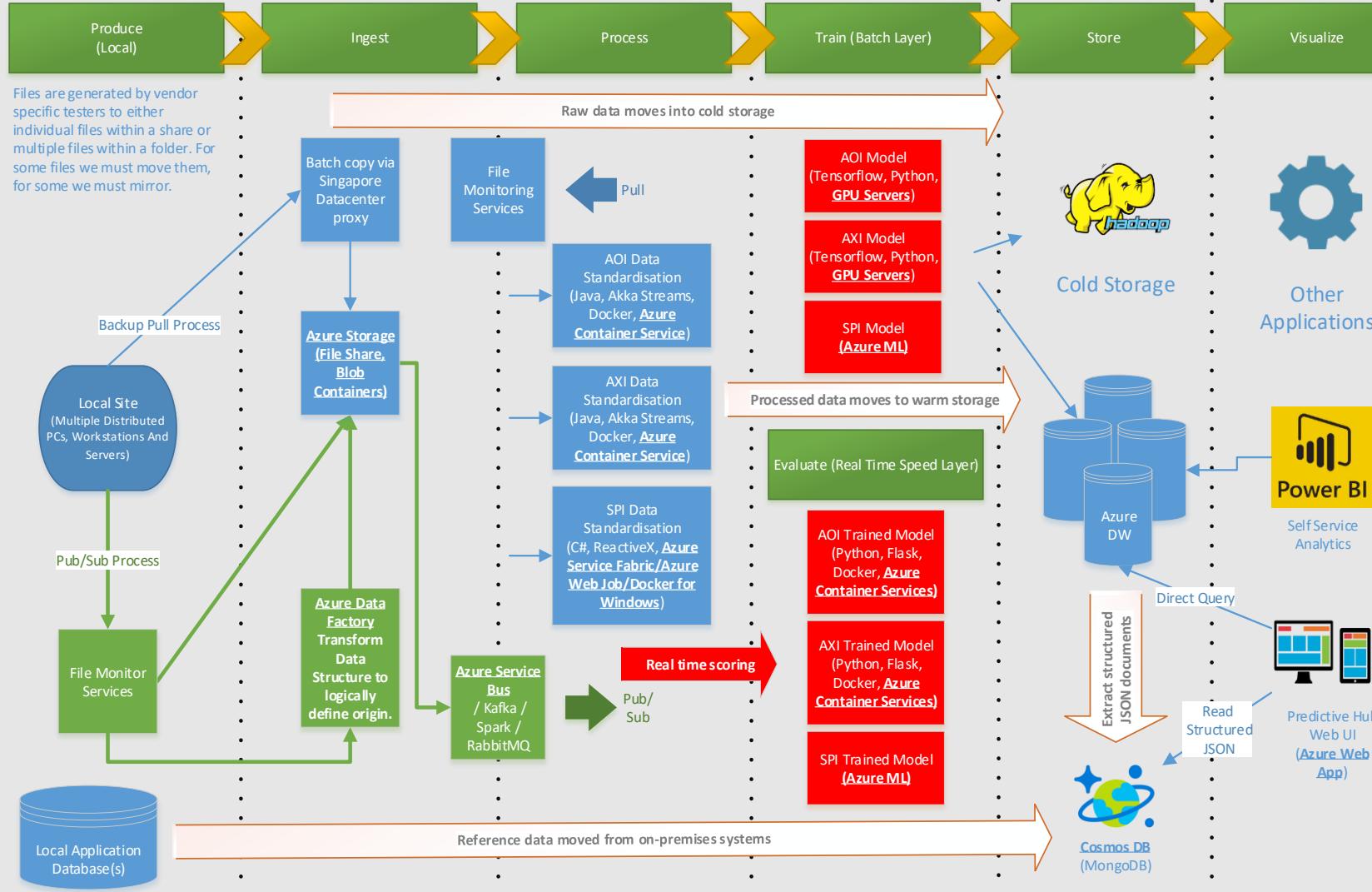


Traditional AOI Manual Inspection Method

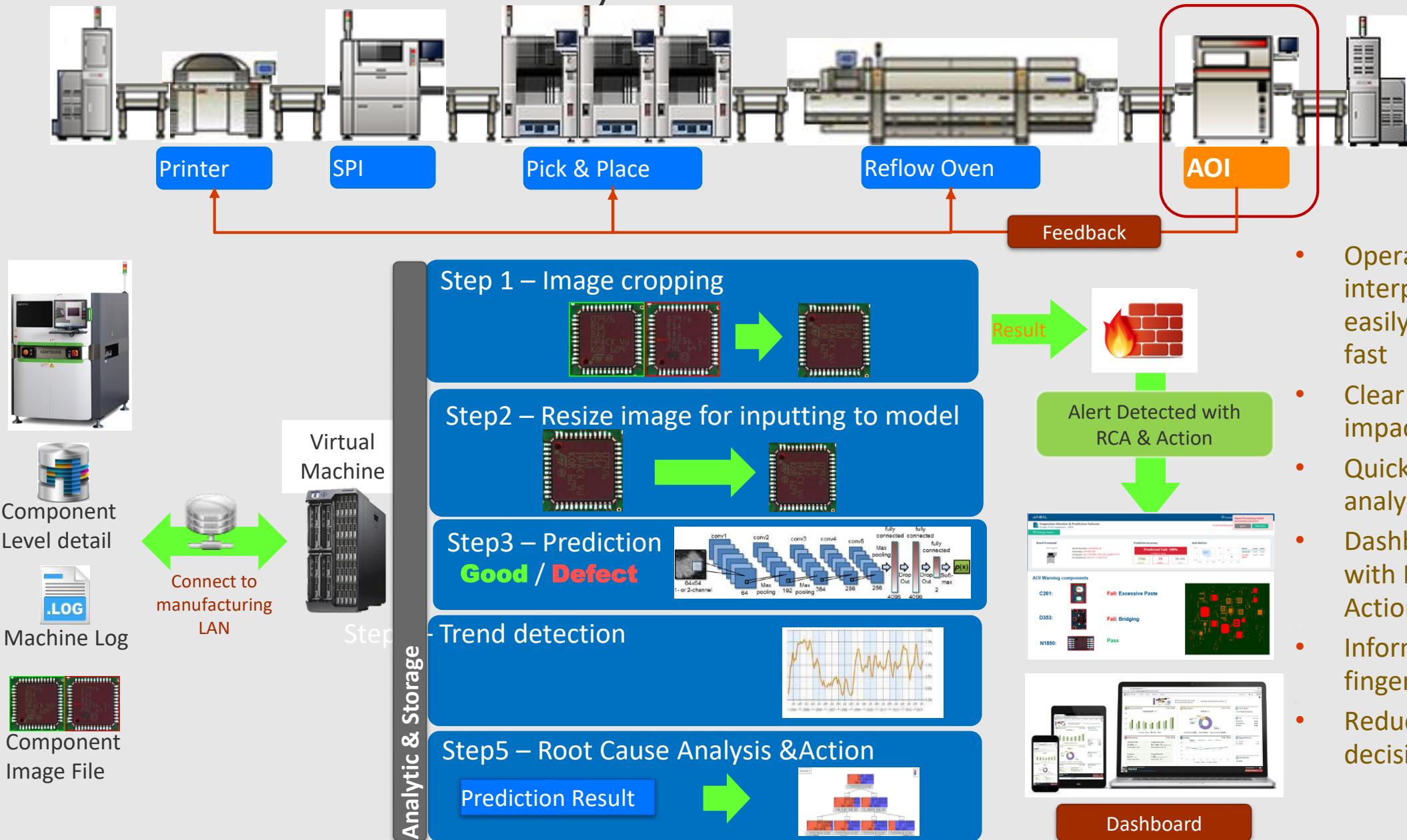
- Once AOI (Automatic Optical Inspection) Flags any number of components as potential failures the operator must cycle through each of the images individually to determine if a true failure has been encountered or if the component placement and solder is OK.
- Due to the number of components in existence on each assembly (2000+), which must be inspected by AOI, this will result in a near 100% manual inspection process after AOI has analysed the board and classified suspected failures.



New Predictive Hub Data Flow and Deep Learning Pipeline



AOI Predictive Analysis



- Operator / MT can interpret the graph easily and take action fast
- Clear visualization on impacted tool
- Quick commonality analysis
- Dashboard updated with Root Cause & Action
- Information on the finger tips
- Reducing human decision

New Jabil AOI User Interface

Predictive Hub - AOI Operator

localhost:39296/Home/Aoi/Operator

JABIL
Enterprise Manufacturing Intelligence

Solder Paste Inspection (SPI)

Automated Optical Inspection (AOI)

Advanced 3D X-ray Inspection (AXI)/5DX

AOI Operator

BOARDS

Manufacturing Area: Bay 03-2
Processed From: 2017-09-01 09:54:48

Predictive
2 pass
1 fail
66.67% yield

Operator
0 pass
0 fail
3 to be checked
NA yield

FAIL
1 of 2092 total components

COMPONENT FAIL

Manufacturing Area: Bay 03-2
Serial Number: CD3T147756

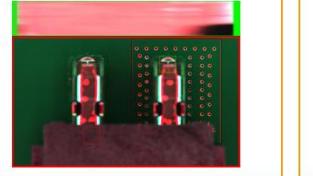
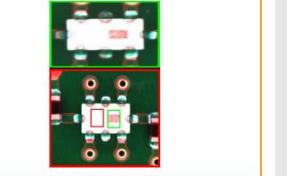
TOP 5 FAILURE

Manufacturing Area: Bay 03-2 CD3T147756 has 1/2092 failure
Processed From: 2017-09-01 component(s) at null, Bay 03-2

Top	Inspection Category	Total
1	Ocvfail	77
2	Liftedlead	39
3	Bad Joint	36
4	Billboard	27
5	Horoff	15

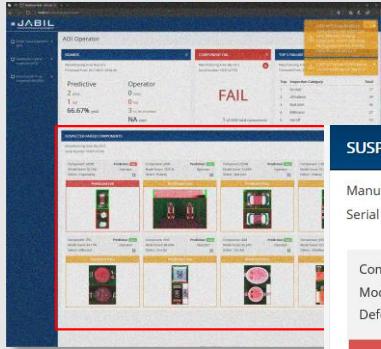
SUSPECTED FAILED COMPONENTS

Manufacturing Area: Bay 03-2
Serial Number: CD3T147756

Component: c6545 Model Score: 92.10% Defect : Coplanarity Prediction: FAIL 	Component: p800 Model Score: 70.01% Defect : Polarity Prediction: PASS 	Component: l5508 Model Score: 72.54% Defect : Bad Joint Prediction: PASS 	Component: z1006 Model Score: 75.15% Defect : Polarity Prediction: PASS 
Component: l701 Model Score: 84.17% Defect : Billboard Predicted Pass 	Component: r943 Model Score: 86.65% Defect : Ocv fail Predicted Pass 	Component: c828 Model Score: 92.24% Defect : Ocv fail Predicted Pass 	Component: j1003 Model Score: 92.81% Defect : Liftedlead Predicted Pass 
Component: 103 Model Score: 99.99% Defect : Coplanarity Predicted Pass 	Component: 220 Model Score: 99.99% Defect : Coplanarity Predicted Pass 	Component: 220 Model Score: 99.99% Defect : Coplanarity Predicted Pass 	Component: 220 Model Score: 99.99% Defect : Coplanarity Predicted Pass 

New Jabil AOI User Interface

Image deep learning predictive analysis.



Predicted and true failures are identified

True passes are identified

Predicted passes are highlighted in orange to allow operator verification.

SUSPECTED FAILED COMPONENTS

Manufacturing Area: Bay 03-2
Serial Number: CD3T093812

Component	Prediction	Model Score	Operator	Defect
d3001a9	FAIL	99.72%	-	Ocv fail
c504p05...	FAIL	89.03%	-	Coplanarity
z1l3a6	PASS	85.61%	-	Bad Joint
r179p01...	PASS	97.35%	-	Billboard
r9a8	PASS	99.97%	-	Billboard

Predictive and Operator result is given for each image

AOI suspected fail reason is given

Image shows expected (green) and actual (red) images

Operator may zoom into image

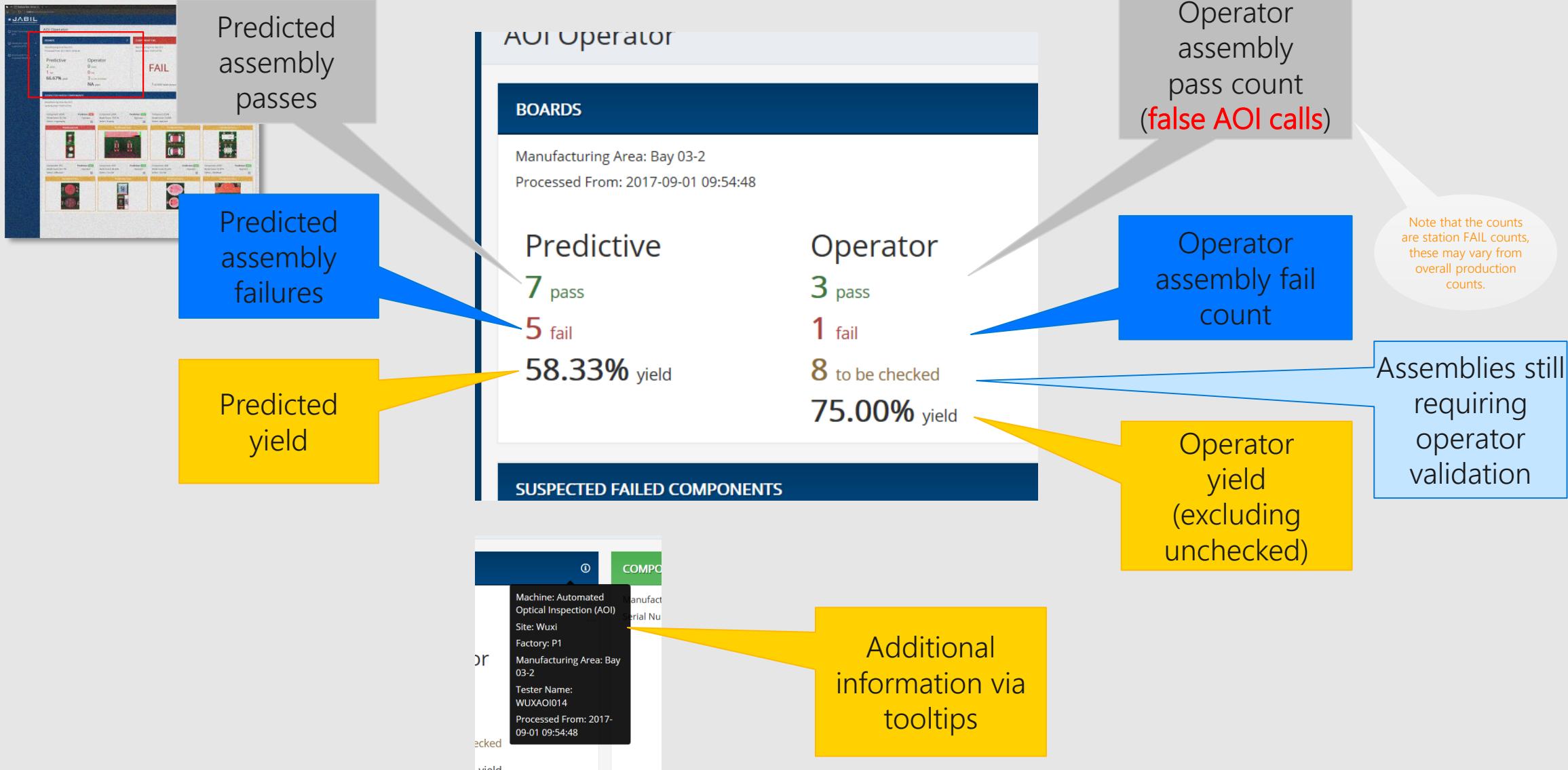
SUSPECTED FAILED COMPONENTS

Manufacturing Area: Bay 03-2
Serial Number: CD3T160613

Component	Prediction	Model Score	Operator	Defect
n600a2az	PASS	98.22%	PASS	Liftedlead
r100a2az	PASS	99.87%	PASS	Coplanarity
d8002a9	PASS	99.89%	PASS	Ocv fail
r15a2az	PASS	99.93%	PASS	Coplanarity

New Jabil AOI User Interface

Feedback on predictive performance verses operator performance.



New Jabil AOI User Interface

Current assembly component fail prediction.



Predicted Fail

Colour and layout
coded
for clarity

Predicted Pass

COMPONENT FAIL

Manufacturing Area: Bay 03-2
Serial Number: CD3T147756

Predicted
result

FAIL

1 of 2092 total components

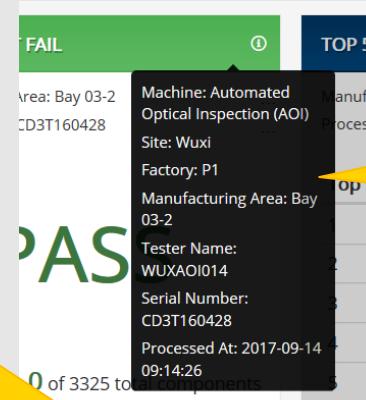
Number of components
predicted to fail verses the
total number of components
tested for that assembly

COMPONENT FAIL

Manufacturing Area: Bay 03-2
Serial Number: CD3P951004

PASS

0 of 3456 total components

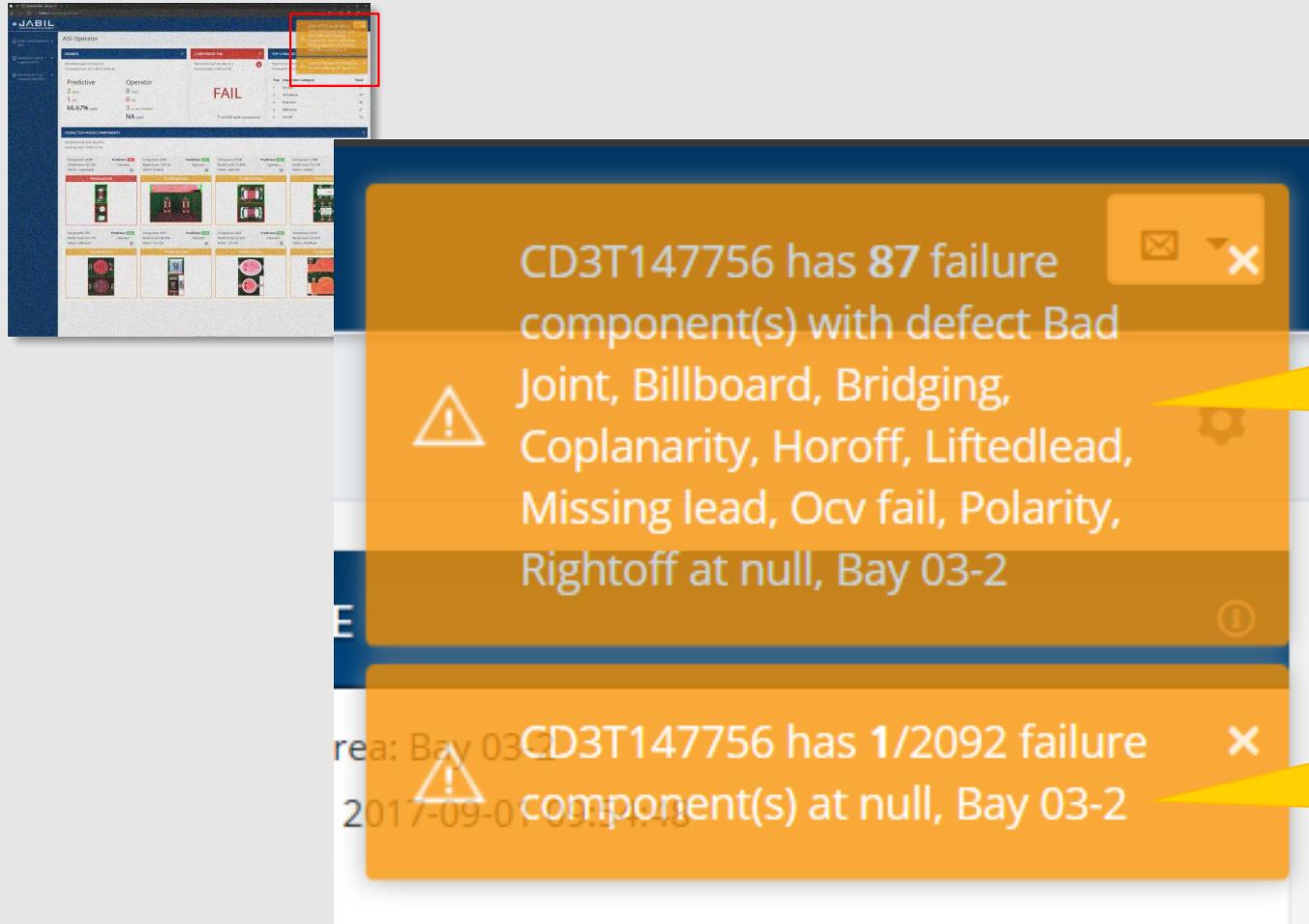


Additional
information via
tooltips

Serial number of
assembly &
manufacturing area
information

New Jabil AOI User Interface

Toaster messages offer additional real time feedback to operator.

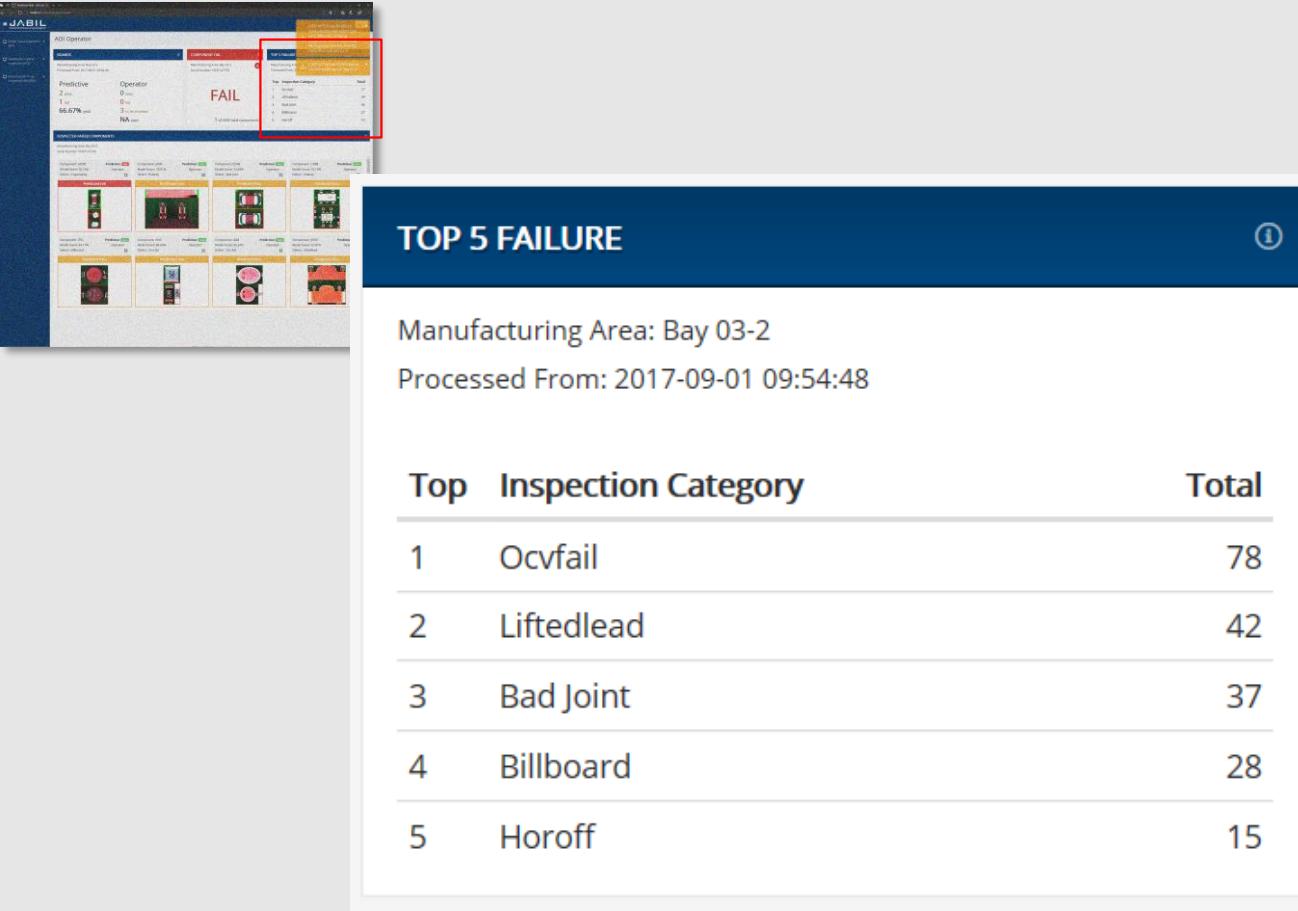


Information is given on AOI results, the number of suspected failed components and the types of failure

Information is given on the predictive result of the AOI failure images

New Jabil AOI User Interface

Total (worst 5) AOI failure types identified, and actionable.



The screenshot shows a dashboard titled "AOI Operation" with various performance metrics and a "FAIL" status indicator. Below it, a "TOP 5 FAILURE" section displays the following data:

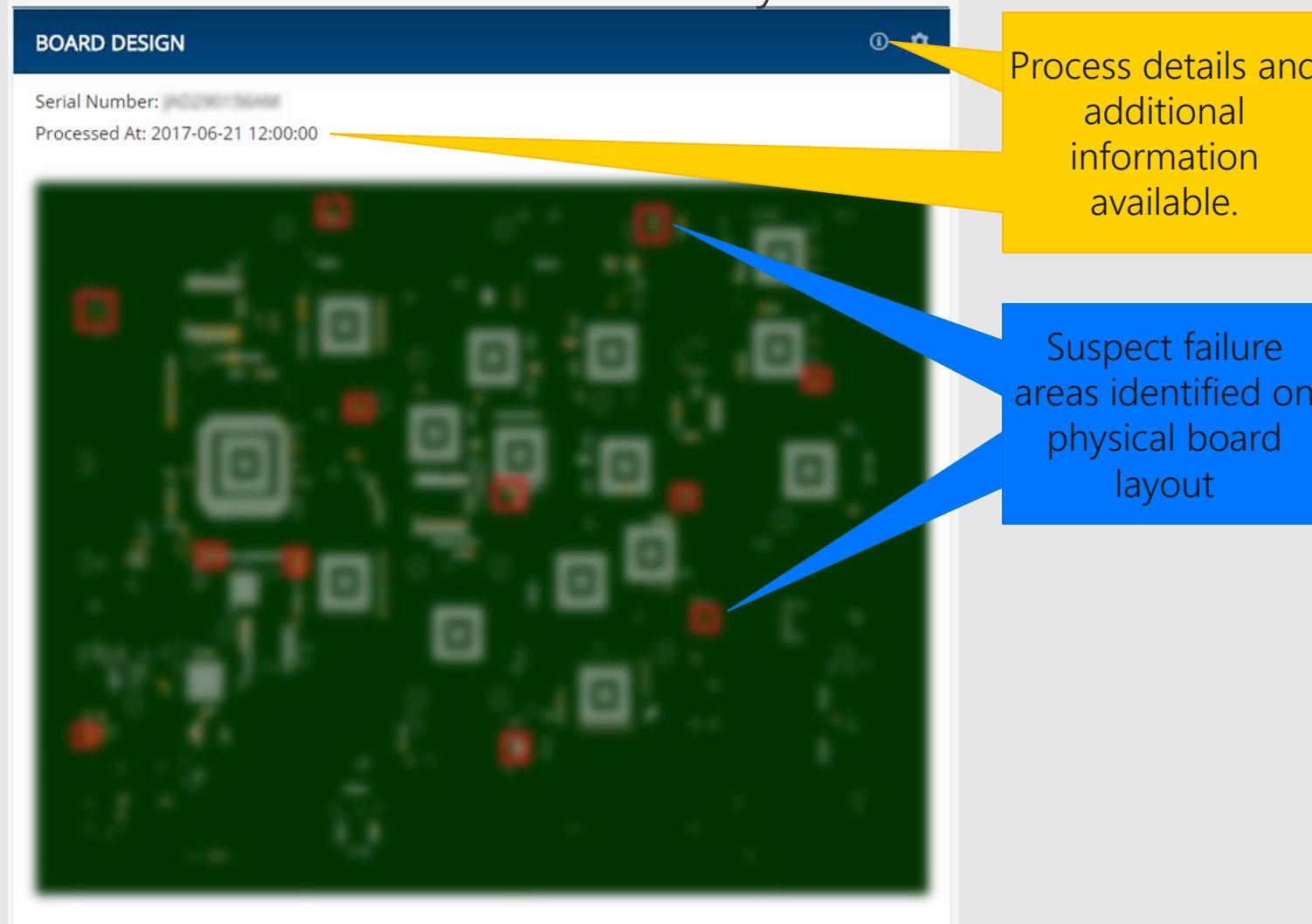
Top	Inspection Category	Total
1	Ocvfail	78
2	Liftedlead	42
3	Bad Joint	37
4	Billboard	28
5	Horoff	15

Information is given on the types of suspected failures being indicated at AOI, so action may be taken against specific root causes

Board Design

(not included in demo due to customer confidentiality – blurred here)

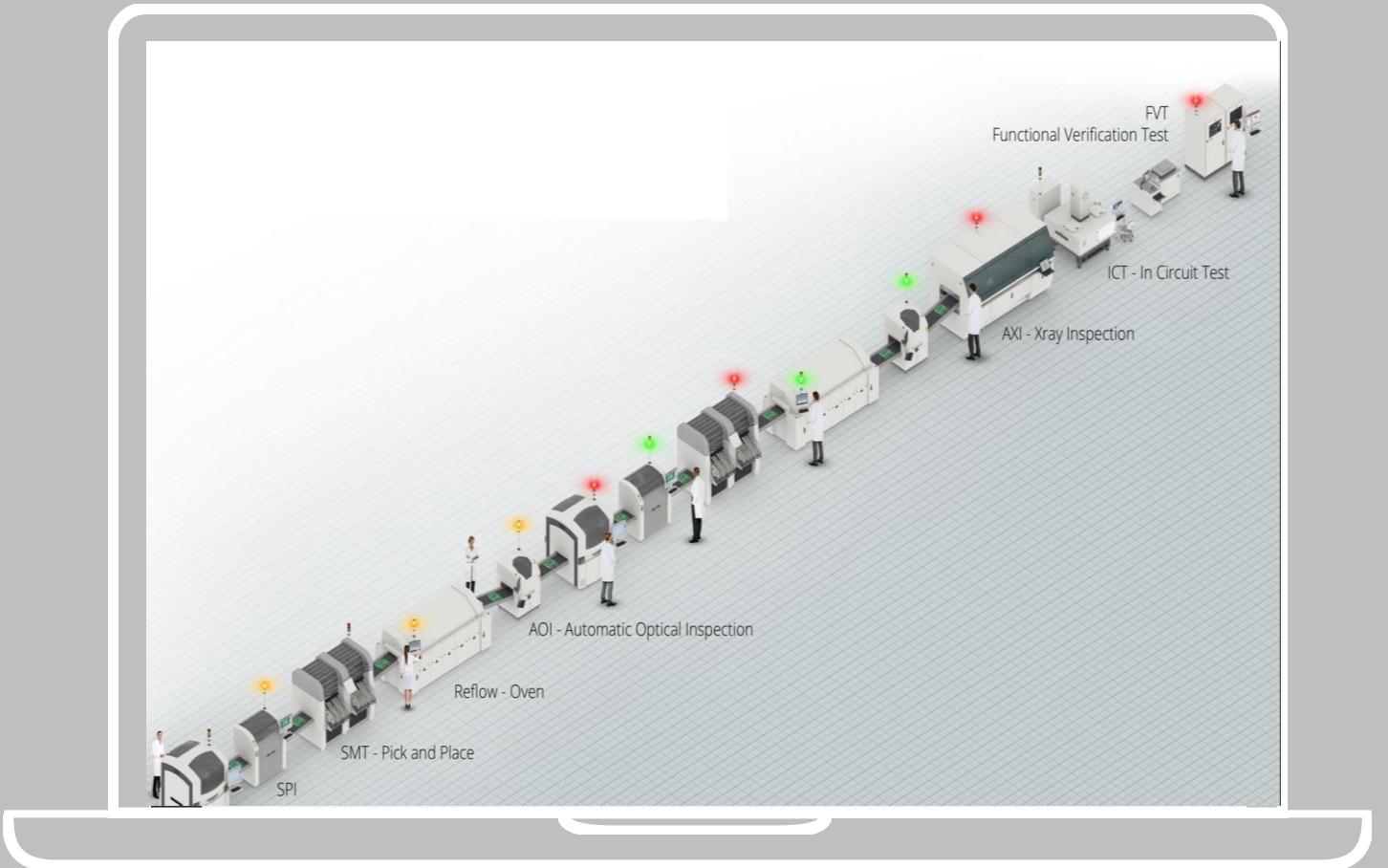
Board design and layout offers operator instant feedback on where to look for identified faults on the assembly.



Demo

Jabil Predictive Hub

Paul Bunting



ESRI

Omar Maher
Practice Lead, Advanced Analytics



GeoAI: ArcGIS + Microsoft AI

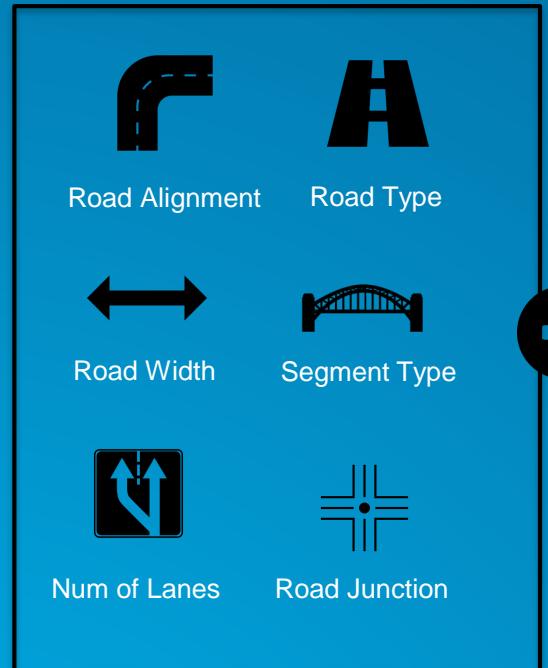
The Power of Geospatial Machine Learning



Incident Management

Predict Incident Locations, Classify by Severity, Optimize Asset Allocation

Road Data



Accidents' Data



Metadata (Optional)



Output



Predict the Location, Type and Time of Accidents before they occur

- Cut Incident Rates
- Explore Accidents Root-Causes
- Optimize Patrols Allocation
- Improve Response Time



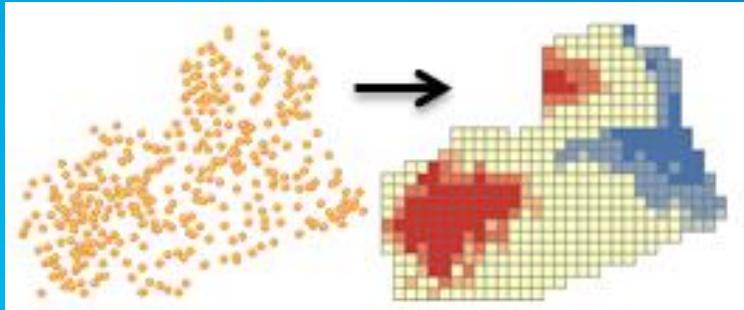
Machine Learning Models over 1+ year of Accidents Data

1. Spatial Data Exploration

Add the Road Network Layer to derive the Road Geo-Features (curvature, speed..)



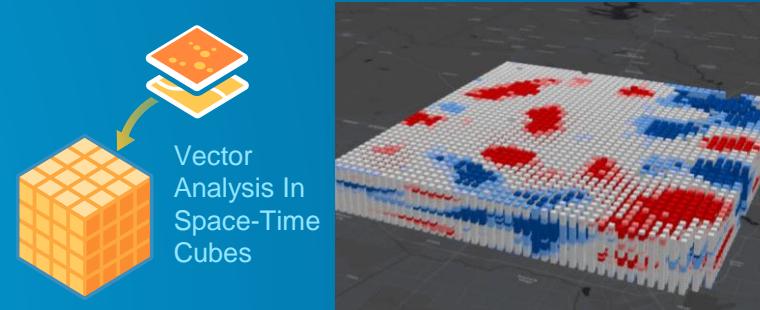
Random VS Non-Random Events?
Run Hotspot Analysis



Data not Accurate: Crashes location slightly way from Road Links



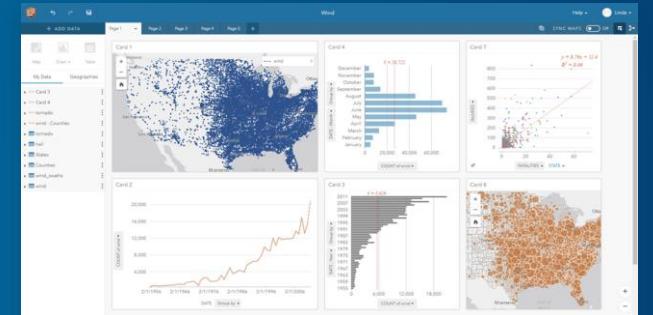
Emerging VS Fading Hotspots?
SpaceTime Cube



Patterns: Crashes mostly on Crossroads.
Add Proximity to Crossroads as a Feature

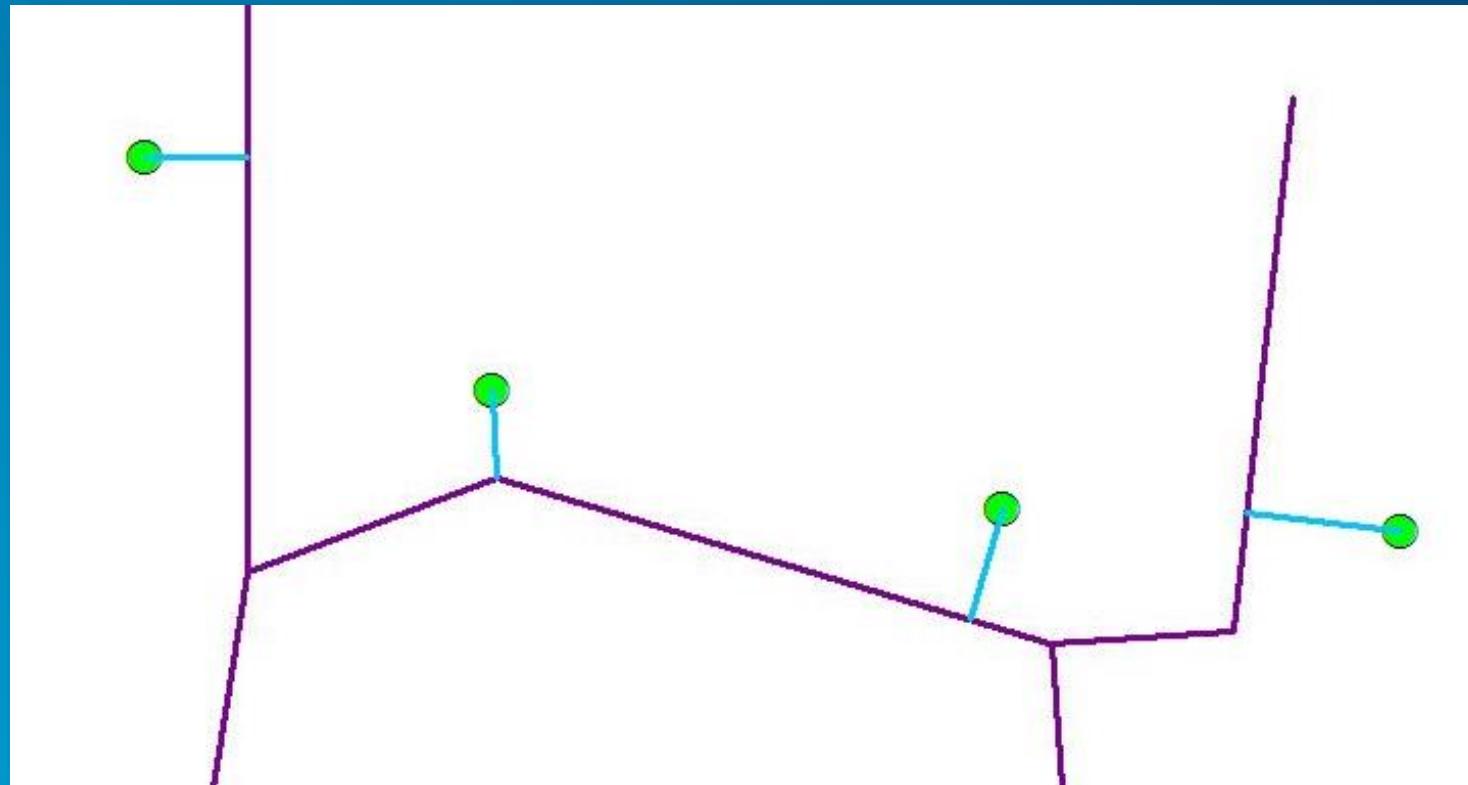


Explore Spatiotemporal Patterns..



2. Spatial Data Preparation

Example: Snapping Crash Points to the Nearest Line (Road Link), GeoEnrichment w/ 10K Features



3. Spatial Feature Extraction



Road Alignment
Straight / Curved



Road Type
*Double / Single
Carriageway*



Num of Lanes
1,2,3..



Road Width
20-30 M



Segment Type
*Highway, Bridge,
Tunnel..*



Road Junction
*Crossroads,
Roundabout..*



Day of the Week
Sun, Mon, Fri..



Time of the Day
12:45, 23:00



Rush Hour
Yes/No



Weekend
Yes/No



Raining
Yes/No



Fog
Yes/No



Temperature
Sun, Mon, Fri..



Nearby Event
Sports, Festivals..



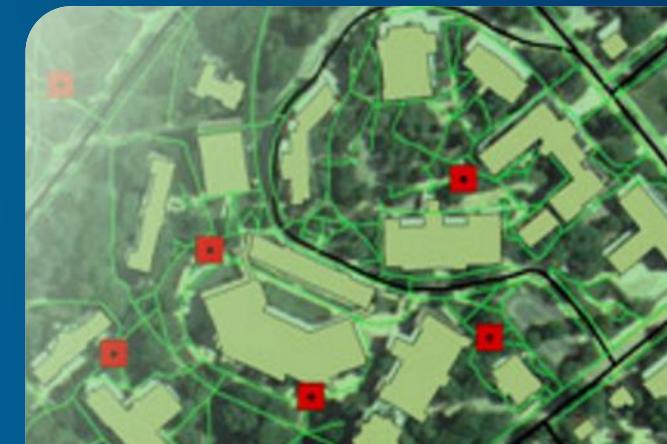
Nearby block
*From related
roads*



Speed Limit
120 km/h



Road Works
Yes/No



4. Feature Selection: listing the values of all independent features variables that were associated with historical accident incidents – and running Feature Selection via Azure Machine Learning

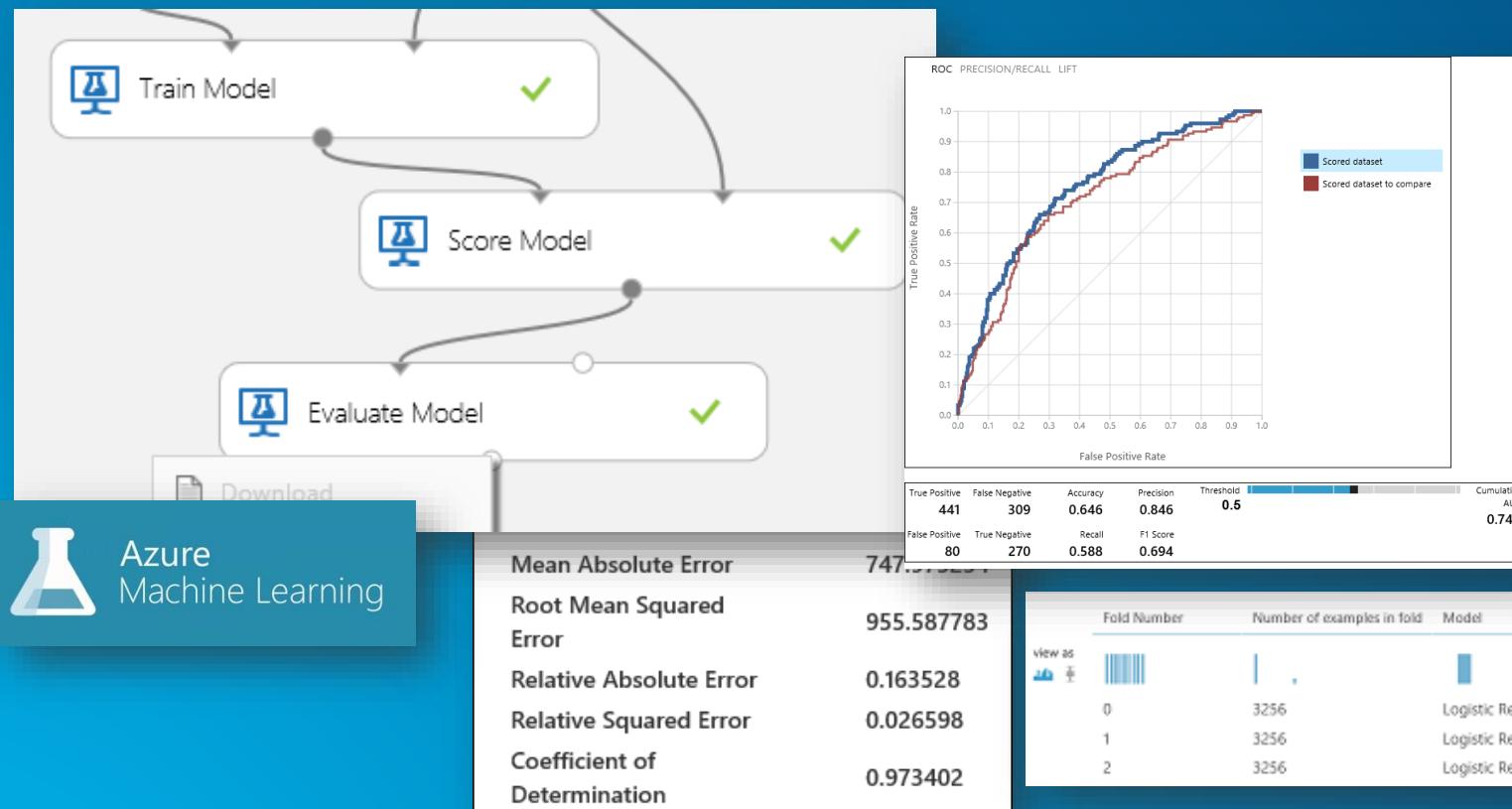
Independent Features

Time	Segment	Curved	Lanes	Type	Width	Rain	Wind	Temp	Prox Inters	Weekend	...	Accident
17:05	A123	Yes	2	Highway	30	Yes	65 km/h	28	0.3	No	...	Two Cars
23:50	B742	No	3	Tunnel	45	No	23 km/h	18	0.1	Yes	...	Road Deviation

Dependent Feature

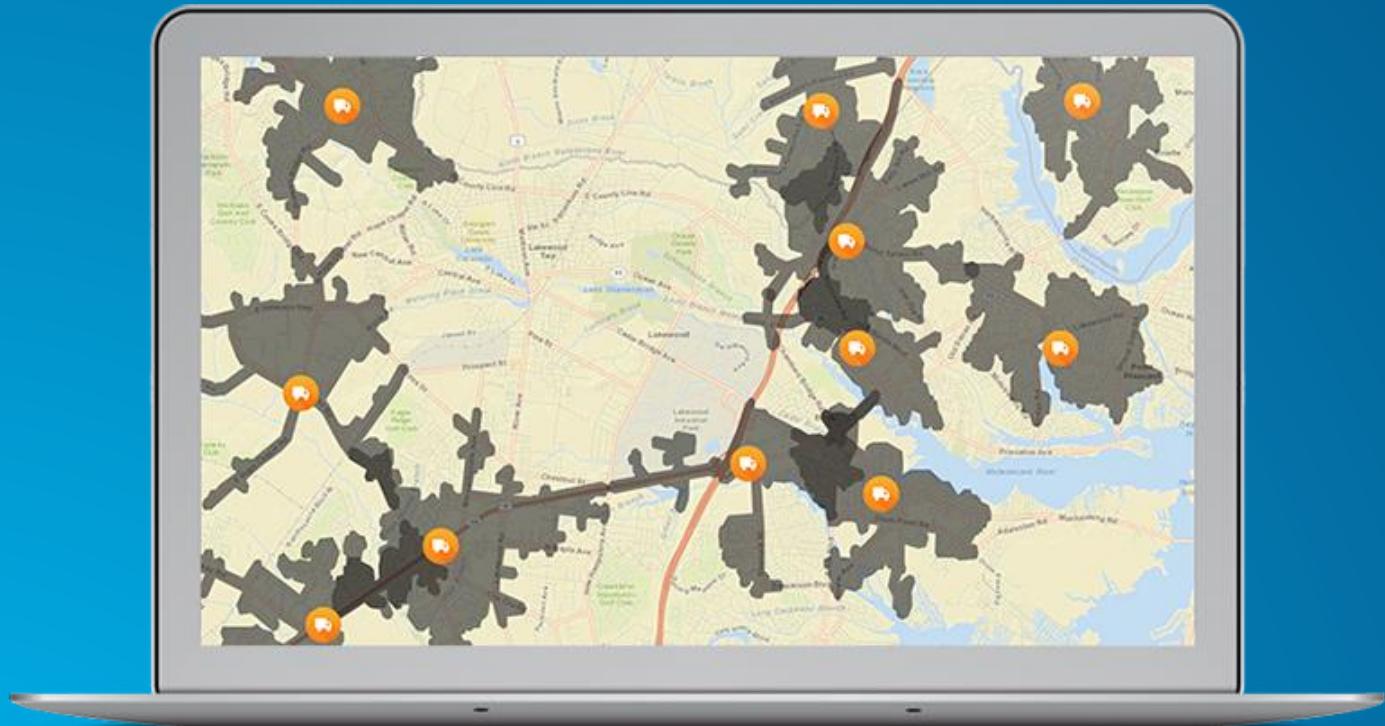
view as	Feature	Score
	Prox to Intersections	3074.528083
	Weather - Rain	3037.526148
	Road Curvature	3003.680839

5. Model Development: Experimenting with different models to see which yield highest Matric of interest (e.g. RMSE) via Azure Machine Learning



Time	Segment	Curved	Lanes	Type	Width	Rain	Wind	Temp	Prox Inters	Weekend	Accident Probability
17:05	A123	Yes	2	Highway	30	Yes	65 km/h	28	No	No	82%
23:50	B742	No	3	Tunnel	45	No	23 km/h	18	No	Yes	25%

6. Post Prediction Actions: Optimizing Asset Allocation: based on Accident Predictions



- Optimize the allocation of Police Patrols & Ambulance Cars based on the Predicted Accident Locations
- Take Accident type into consideration (Critical, Medium..)
- Inputs: Predictions, Asset types, numbers, allocation constraints, shifts.. Etc
- Root-cause Analysis

GeoAI Project Lifecycle

Spatial Data Exploration



How is the Data distributed Spatially?
Any Spatial Patterns of interest?

Spatial Data Preparation



Example:
Snapping Car Crashes to Road Links, Geo-enrichment

Spatial Feature Extraction



Example:
Road Curvature, Number of Lanes, Proximity to Crossroads

Feature Engineering



Exploring Input Feature Correlation with the Output Feature. Feature Selection Techniques

Model Development



Iterating with different Models. Choosing Models per Metric of choice (e.g. RMSE)

Spatial Action Facilitation



Facilitating Post-Prediction Actions
E.g.: Optimizing Ambulance Allocation based on Crash Prediction



Retail GeoPrediction

Boost Annual Revenues Predictive Geo-Targeted Marketing

ESRI Business Data

10,000 Features
per Block Group
(217K across US)



Demographics

More than 2,000 variables on accurate current-year estimates and five-year projection of US demographics including population, households, income, age, and ethnicity. Data on education, labor force, journey to work, marital status, languages spoken, home value, and more.

Data Source: Esri, US Census Bureau, and ACS

Sample Report: Demographics and Income Profile → [\[PDF\]](#)



Consumer Spending

What products and services US consumers are buying where? Includes expenditures by households for 20+ categories including apparel, food and beverage, financial, and more.

Data Source: Esri, Consumer Expenditure Surveys, and Bureau of Labor Statistics

Sample Report: Retail Goods and Services Expenditure → [\[PDF\]](#)



Tapestry Segmentation

Detailed market segmentation of US residential neighborhoods divided into 67 segments based on socioeconomic and demographic characteristics. Provides insights into customer lifestyle and behaviors.

Data Source: Esri

Sample Report: Tapestry Segmentation Area Profile → [\[PDF\]](#)



Market Potential

Data on products and services consumers use and demand. Expected number of consumers and Market Potential Index data for goods, services, attitudes, and activities.

Data Source: Esri and GfK MRI

Sample Report: Retail Market Potential → [\[PDF\]](#)



Business Locations and Summary

Data on more than 13 million businesses including name and location, industrial classification code, number of employees, and sales.

Data Source: Esri and Infogroup

Sample Report: Business Locations → [\[PDF\]](#)



Major Shopping Centers

Statistics for nearly 7,000 major US shopping centers including name, total retail sales, gross leasable area, location, and more.

Data Source: Directory of Major Malls, Inc.

Sample Report: Major Shopping Center Locator → [\[PDF\]](#)



Traffic Counts

Extensive US traffic data for trade analysis, routing, and mathematical modeling. Data on the number of vehicles that cross a certain point or street location in the US.

Data Source: Market Planning Solutions Inc.

Sample Report: Traffic Count Map - Close Up → [\[PDF\]](#)



Retail MarketPlace

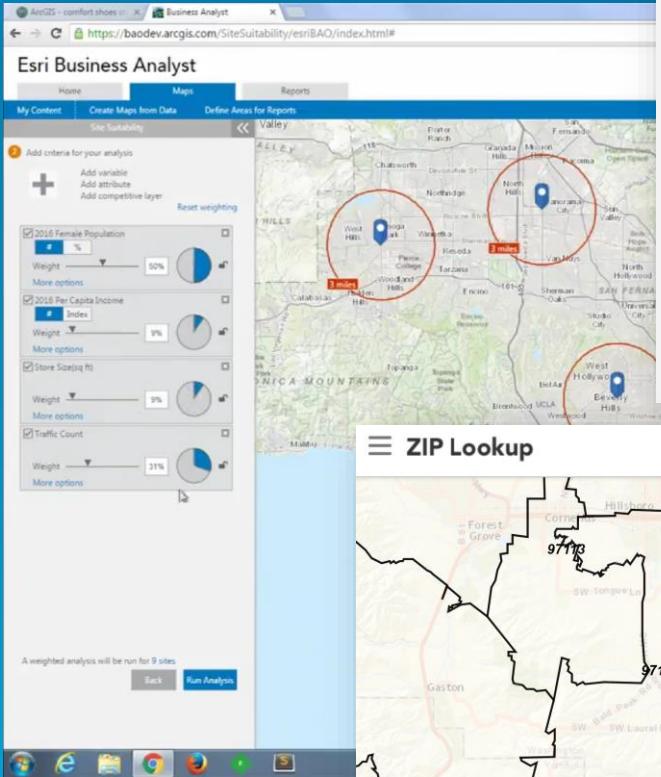
Compare retail sales and consumer spending to measure the gap between supply and demand. Learn where consumer spending might be leaking out of a market area and find opportunities to supply goods and services to new customers.

Data Source: Esri and Infogroup

Sample Report: Retail MarketPlace Profile → [\[PDF\]](#)

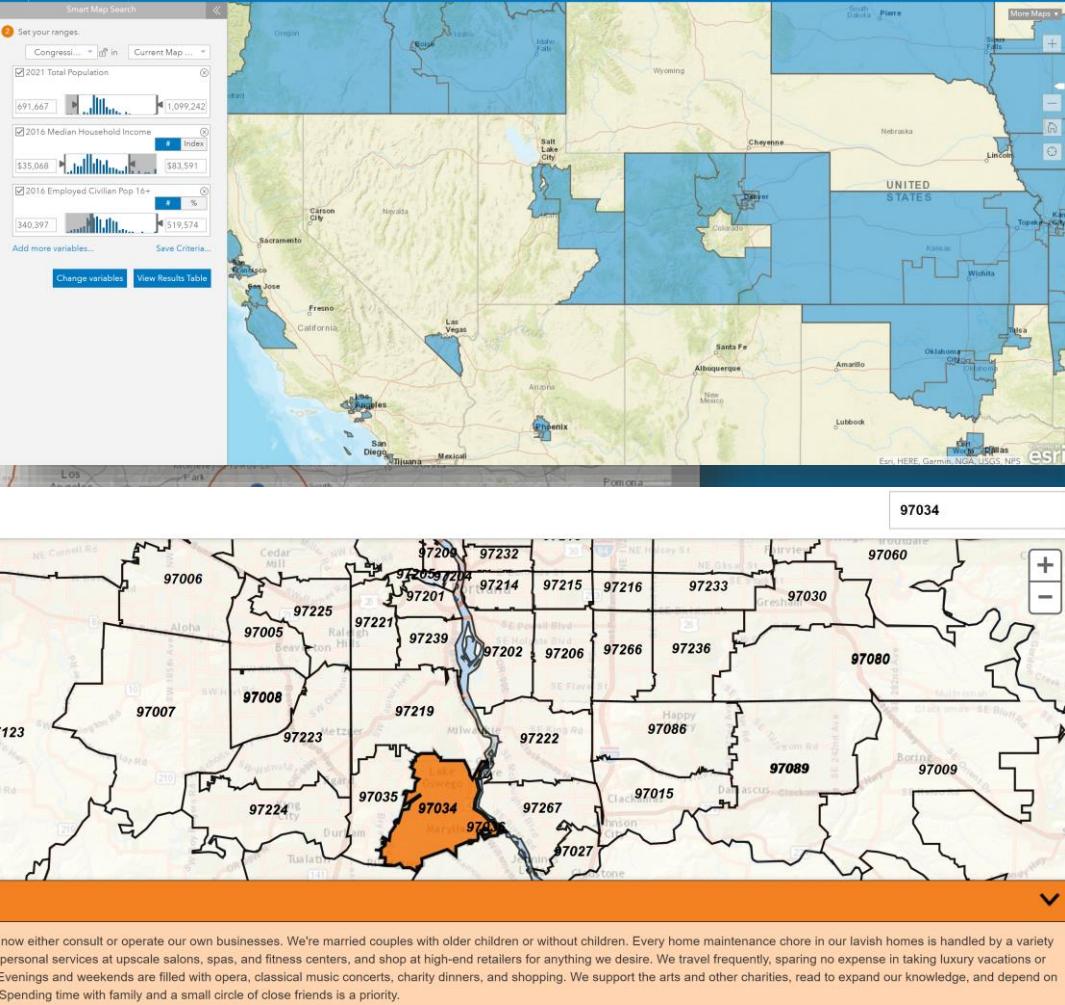
ESRI Business Data: Current Utilization

**Manual Assignment of 10K+ Features w/
Weights (Best Locations, Site Selection)**



**Exploring Tapestry
Segmentation per
Area
(Market Expansion)**

**Manual Filtering for 217K+ areas based Feature Values
(Targeted Promotion)**



Challenges

- Which Features are mostly affecting my Sales/Marketing? (10K+)
- What Weights should I assign per Feature?
- Do the Features differ per Area? To which extent?

Outcome

Significant Time spent on Manual Analysis with good value realized – but could be missing a complex pattern

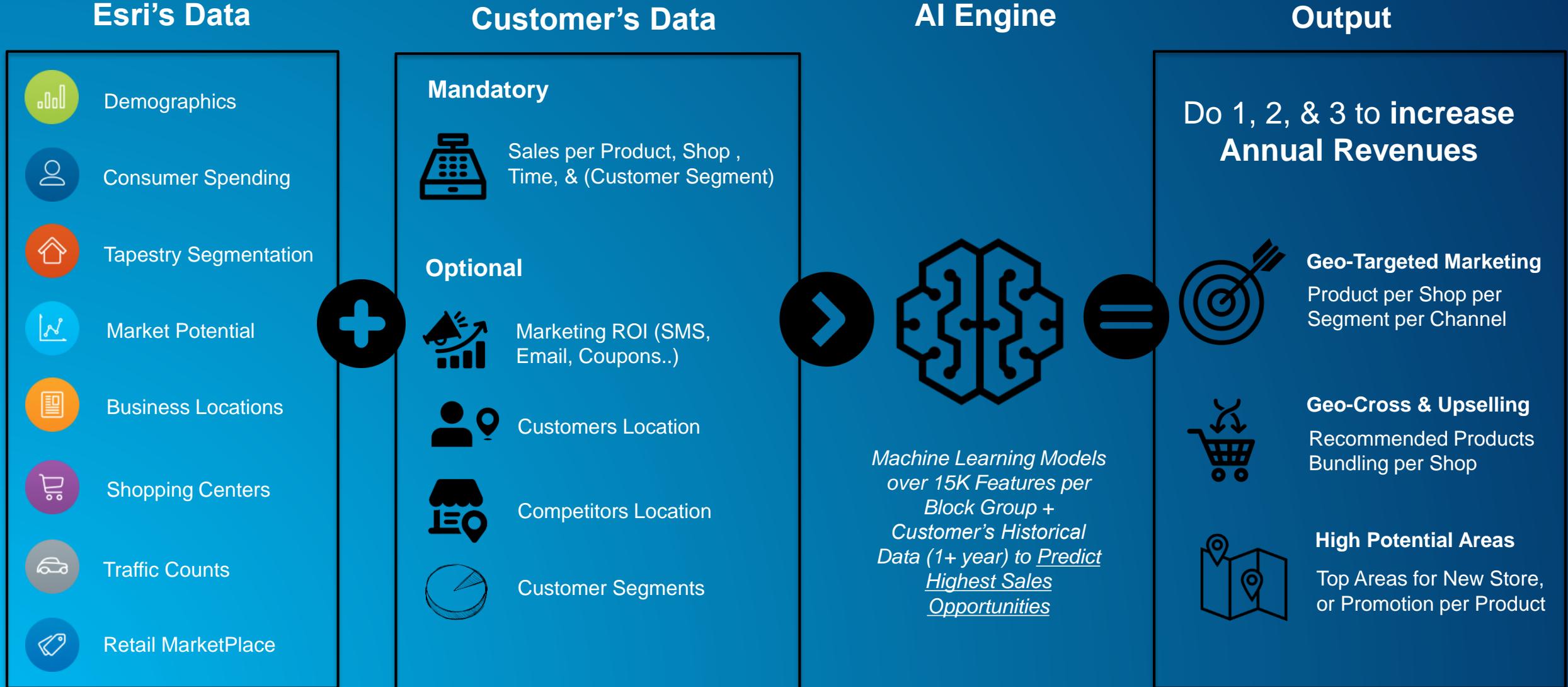
Solution

Machine Learning over Esri's data + Customer's Historical Sales data to **Predict Highest Sales Opportunities per 217K US Block Groups with Justification**

Machine Learning will be able to:

- **Accurately identify** which **features** among the 10,000 are most correlating to success/failure per product per area per time/season, together with the weight to be assigned per each of the selected features
- **Accurately Predict the locations** where:
 - A store is likely to generate high sales
 - Product A will sell great (in some cases we can predict the sales amount)
 - Product B could be bundled with Product C to generate higher sales
- Continuously learn from new data, optimize the algorithm, and increases the prediction accuracy

Esri's Data + Customer's Data = Predictions for boosting Annual Revenues

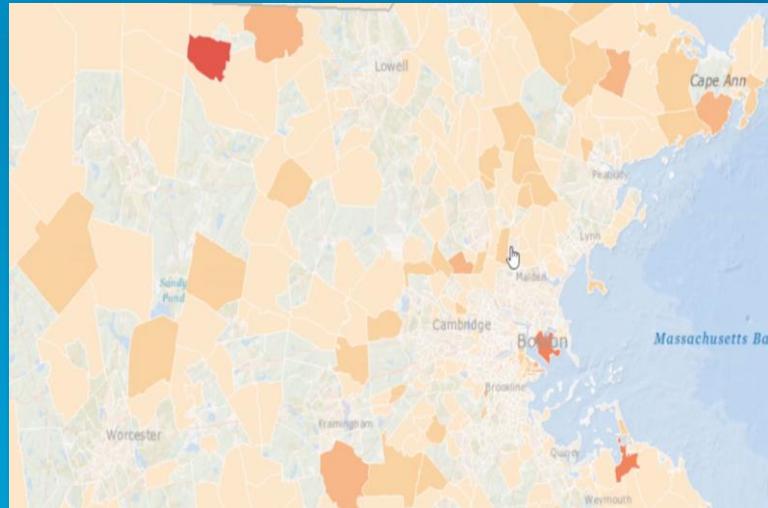


Predict Sales, Marketing ROI, Cross-Selling, and Potential per Location

With high Accuracy, Speed & Scale

Where should I Promote for Product XYZ?

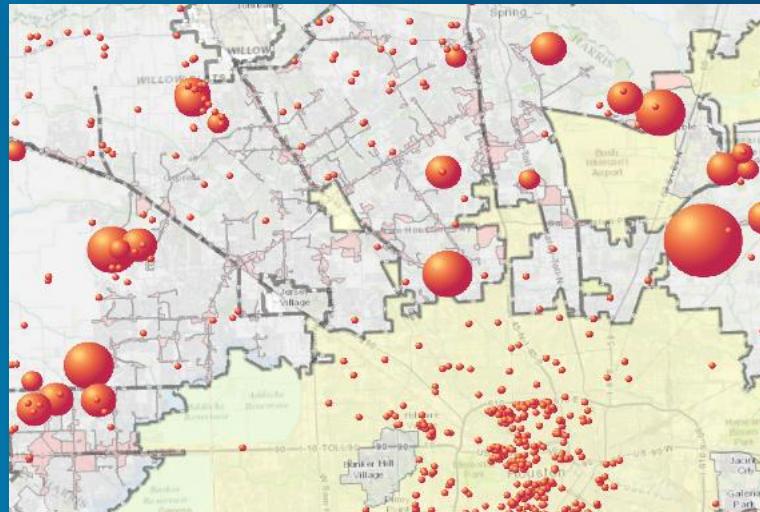
What's the expected Return? Why?



- **Areas:** 1,2 & 3
- **Predicted Sales:** \$12.7M
- **Reasons:** 55K Hispanic Male Teenagers 12-17, 75% Comfortable Empty Nesters (Tap), 10+ mins Away from 70% of Competitors

Where should I open a new store?

What's the expected Sales? Why?



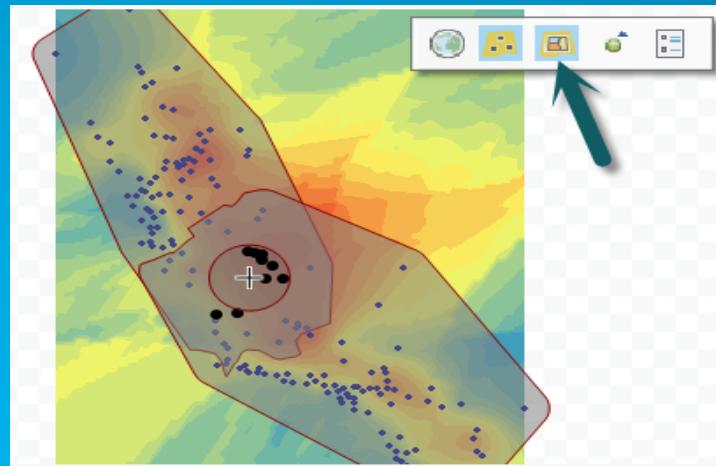
- **Block G: 31B**
- **Predicted Sales:** \$2.1M
- **Reasons:** 125K Employed Female 25-28s whose income > \$80K, 2016 “Food Away from Home - Dinner at Full Service Restaurant” Index 125, 67% Diners & Miners

Predicting Seagrass Density Globally

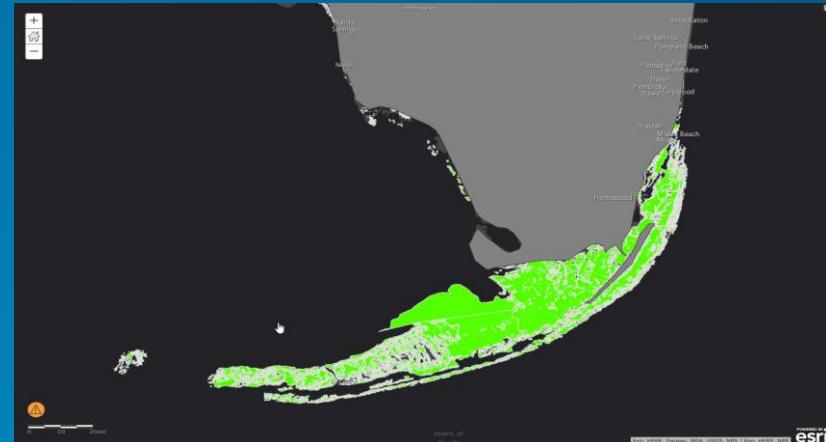
Seagrass reduces global warming, storing up to 100 times more CO₂ vs tropical forests.



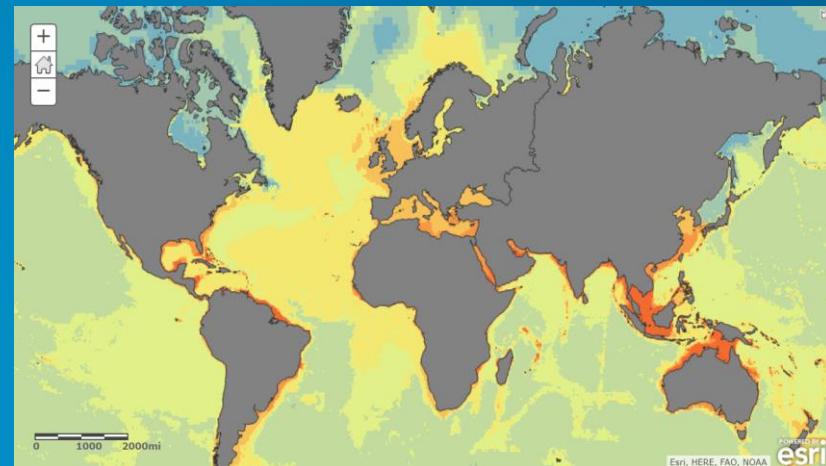
EBK interpolates discrete measurement data into statistically valid continuous surfaces



Seagrasses around Florida Coast



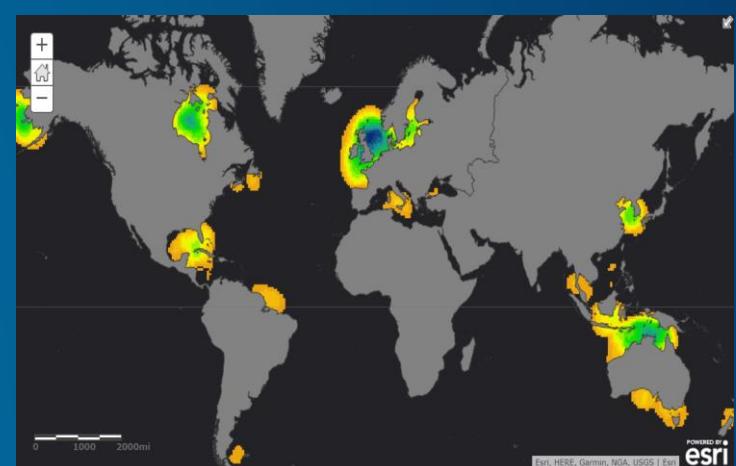
Interpolated Ocean Conditions Globally
(Temperature, Salinity, Phosphate..)



Features Highly correlated with Seagrass Existence



Global prediction for seagrass occurrence via Random Forests





esri

**THE
SCIENCE
OF
WHERE**

MICROSOFT AI PLATFORM

Azure & AI

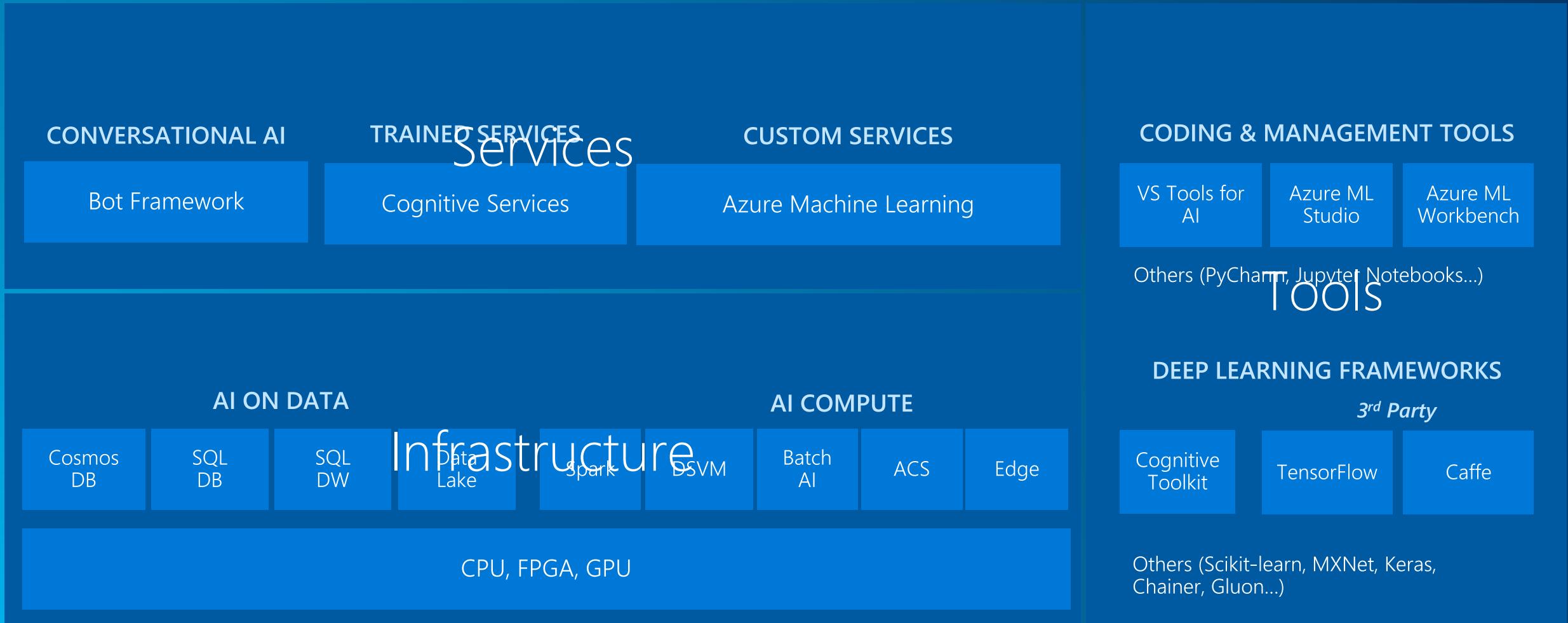
Bringing AI productivity
to every organization
and every scenario

Services

Infrastructure

Tools

The Microsoft AI platform



I want to use AI – How can I get started?

I want to use AI – How can I get started?



Cognitive
Services



Custom Services
Bring your Own Data



Custom
AI Models



Flexibility

Cognitive Services



Vision



Language



Speech



Search



Knowledge

Computer vision

Face

Emotion

Content Moderator

Video

Video Indexer

Text analytics

Spell check

Web language model

Linguistic analysis

Translator

Speaker recognition

Speech

Web search

Image search

Video search

News search

Autosuggest

Academic knowledge

Entity linking service

Knowledge exploration

Recommendations

QnA maker

Cognitive Services Labs

Custom

Vision Service

Custom

Language
Understanding

Custom

Speech Service

Custom

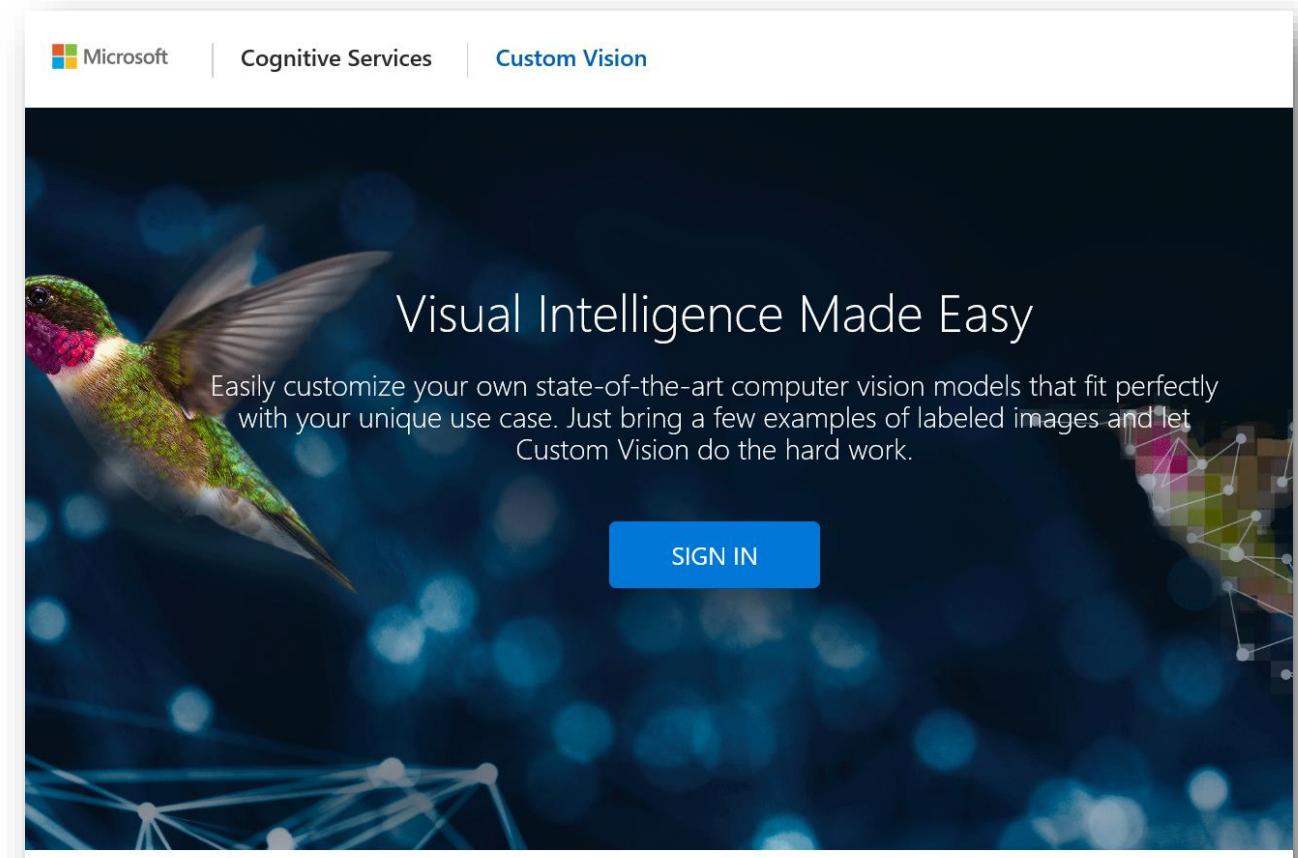
Search

Custom

Decision Service

Demo Custom Vision

Wee Hyong Tok



Big Data, Cloud and AI
Customizing for your use case

Recipe for Jumpstarting AI

- Lots of experimentation
- Lots of compute power, often GPUs
- Lots of training data

Need to run Training at scale!

Example Tools / Dev Environments

- Azure Machine Learning Studio: Serverless collaborative drag-and-drop tool for graphical machine learning development
- Azure Machine Learning Workbench: Visual AI powered data wrangling, experimentation, and lifecycle management
- Visual Studio Code Tools for AI: Build, debug, test, and deploy AI with Visual Studio Code on Windows and Mac
- Azure Notebooks: Organize your datasets and Jupyter Notebooks in a centralized library for Data Science and Analysis
- Deep Learning Virtual Machine! Session this afternoon: A pre-configured environment for deep learning using GPU instances

What do I need to build custom models?

Toolkits

Cognitive Toolkit

Caffe



TensorFlow™

mxnet

And more....

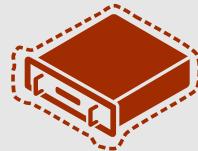
Environment



Desktop / Laptop



Cloud



Virtual Machine

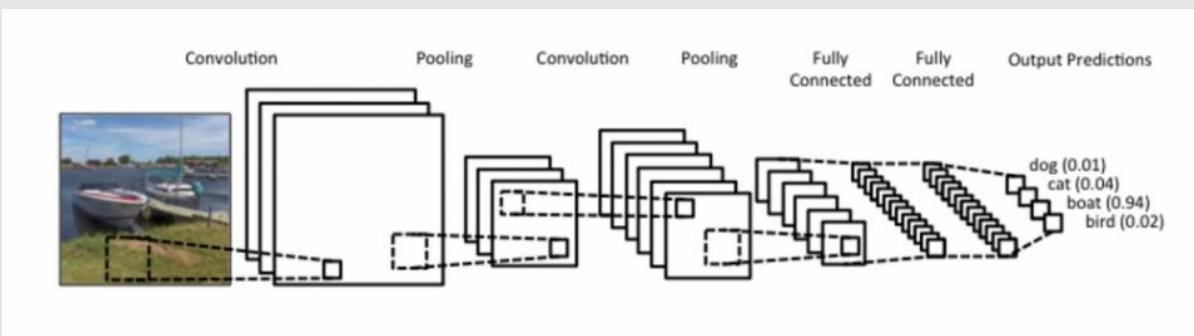


Devices / Edge

What are the common models?

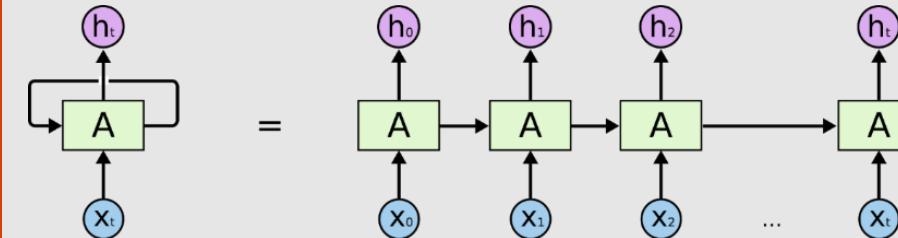
CNN

Convolutional Neural Network

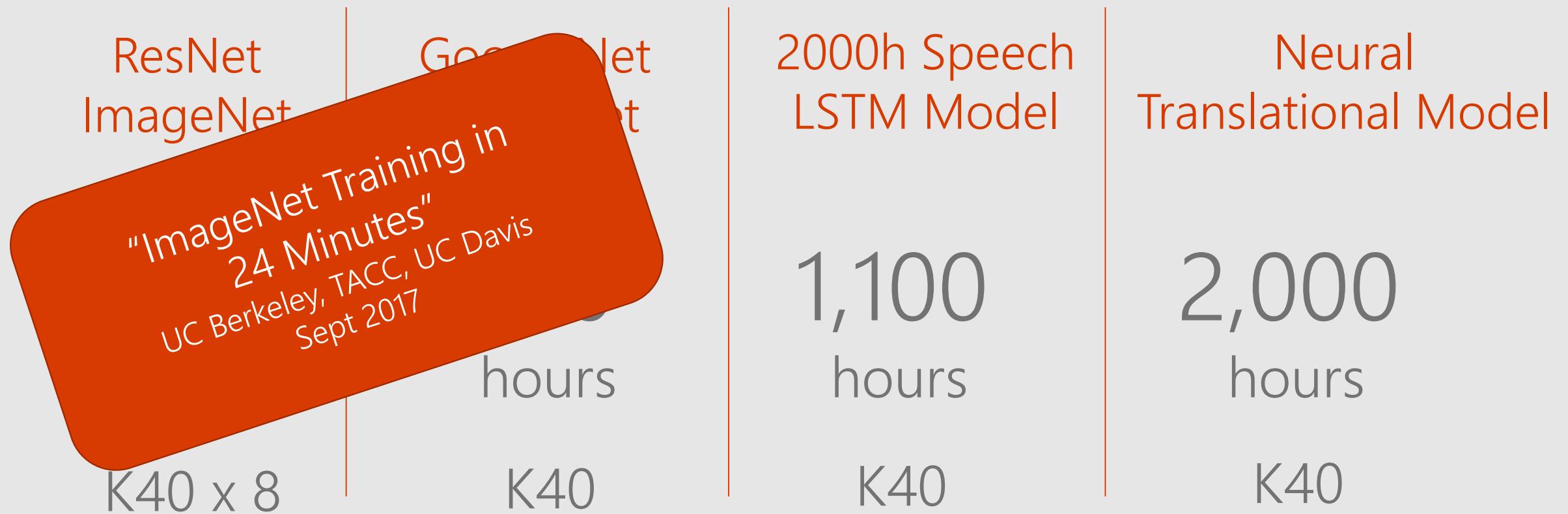


RNN

Recurrent Neural Network



How long does it take to train DNN models?



Imagenet: 1M Images, 1K Classes

What's the magic behind
getting started?

Lung Cancer Detection

Based on work by
Miguel Fierro, Ye Xing, Tao Wu



Kaggle Data Science Bowl 2017



Determine
whether a
patient
has **cancer**

Jumpstart in less than 1h
Top 10% (Jan 19th)

Solution Pieces

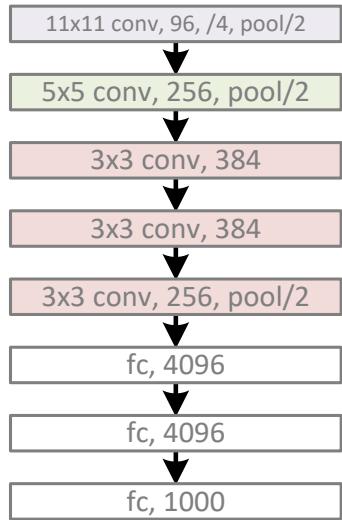
Data Science VM (GPU)

Microsoft Cognitive Toolkit (CNTK)

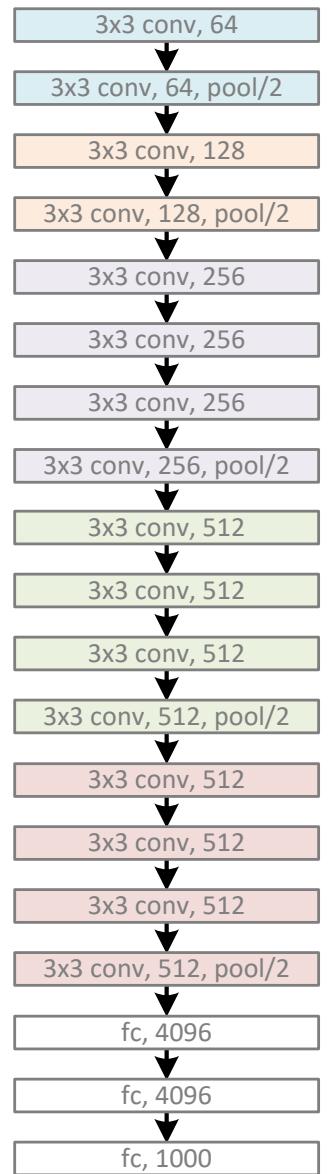
LightGBM for classification

Deep Neural Nets

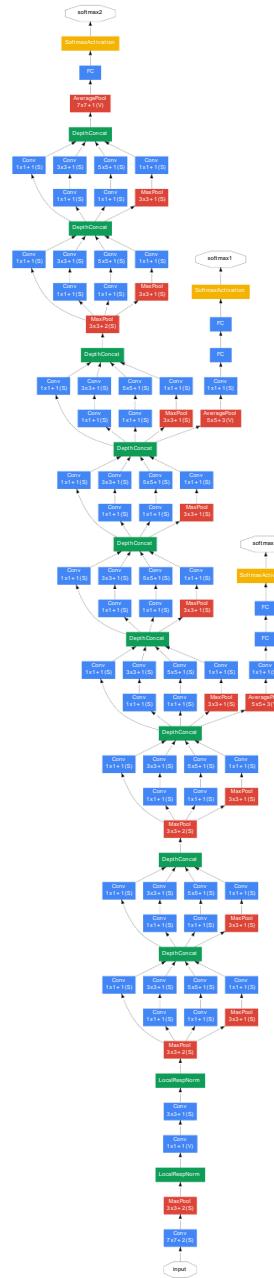
AlexNet, 8 layers
(ILSVRC 2012)



VGG, 19 layers
(ILSVRC 2014)

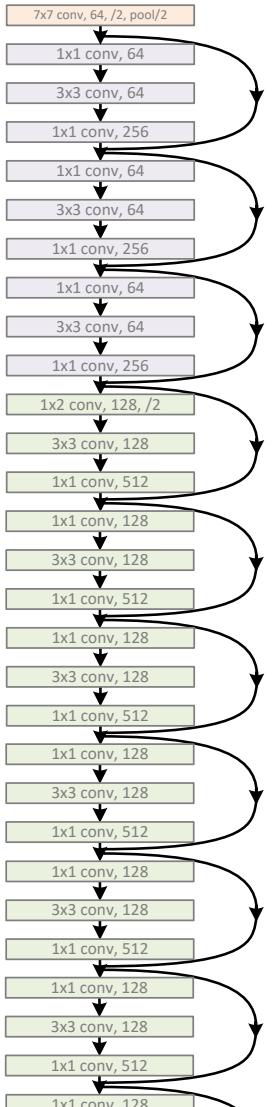


GoogleNet, 22 layers
(ILSVRC 2014)

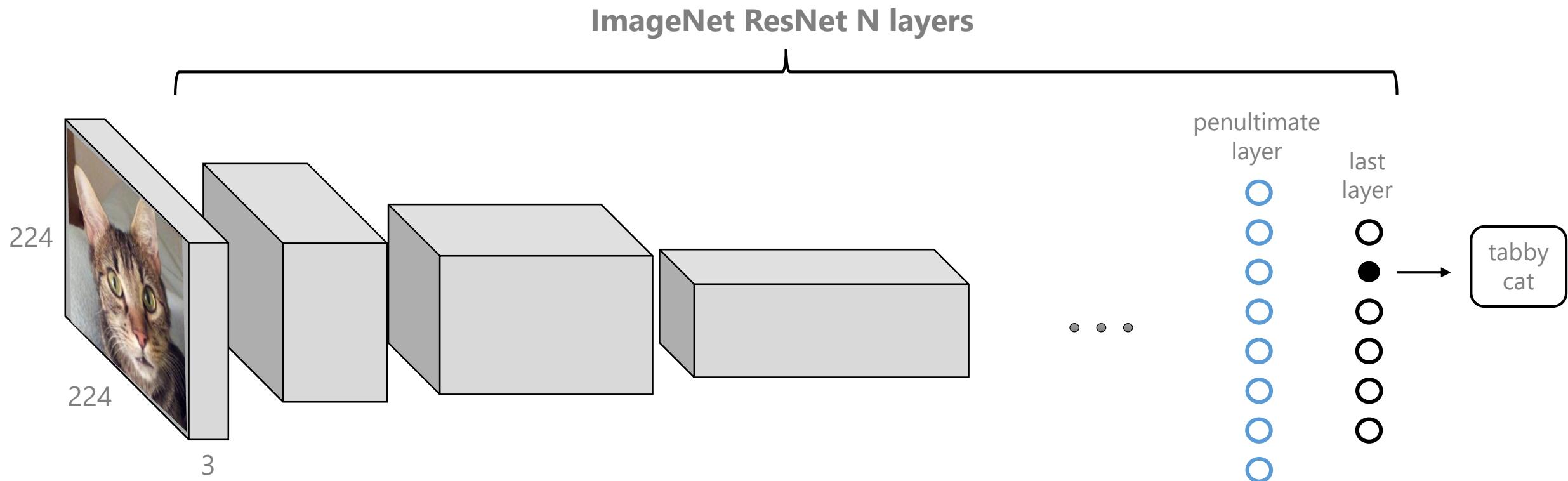


ResNet, 152 layers

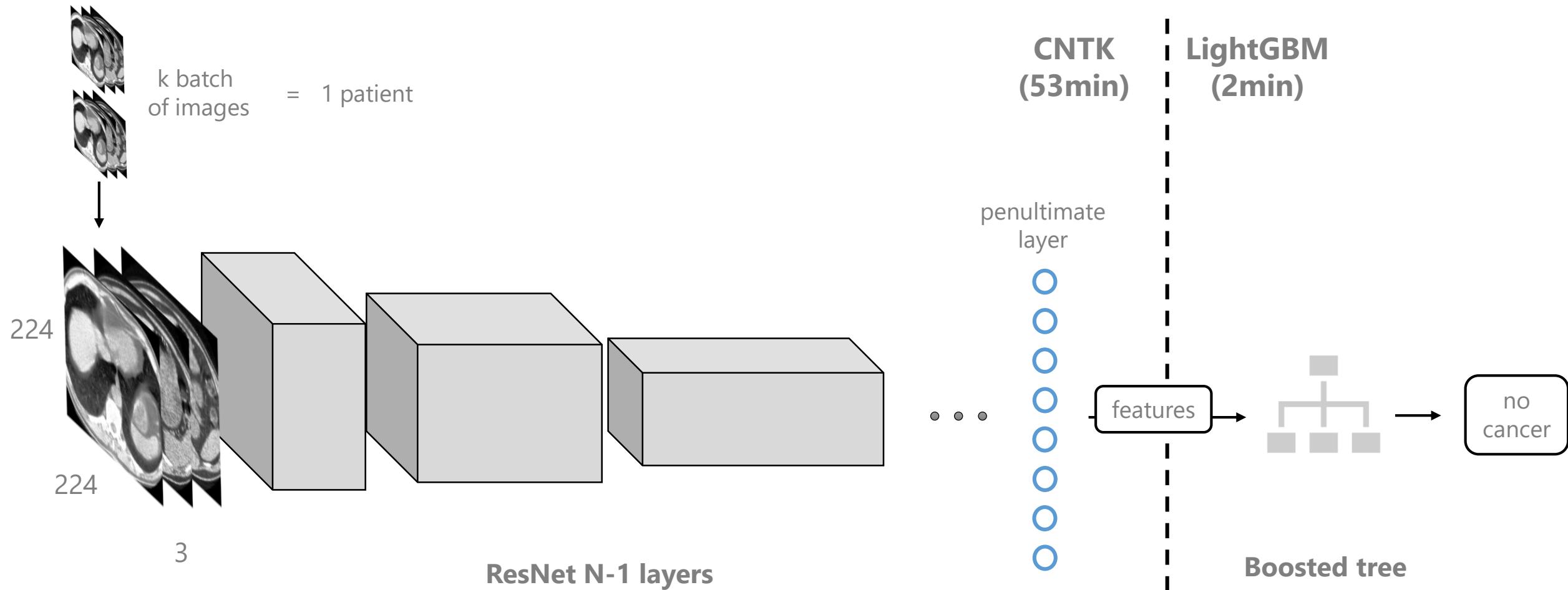
Microsoft



Solution: Transfer Learning

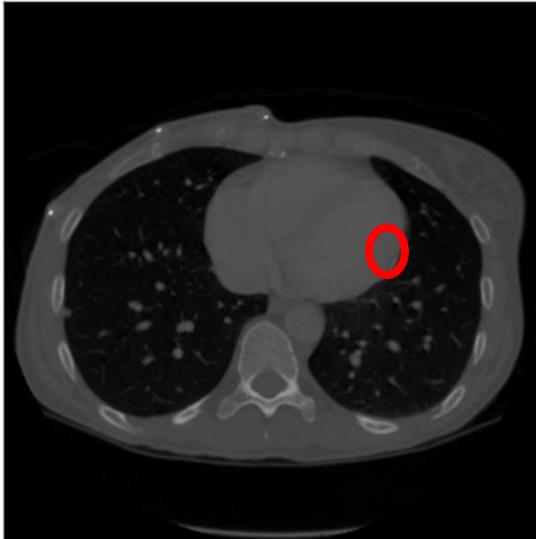


Solution: Transfer Learning



Challenges and Improvements

Nodule detection
+
classification



Src:
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.705.8152&rep=rep1&type=pdf>

- Use other pretrained CNNs such as [AlexNet](#), [AlexNet with Batch Normalization](#) and [ResNet with 18 layers](#).
- Perform image augmentation: You can try to increment the training set by performing transformations in the images. For that you can apply different filters or rotate them.
- Use a customize network with 3D images: CNTK allows [3D convolutions](#). On GPU, 1D, 2D and 3D convolutions will use cuDNN (fast), all other convolutions will use reference engine (slow). You can create a CNN that accepts 3D images.
- Try some feature engineering: Before training the tree, the features are averaged. You can try to feed the tree without this operation or try a different one.

How do I get started?



Kaggle script

<https://aka.ms/dsb2017-cntk-script>



Cortana Gallery
notebook

<https://aka.ms/dsb2017-cntk-notebook>



Microsoft | TechNet

Blog in TechNet

<https://aka.ms/dsb2017-cntk-blog>

Scoring using SQL Server 2017

SQL Server 2017 is the latest version of the Microsoft SQL database.

Supports Python and R scripts as stored procedures

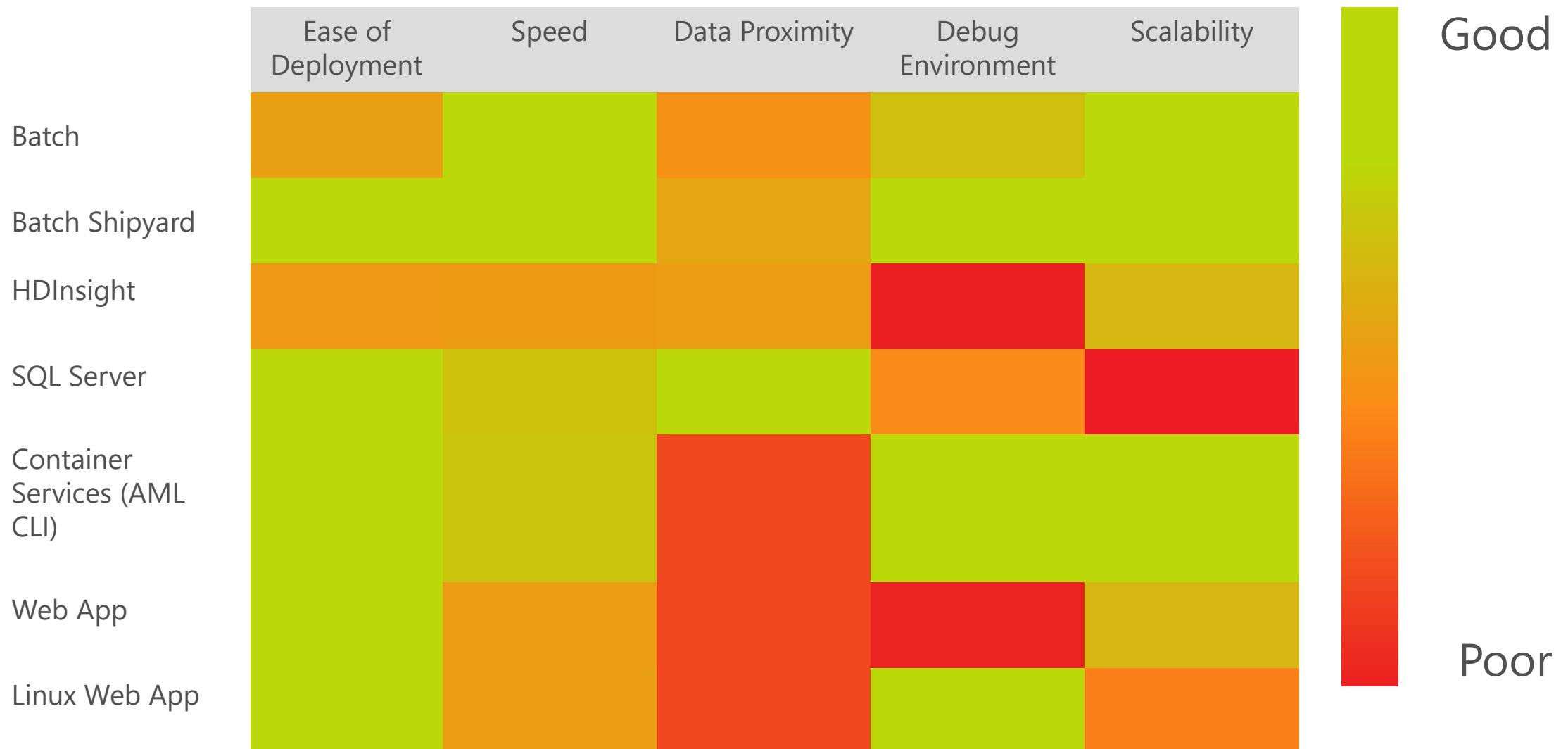
Fully leverage the data management of the database to minimize I/O latency

Example: [Lung cancer detection on SQL Server 2017](#)

Operationalization

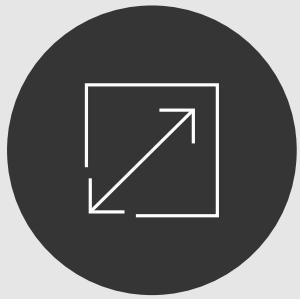


Operationalization



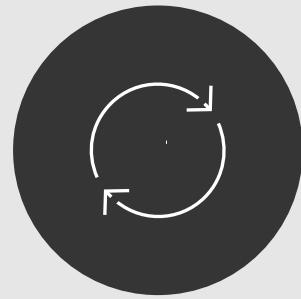
Azure Machine Learning

Benefit from the fastest AI developer cloud



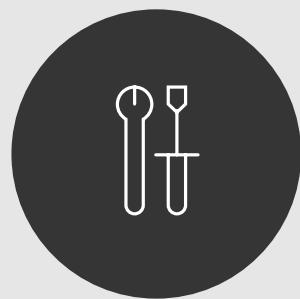
Workbench

Build, deploy, and manage at scale



Experimentation

Boost productivity with agile development

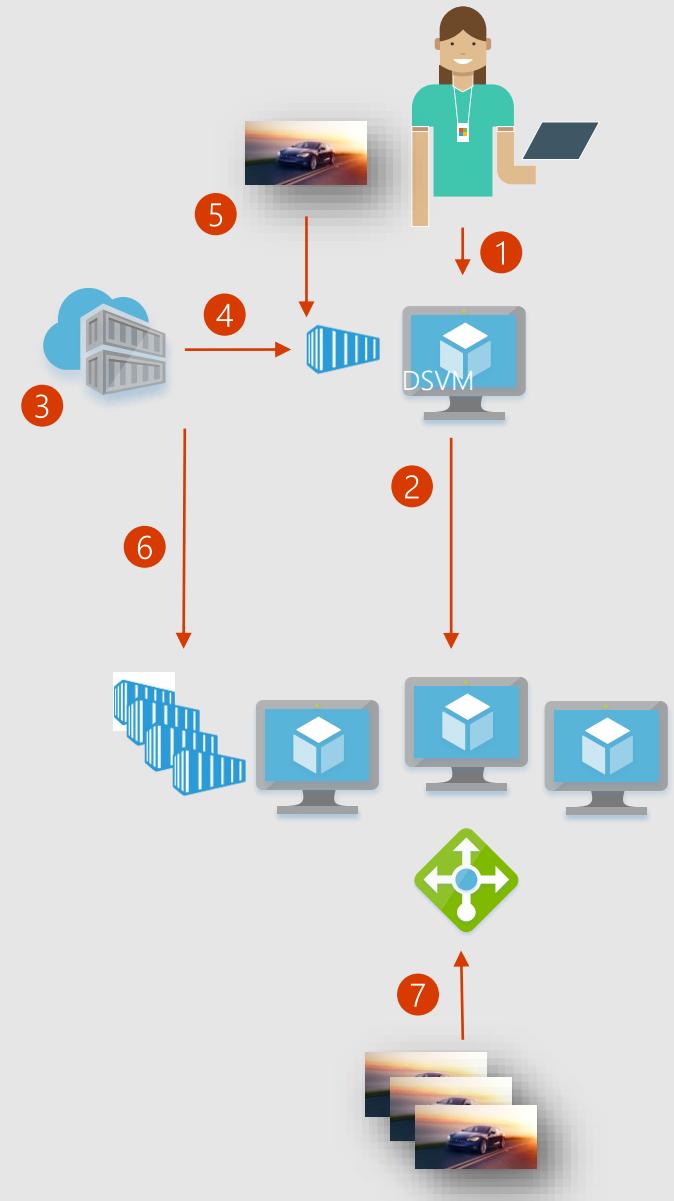


Model Management

Begin building now with the tools and platforms you know

Example Deployment workflow

1. Create a driver file for a trained DNN
2. Setup of the cluster from AZ ML CLI
3. Create a web-service, image uploaded to Docker Registry
4. Deploys web-service locally
5. Test locally
6. Deploy to cluster
7. Send requests to web-service



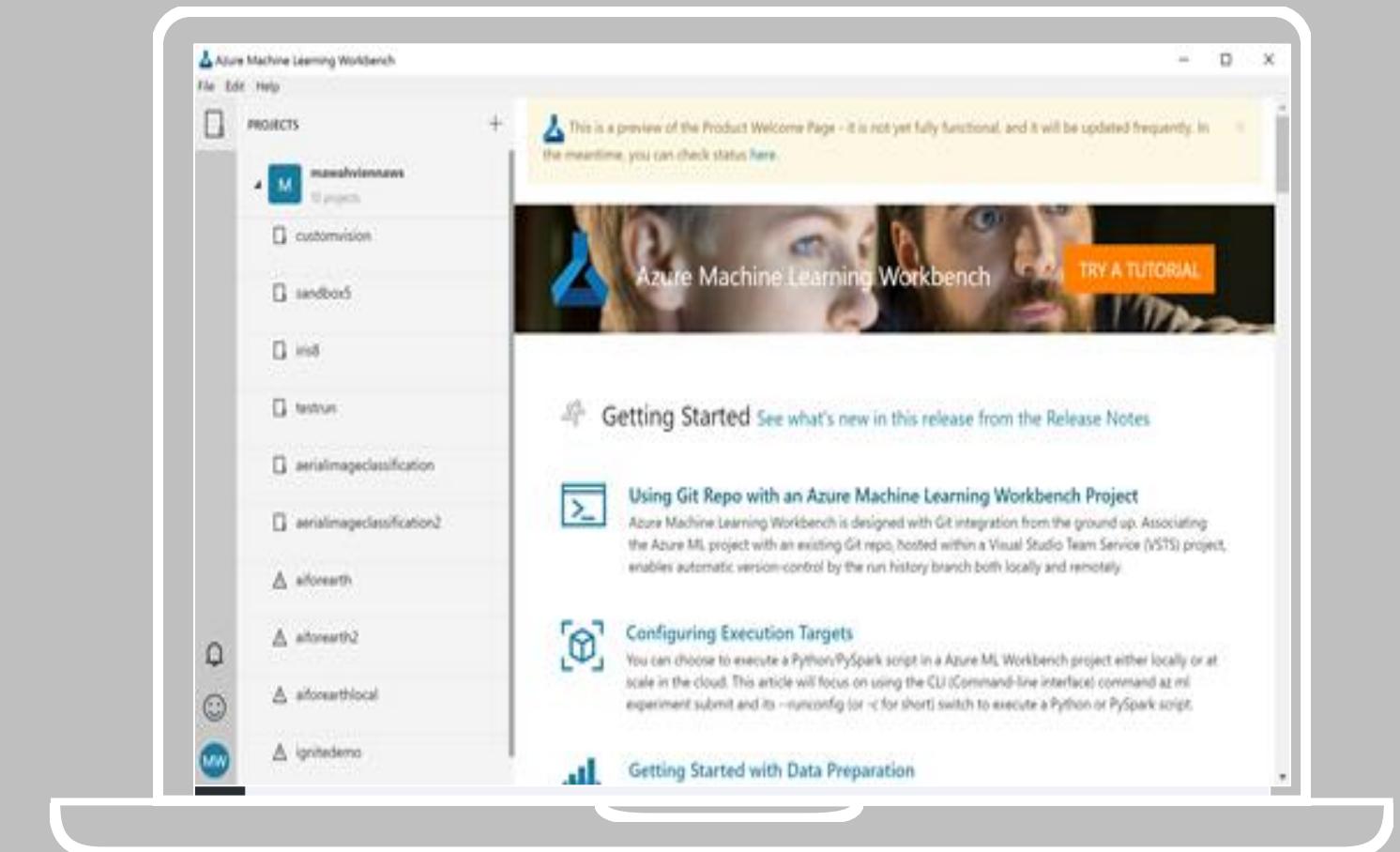
Key Recommendations

- Newly released Azure Machine Learning for development, either with jupyter Notebooks or Visual Studio Code Tools for AI
- Azure Batch AI Service for programmatic and scheduled batch workloads
- Azure ML Operationalization for real-time scoring
- Bring compute to your data with HDInsight on Spark or SQL Server 2017

Demo

Azure Machine Learning

Danielle Dean



Practical tips for ...

Putting Big Data, Cloud and AI to use

Being Obsessed with Data

Can only complete the process with the right data!

Question is sharp.

Data measures what you care about.

Data is accurate.

Data is connected.

A lot of data.

E.g. Predict whether component X will fail in the next Y days

E.g. Identifiers at the level you are predicting, relevant data collected & feature engineering using domain knowledge

E.g. Failures are really failures, human labels on root causes

E.g. Machine information linkable to usage information

E.g. Will be difficult to predict failure accurately with few examples

Examples for a predictive maintenance project

Asking the right questions... and how will it be used??

Business scenario	Key decision	Data Science question
Energy forecasting	Should I buy or sell energy contracts?	What will be the long/short-term demand for energy in a region?
Customer churn	Which customers should I prioritize to reduce churn?	What is probability of churn within X days for each customer?
Personalized marketing	What product should I offer first?	What is the probability that customer will purchase each product?
Product feedback	Which service/product needs attention?	What is social media sentiment for each service/product?

Defining Performance Metrics

Establish a
**Qualitative
Objective**

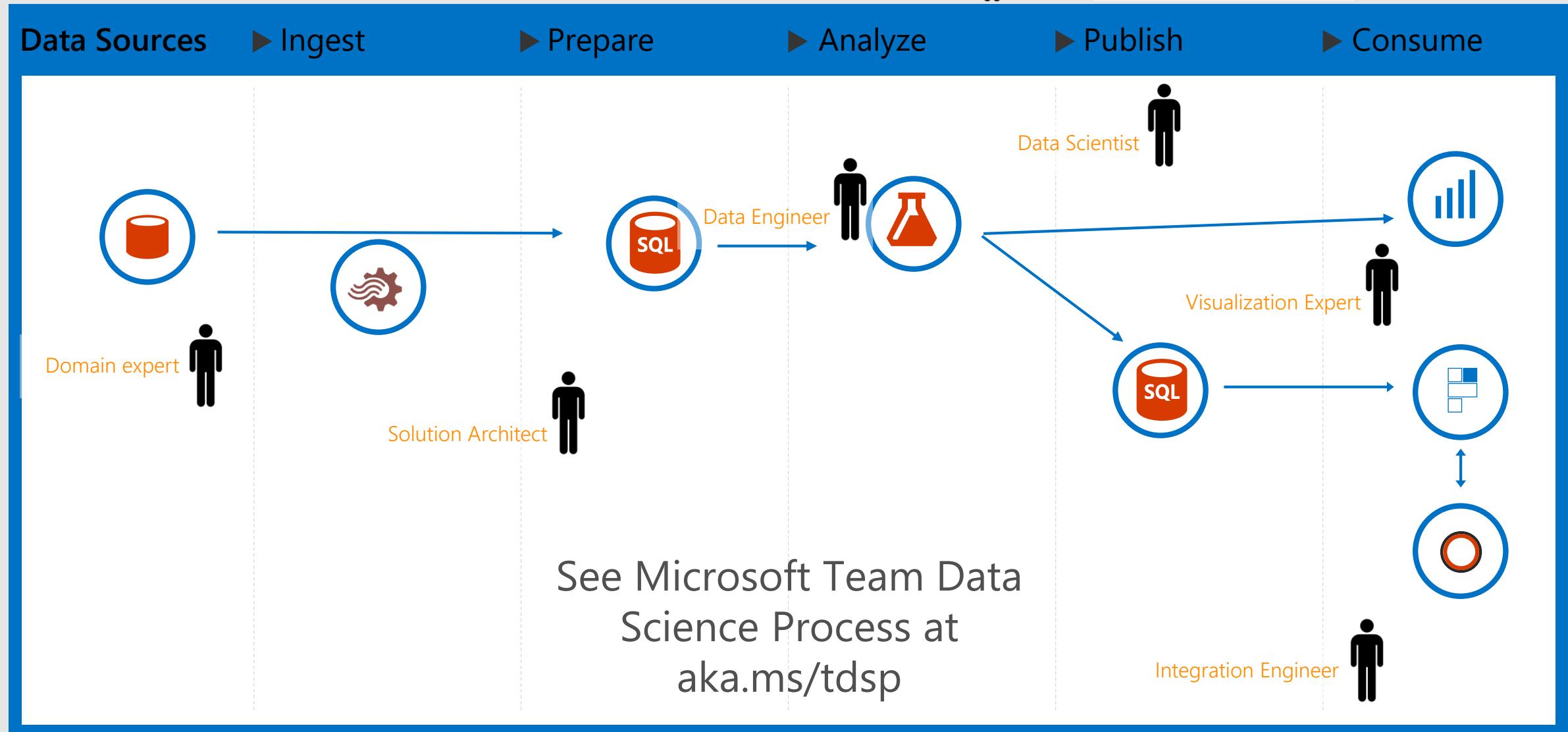
Translate into
**Quantifiable
Metric**

Quantify the
**metric value
improvement**
useful for
customer scenario

Establish a
baseline

Establish how to
measure the
improvement in
the metric with
the data science
solution

It's a team!



AI Solution Templates

A thumbnail showing a hand writing on a document labeled 'LOAN' with a green checkmark and a red X, connected to a computer monitor displaying a chart and a database icon.

SOLUTION

Loan Credit Risk with SQL Server

Using SQL Server 2016 with R Services, a lending institution can make use of predictive analytics to reduce number of loans they offer to those borrowers most likely to default, increasing the profitability of their loan ...

1.3K 343 24 days ago

Microsoft

A thumbnail showing a computer monitor displaying a chart and a person icon.

SOLUTION

Personalized Offers

In today's highly competitive and connected environment, modern businesses can no longer survive with generic, static online content. Furthermore, marketing strategies using traditional tools are often expensive, ...

4.6K 399 3 months ago

Microsoft

A thumbnail showing a hand holding a pen over a document, a clock, a calendar, and a dog icon.

SOLUTION

Campaign Optimization with SQL Server

This solution demonstrates how to build and deploy a machine learning model with SQL Server 2016 with R Services to recommend actions to maximize the purchase rate of leads targeted by a campaign.

8.1K 1.2K 2 hours ago

Microsoft

A thumbnail showing a hand holding a pen over a document, a clock, a calendar, and a hospital building icon.

SOLUTION

Campaign Optimization with Azure HDInsight Spark Clusters

This solution demonstrates how to build and deploy a machine learning model with Microsoft R Server on Azure HDInsight Spark clusters to recommend actions to maximize the purchase rate of leads targeted by a...

1.2K 157 18 days ago

Microsoft

A thumbnail showing a hospital building, an ambulance, and a medical icon.

SOLUTION

Predicting Length of Stay in Hospitals

This solution enables a predictive model for Length of Stay for in-hospital admissions. Length of Stay (LOS) is defined in number of days from the initial admit date to the date that the patient is discharged from any ...

10K 1.2K 25 days ago

Microsoft

A thumbnail showing a whiteboard with a line graph.

SOLUTION

Demand Forecasting and Price Optimization

Pricing is recognized as a pivotal determinant of success in many industries and can be one of the most challenging tasks. Companies often struggle with several aspects of the pricing process, including accurately fo...

4.4K 670 3 months ago

Microsoft

A thumbnail showing a clipboard with a checklist.

SOLUTION

Quality Assurance

Quality assurance systems allow businesses to prevent defects throughout their processes of delivering goods or services to customers. Building such a system that collects data and identifies potential problems alone...

1.9K 313 3 months ago

Microsoft

A thumbnail showing a car icon with a cloud and network connections.

SOLUTION

Telemetry Analytics

Super computers have moved out of the lab and are now parked in our garage! These cutting-edge automobiles contain a myriad of sensors, giving them the ability to track and monitor millions of events every sec...

9.2K 1.4K 3 months ago

Microsoft

A thumbnail showing a wind turbine and a line graph.

SOLUTION

Demand Forecasting

Accurately forecasting spikes in demand for products and services can give a company a competitive advantage. This solution focuses on demand forecasting within the energy sector

9.5K 1.6K 3 months ago

Microsoft

A thumbnail showing an airplane icon with gears.

SOLUTION

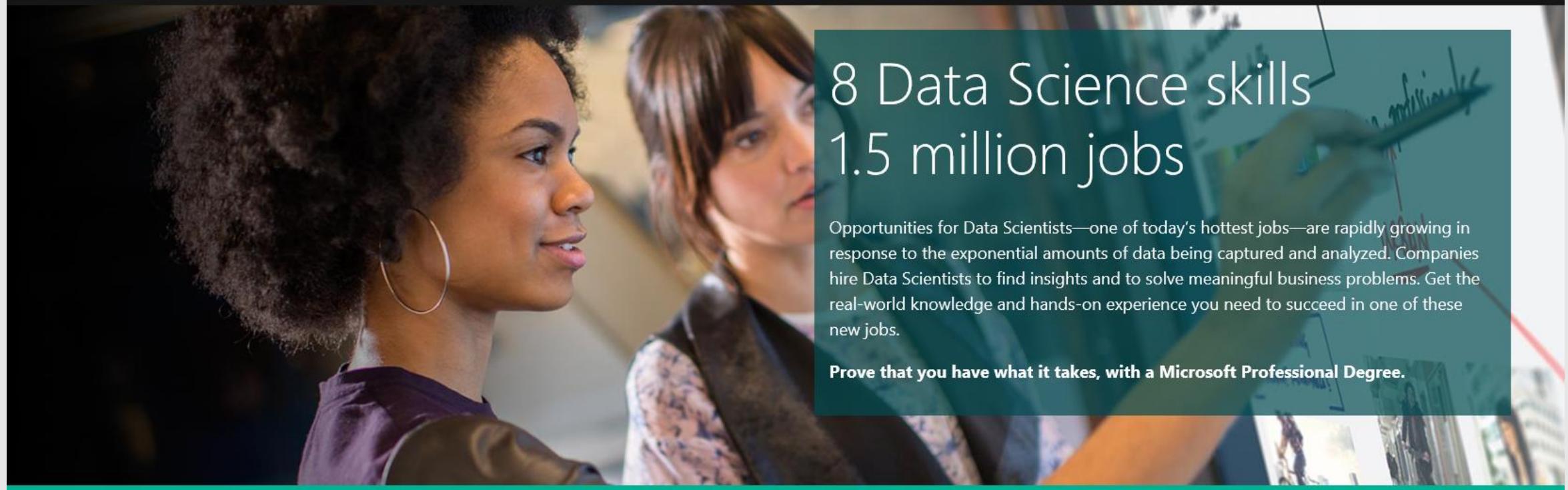
Predictive Maintenance

This Predictive Maintenance solution monitors aircraft and predicts the remaining useful life of aircraft engine components

11K 2.1K 3 months ago

Microsoft

- Discover on [AppSource](#) [Azure.Microsoft.com](#) [CI Gallery](#)
- Deploy in minutes
- Integration with VS Code
- GitHub Samples
- Configure/Customize
- Implement with Partners



8 Data Science skills 1.5 million jobs

Opportunities for Data Scientists—one of today's hottest jobs—are rapidly growing in response to the exponential amounts of data being captured and analyzed. Companies hire Data Scientists to find insights and to solve meaningful business problems. Get the real-world knowledge and hands-on experience you need to succeed in one of these new jobs.

Prove that you have what it takes, with a Microsoft Professional Degree.



BETA RELEASE

Microsoft Professional Degree in Data Science

Microsoft consulted Data Scientists and the companies that employ them to identify the requisite core skills. We then developed a curriculum to teach these functional and technical skills, combining highly rated online courses with hands-on labs, concluding in a final capstone project. Graduates earn a Microsoft Professional Degree in Data Science—a digitally sharable, résumé-worthy credential.

The program opens soon! To be notified, submit your email address.

Enter your email address

Submit

<https://academy.microsoft.com/en-US/professional-degree/data-science/>

Summary

Big Data, Cloud and AI Use Cases

What technology/service
should I use to get started?

Going deeper – custom
solutions

Putting it together