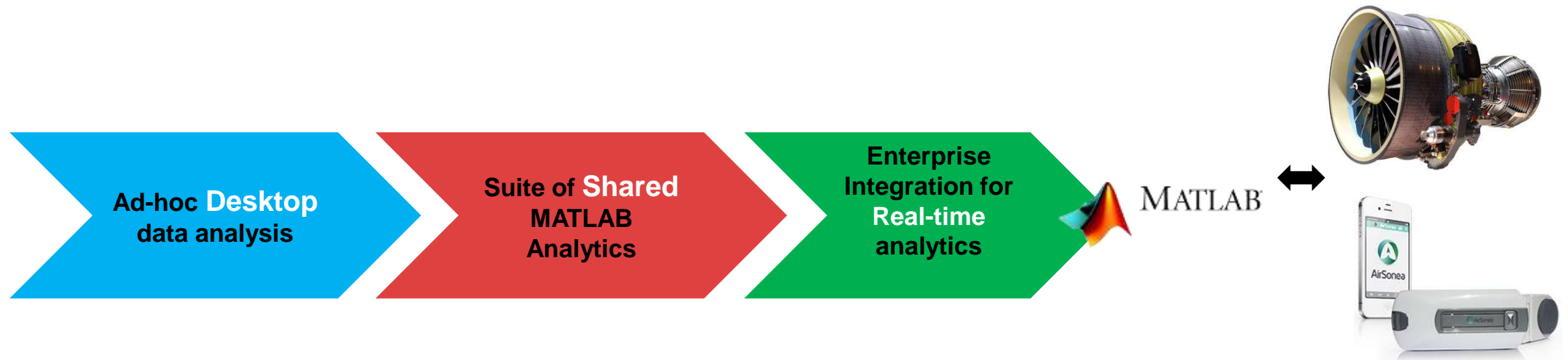
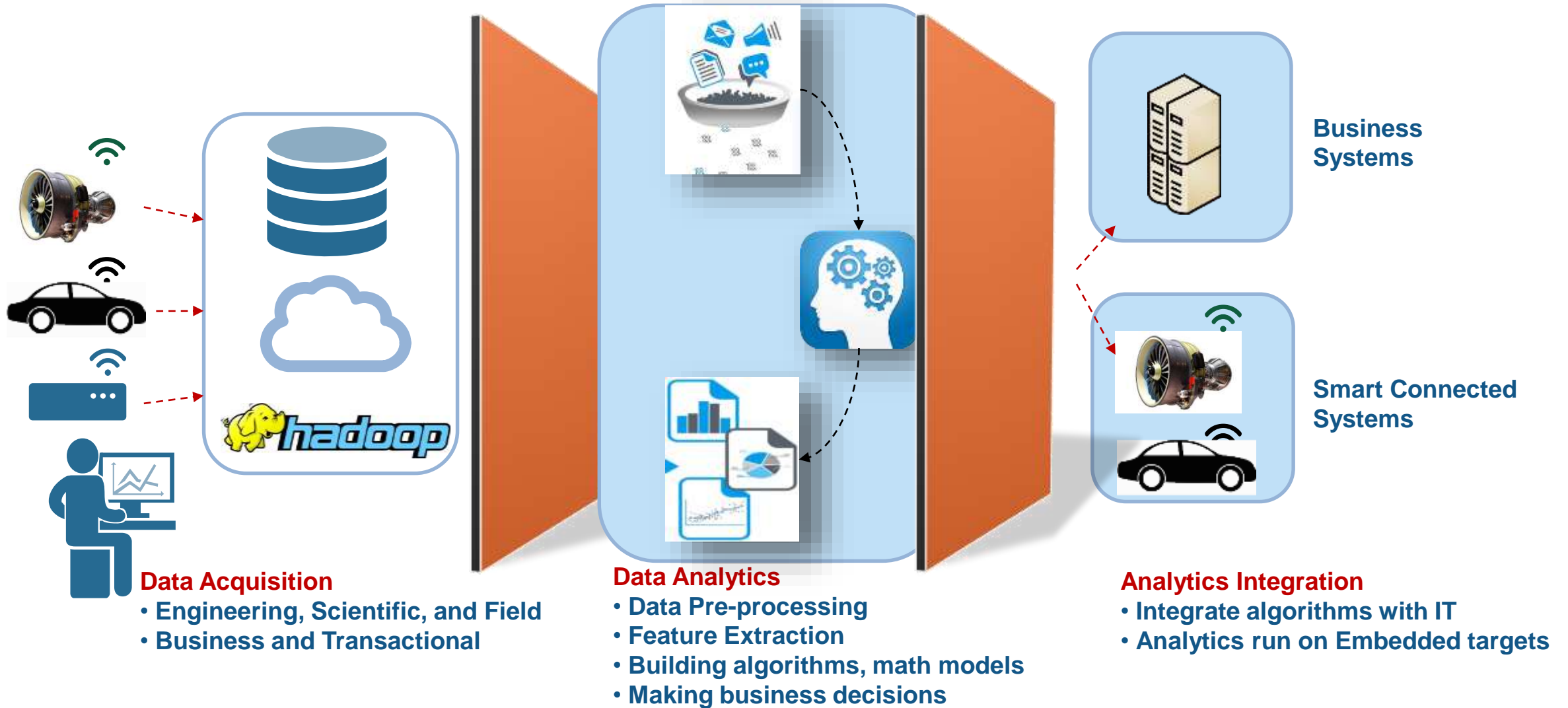


# MATLAB for Data Analytics and Machine Learning

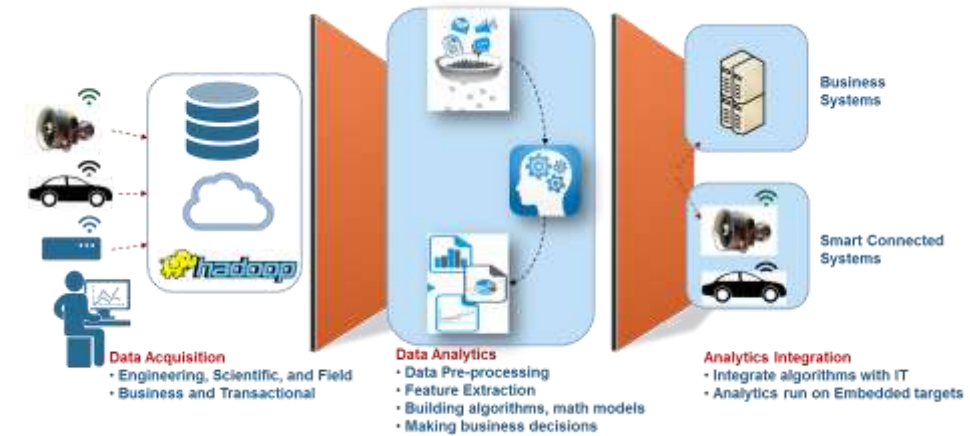


**Sundar Umamaheshwaran**  
**Amit Doshi**  
**Application Engineer-Technical Computing**

# Data Analytics Workflow



# Integrated Analytics— Success Stories



1. Taking Business Decisions Using **Historical** Data
2. Condition Monitoring On **Live** Data
3. Taking Analytics To **Embedded** Device

# Success Story 1: Daimler - Data Driven Fuel Cell Vehicle Design

## Challenge

- Understand vehicle **usage patterns**
- **Plan** hydrogen **refueling** infrastructure
- **Understand** how driving **patterns** affect vehicle **performance**

## Solution

- Connect to data using Database Toolbox
- Use MATLAB to explore data and identify insights
- Visualize data on charts and maps and share via automated reports and web applications

## Results

- Millions of miles of drive files translated into meaningful insights

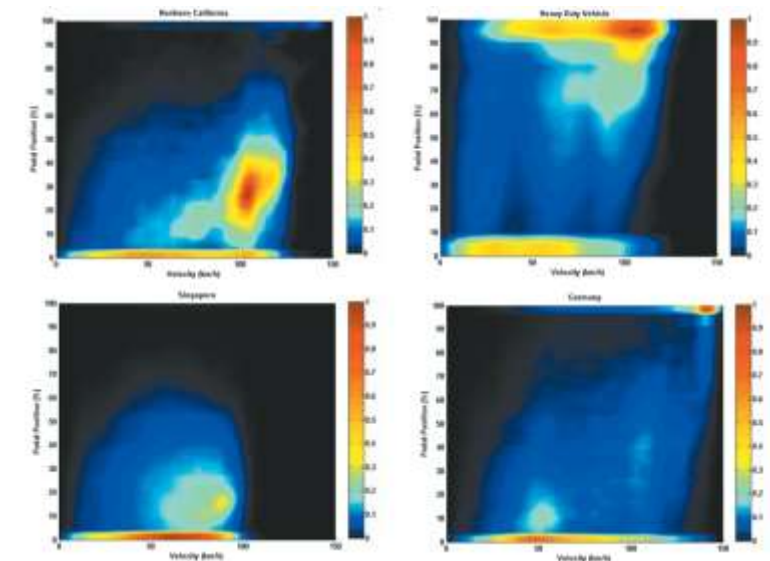


Vehicle health & troubleshooting



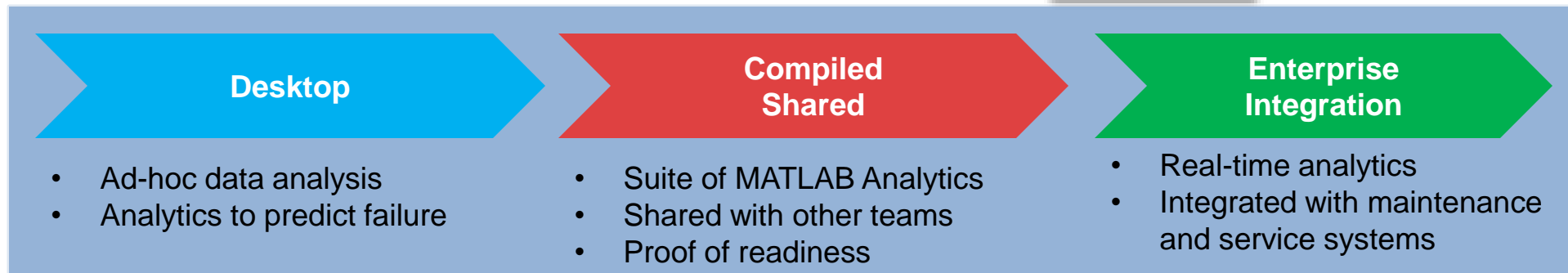
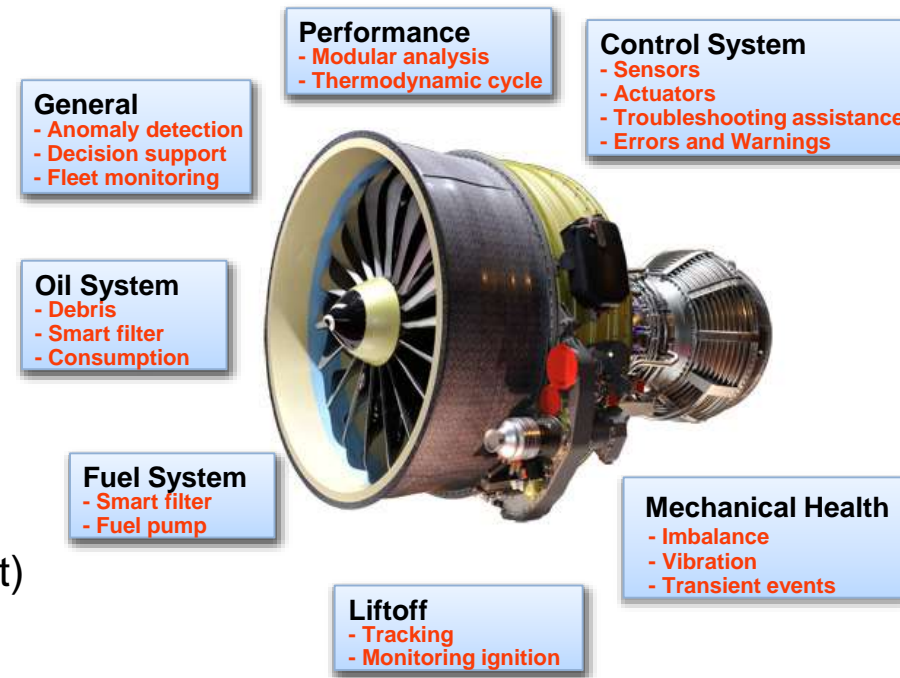
Planning hydrogen fuel-station locations

Optimized engine control systems based on how people drive



# Success Story 2: Safran **Online** Engine Health Monitoring Solution

- Monitor Systems
  - Detect failure indicators
  - Predict time to maintenance
  - Identify components
- Improve Aircraft Availability
  - On time departures and arrivals
  - Plan and optimize maintenance
  - Reduce engine out-of-service time
- Reduce Maintenance Costs
  - Troubleshooting assistance (isolate faulty element)
  - Limit secondary damage



# Success Story 3: iSonea Cloud and Embedded Analytics

## Challenge

- Develop an acoustic respiratory monitoring system for wheeze detection and asthma management

## Solution - Analytics in cloud and embedded

- Captures 30 seconds of windpipe sound and processes the data locally to clean up and reduce ambient noise
- Invokes spectral processing and pattern-detection analytics for wheeze detection on iSonea server in the cloud
- Provides feedback to the patient on their smartphone

## Results

- Eliminates error-prone self-reporting and visits to the doctor



# Aeronautics

## Automotive



## Retail



## Finance



## Internet



## Logistics



## Healthcare Management



## Medical Devices



## Clean Energy



## Oil & Gas



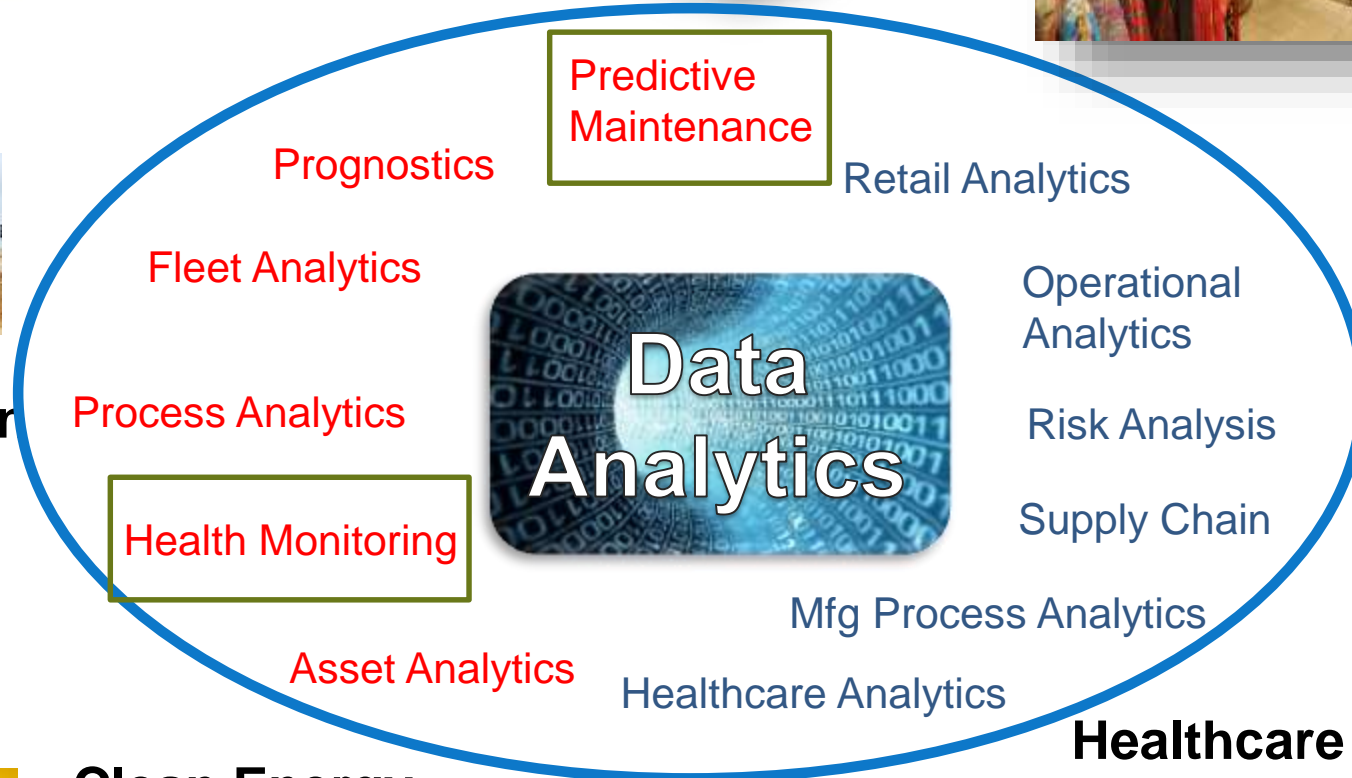
## Industrial Automation



## Off-highway vehicles



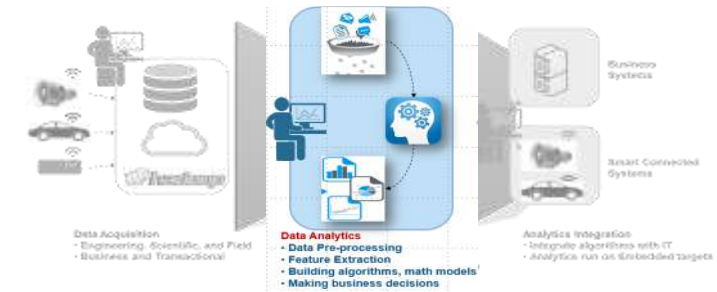
## Railway Systems



# Example 1: Predictive Maintenance of Turbofan Engine

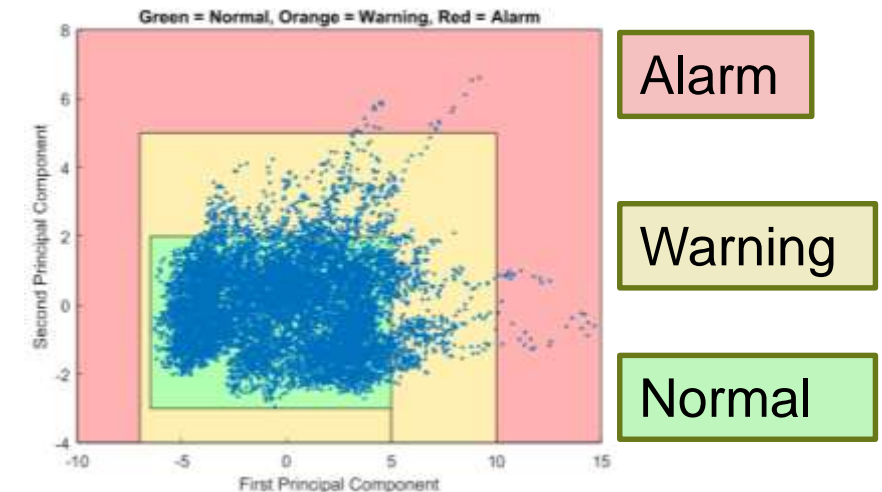
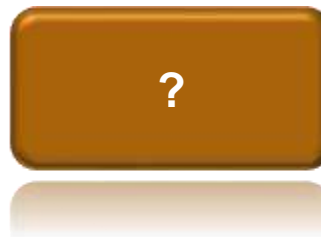
## Background:

- Sensor data from 100 engines of the same model
- The manufacturer recommends that we perform maintenance after every 125 flights



## Questions:

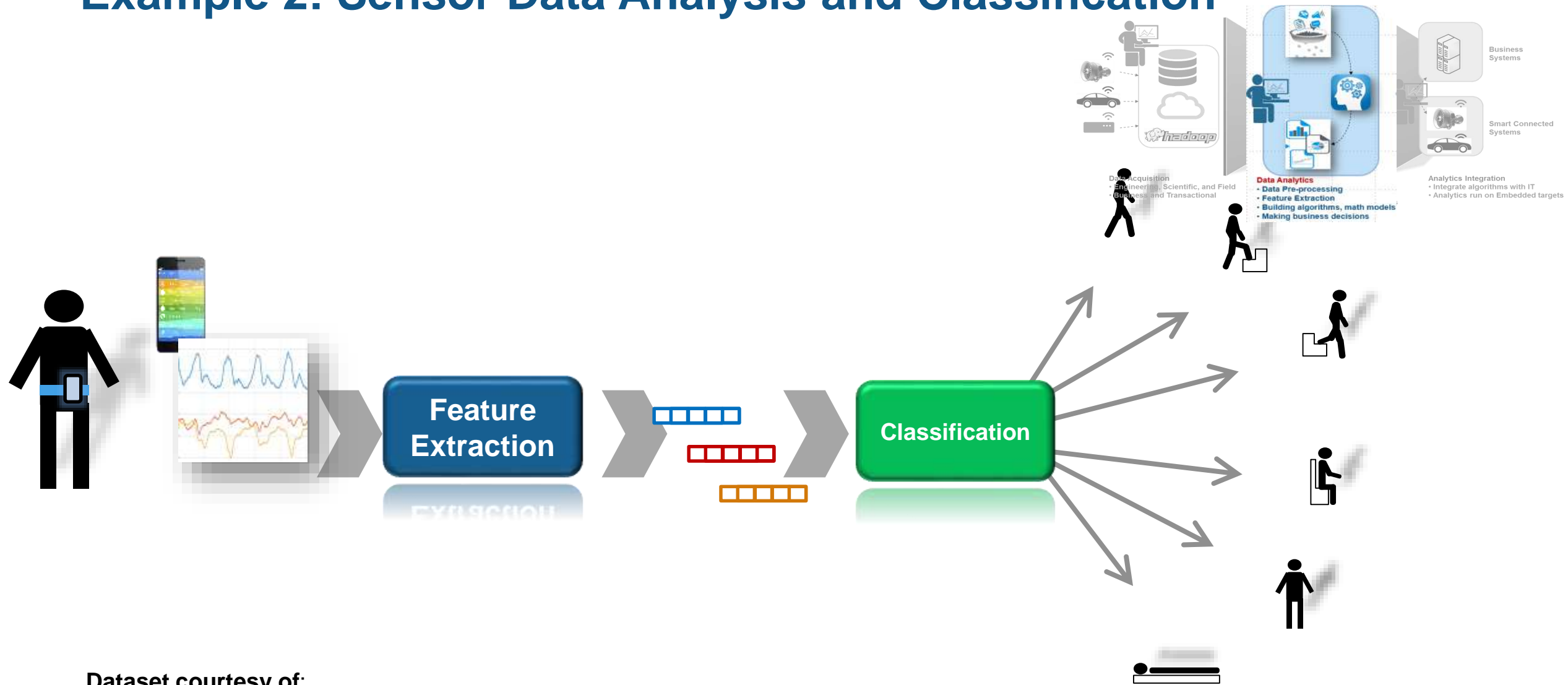
- Are we wasting money by doing maintenance more often than needed?
- Is there a better way to identify when servicing is needed so we can be smarter about scheduling our maintenance.



- Data provided by NASA PCoE
- <http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/>



# Example 2: Sensor Data Analysis and Classification



## Dataset courtesy of:

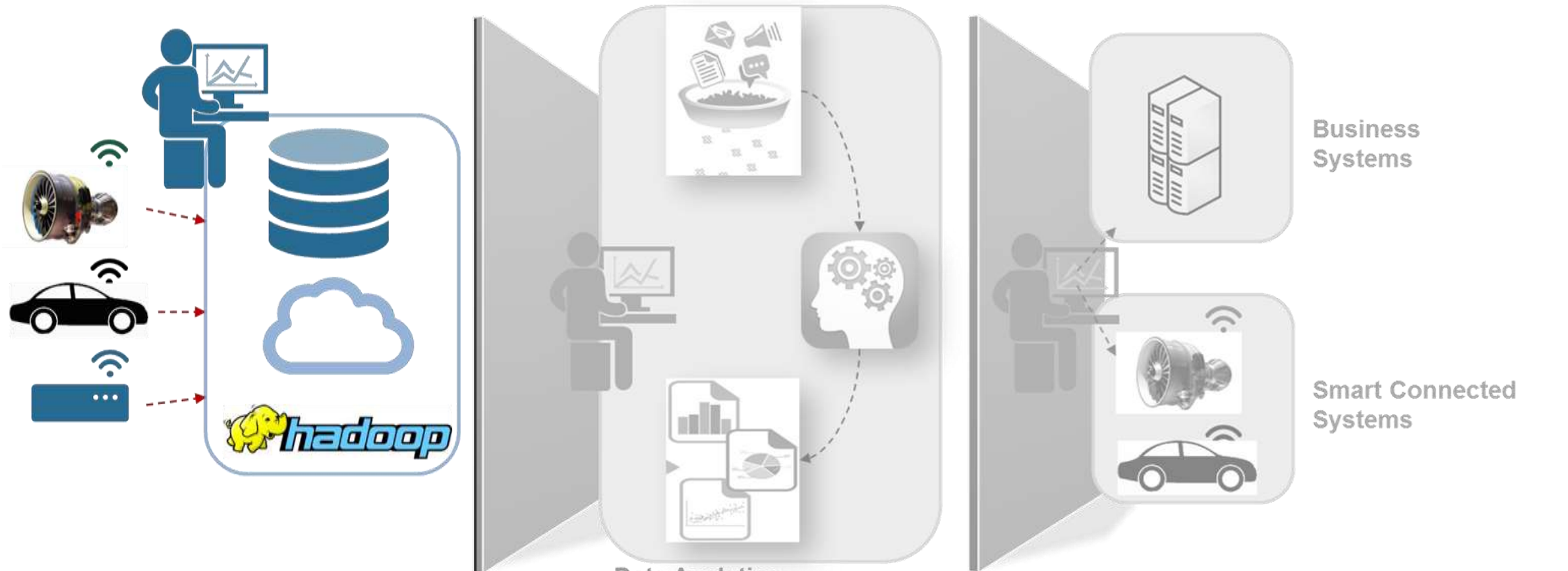
Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz.

*Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine.*

International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012

<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

# Data Analytics Workflow: Data Acquisition



## Data Acquisition

- Engineering, Scientific, and Field
- Business and Transactional

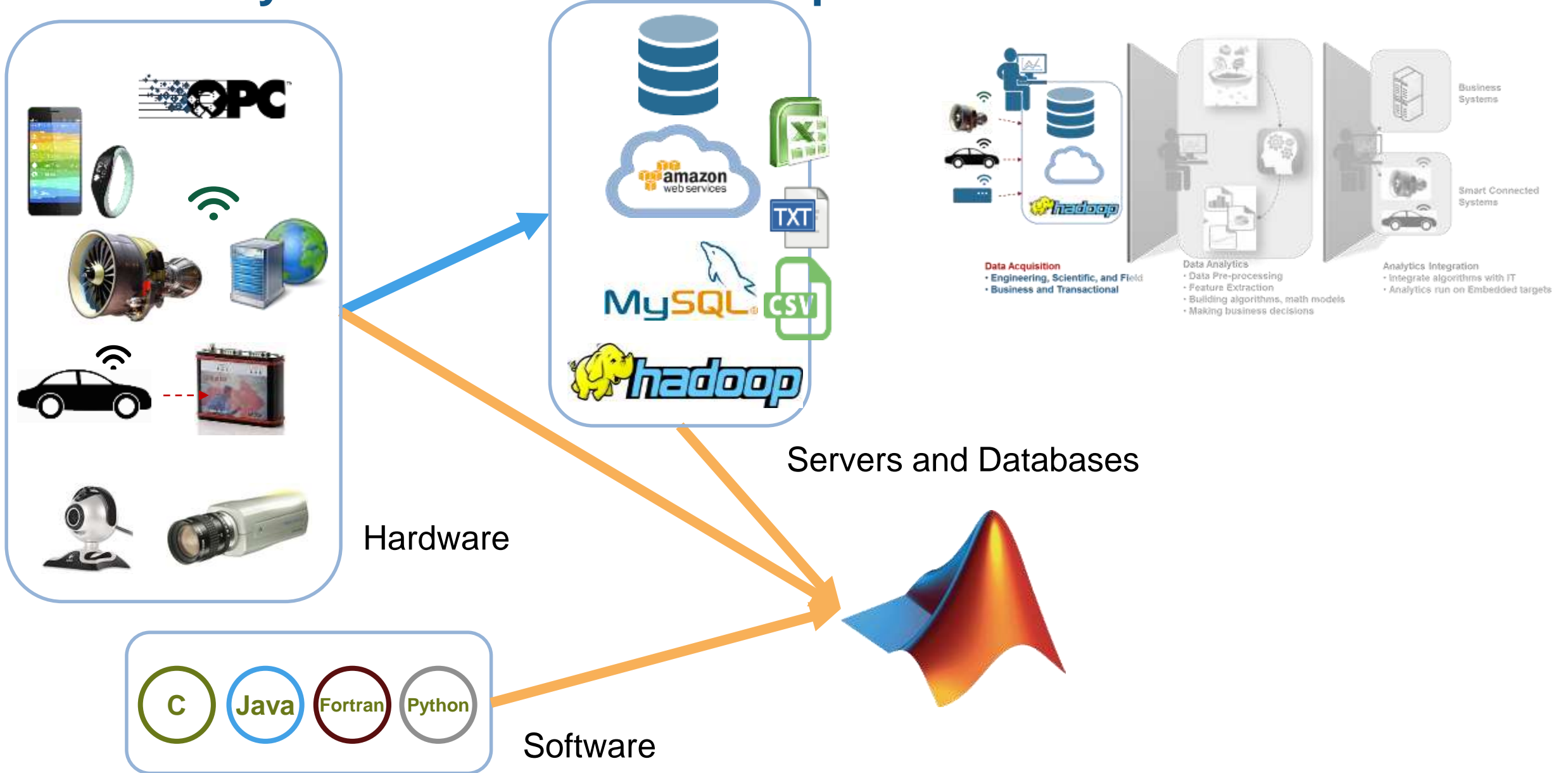
## Data Analytics

- Data Pre-processing
- Feature Extraction
- Building algorithms, math models
- Making business decisions

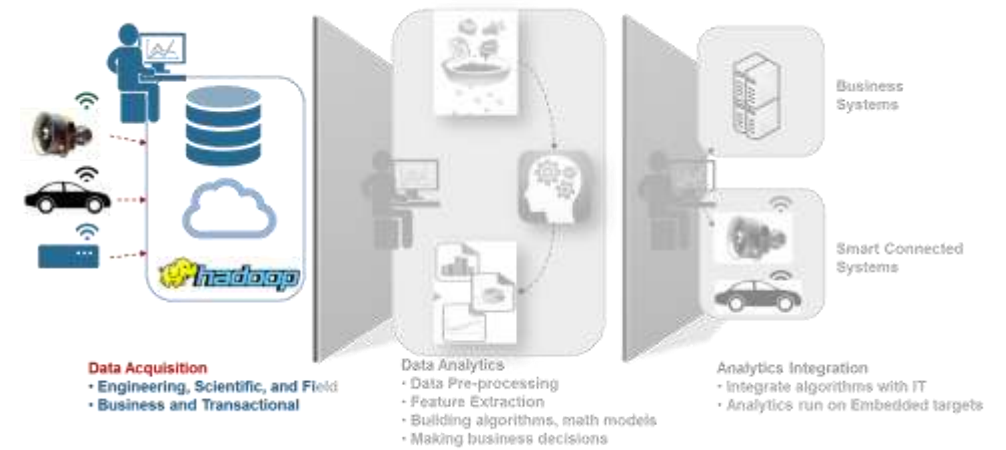
## Analytics Integration

- Integrate algorithms with IT
- Analytics run on Embedded targets

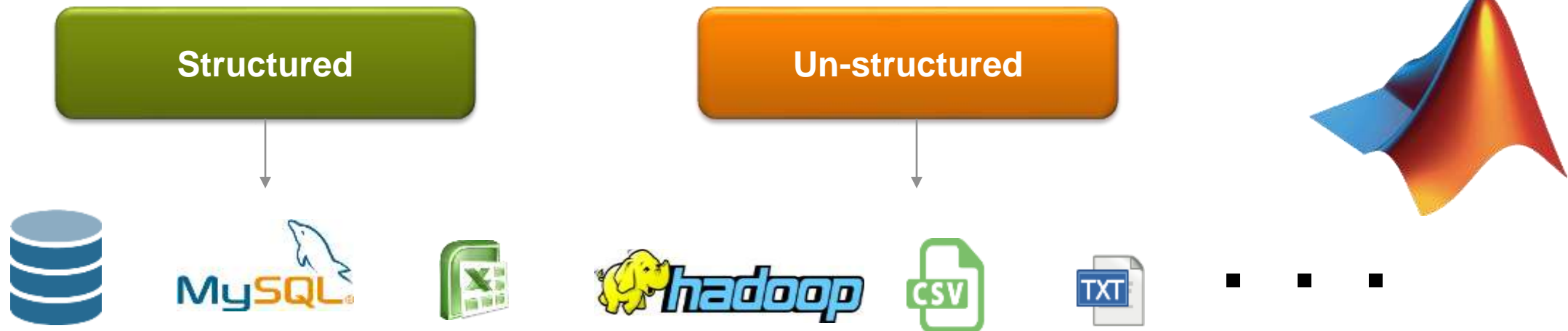
# Data Analytics Workflow: Data Acquisition



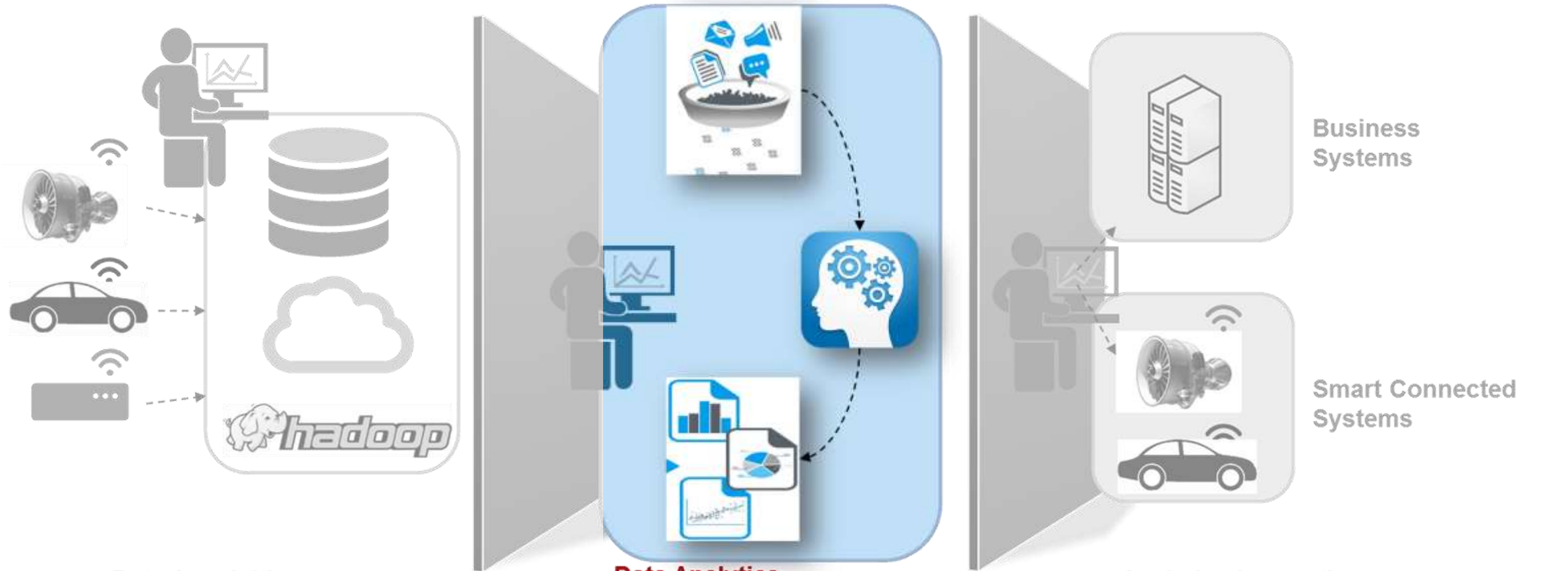
# Data Analytics Workflow: Data Acquisition



## Servers and Databases



# Data Analytics Workflow: **Data Analytics**



## Data Acquisition

- Engineering, Scientific, and Field
- Business and Transactional

## Data Analytics

- Data Pre-processing
- Feature Extraction
- Building algorithms, math models
- Making business decisions

## Analytics Integration

- Integrate algorithms with IT
- Analytics run on Embedded targets

# Example 1: Predictive Maintenance of Turbofan Engine

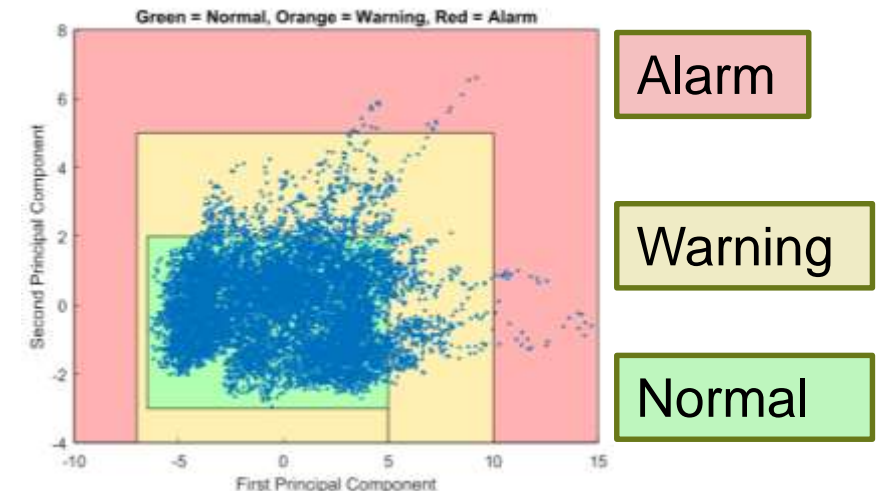
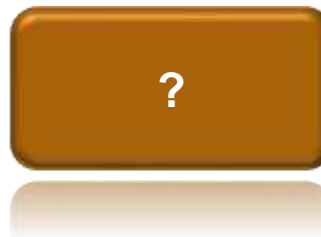
## Background:

- Sensor data from 100 engines of the same model
- The manufacturer recommends that we perform maintenance after every 125 flights



## Questions:

- Are we wasting money by doing maintenance more often than needed?
- Is there a better way to identify when servicing is needed so we can be smarter about scheduling our maintenance.



- Data provided by NASA PCoE
- <http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/>

# Why perform predictive maintenance?

- Example: faulty braking system leads to windmill disaster
  - <https://youtu.be/-YJuFvjtM0s?t=39s>
- What could have caused this?
  - No scheduled maintenance OR
  - Edge case scenarios might not taken into account OR
  - Anything else
- Things under control:
  - Carry on maintenance

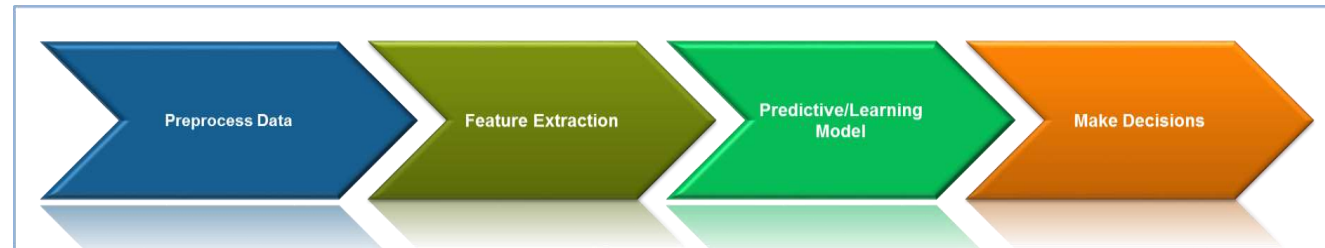
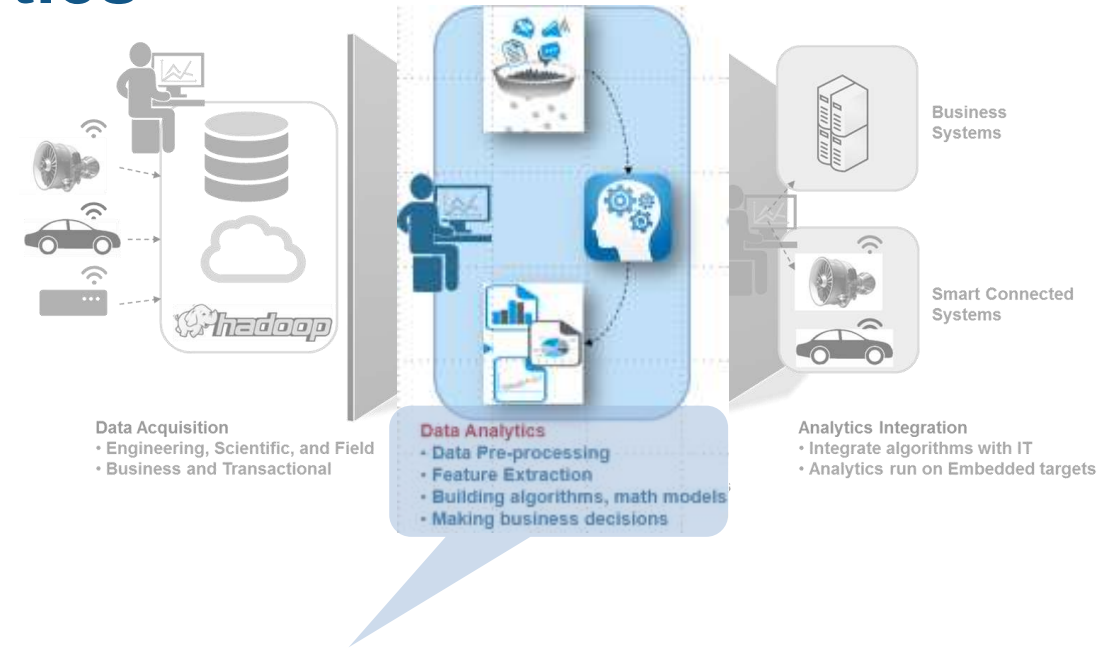
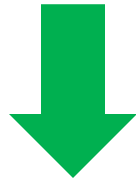


# Types of Maintenance

- Reactive – Do maintenance once there's a problem
  - Example: replace car battery when it has a problem
  - **Problem:** unexpected failures can be expensive and potentially dangerous
  
- Scheduled – Do maintenance at a regular rate
  - Example: change car's oil every 5,000 miles
  - **Problem:** unnecessary maintenance can be wasteful; may not eliminate all failures
  
- Predictive – Forecast when problems will arise
  - Example: certain GM car models forecast problems with the battery, fuel pump, and starter motor
  - Problem: difficult to make accurate forecasts for complex equipment



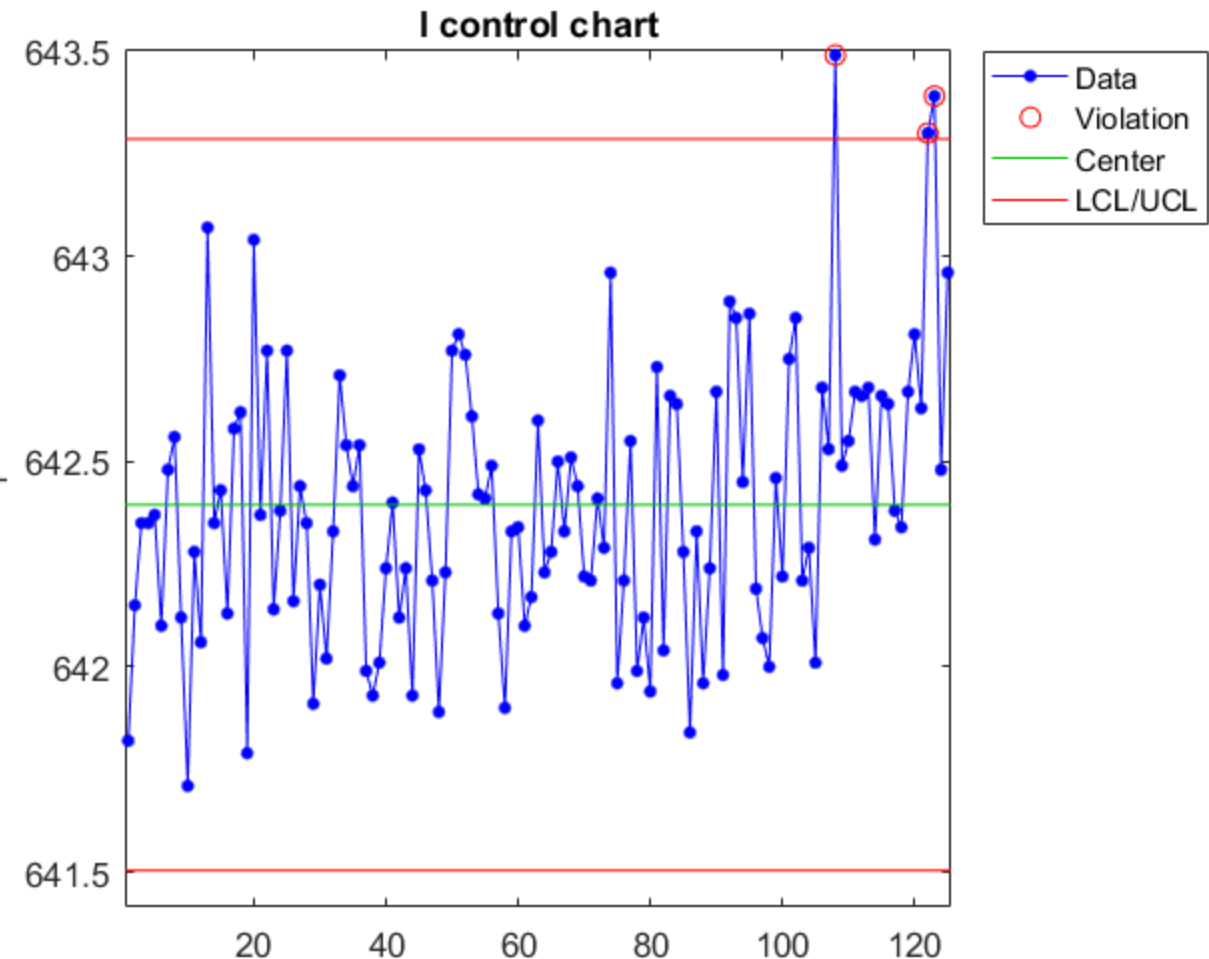
# Data Analytics Workflow: Data Analytics



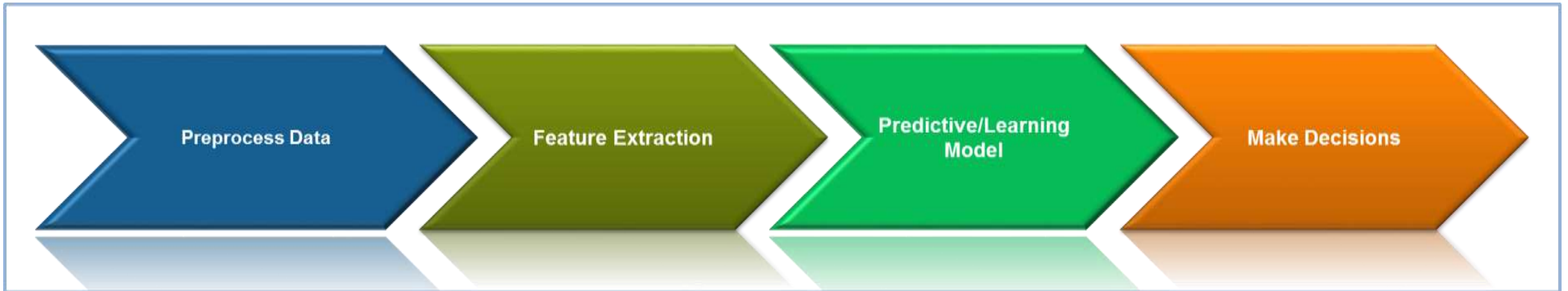
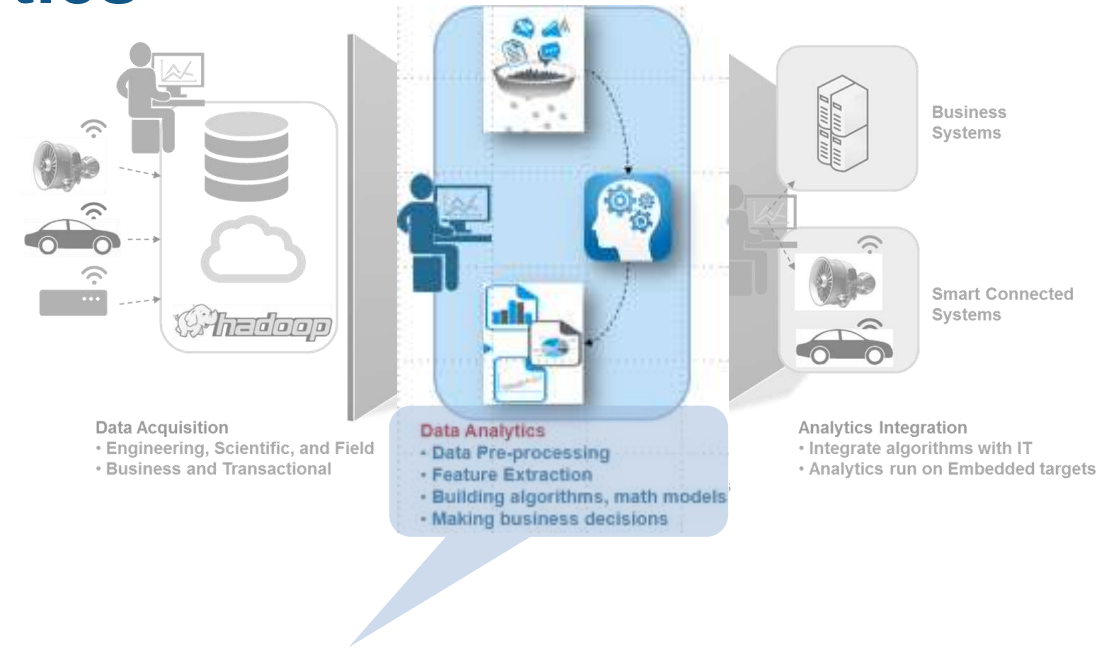
# Monitoring Equipment Health

- We have ***clean data***. How can we use these signals to determine if the equipment is in normal conditions?
  - Control Charts
- Challenge:
  - Number of signals -14
  - Difficult to say when do we have a problem
    - Is 1 sensor going outside the bounds for 1 point a problem?
    - 5 sensors for 3 points?
    - 10 sensors for 20 points?
  - Control charts become difficult to use in these cases, so we will bring in **dimension reduction techniques** to help us.

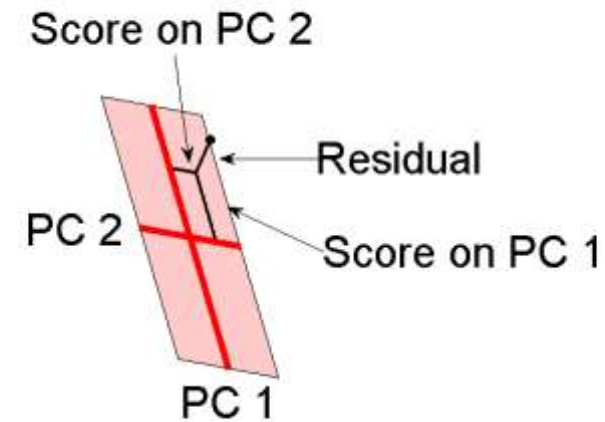
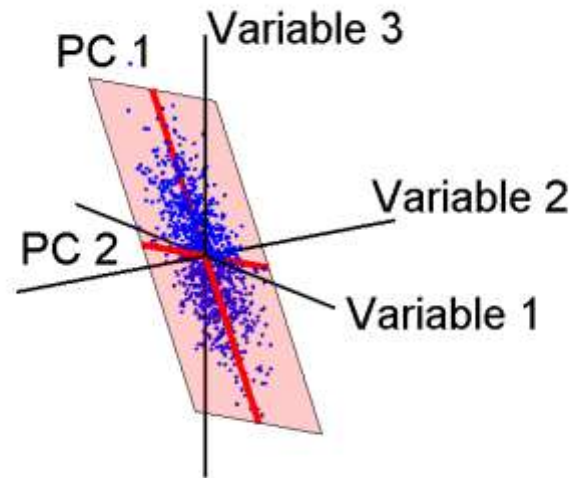
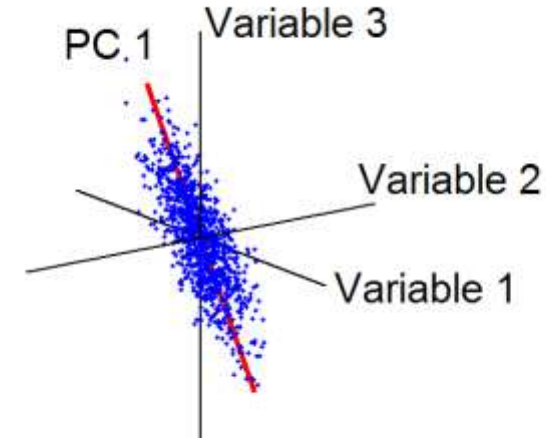
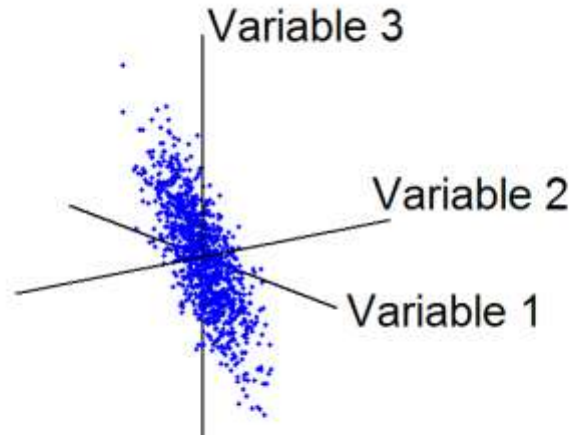
```
>> controlchart(sensorData.LPCOutletTemp,'chart','i')
```



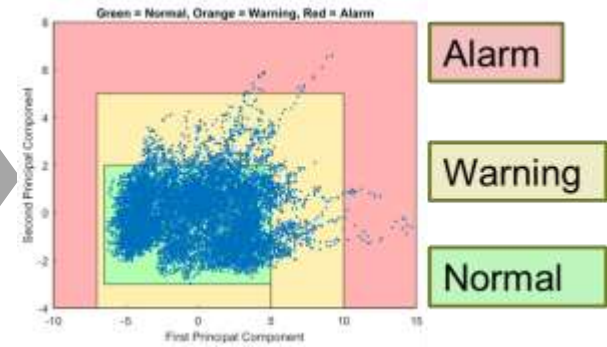
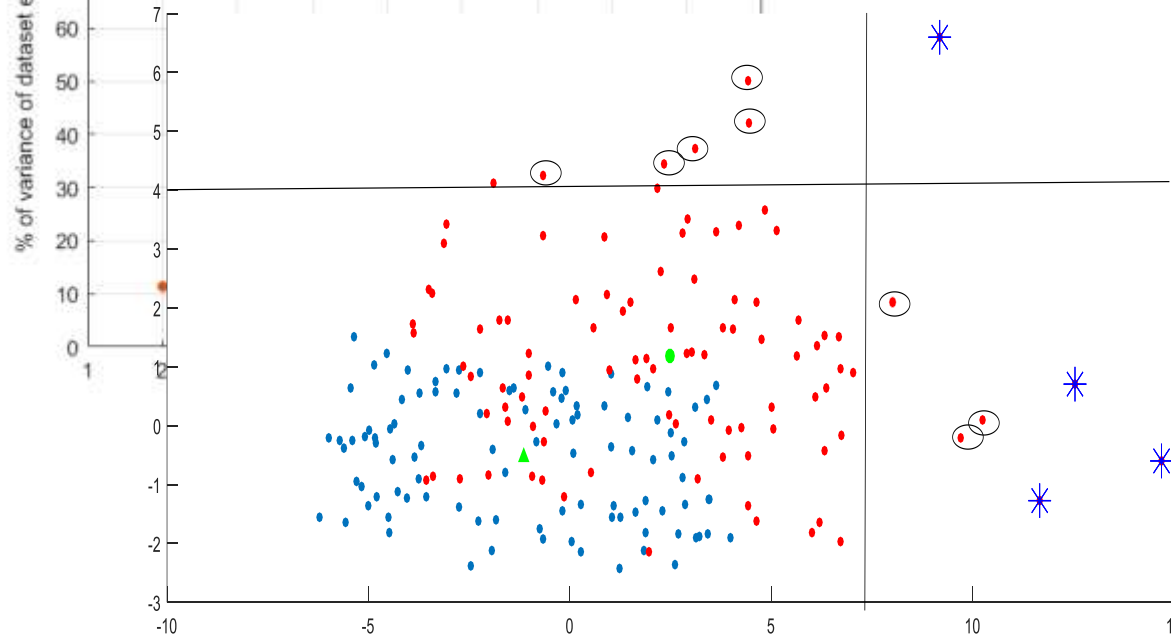
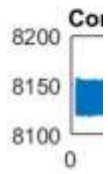
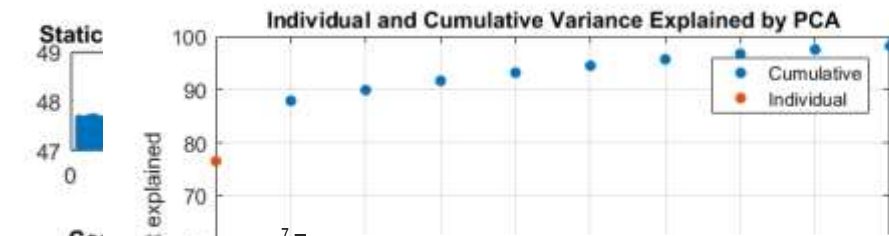
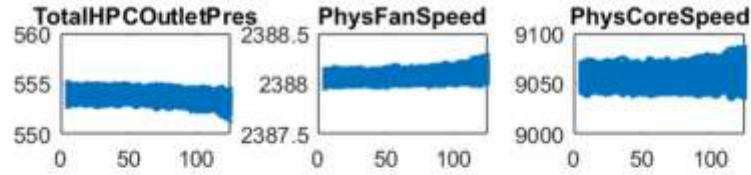
# Data Analytics Workflow: Data Analytics



# Principal Components Analysis – what is it doing?



# Summary: Data Analytics for Predictive Maintenance of Turbofan Engine

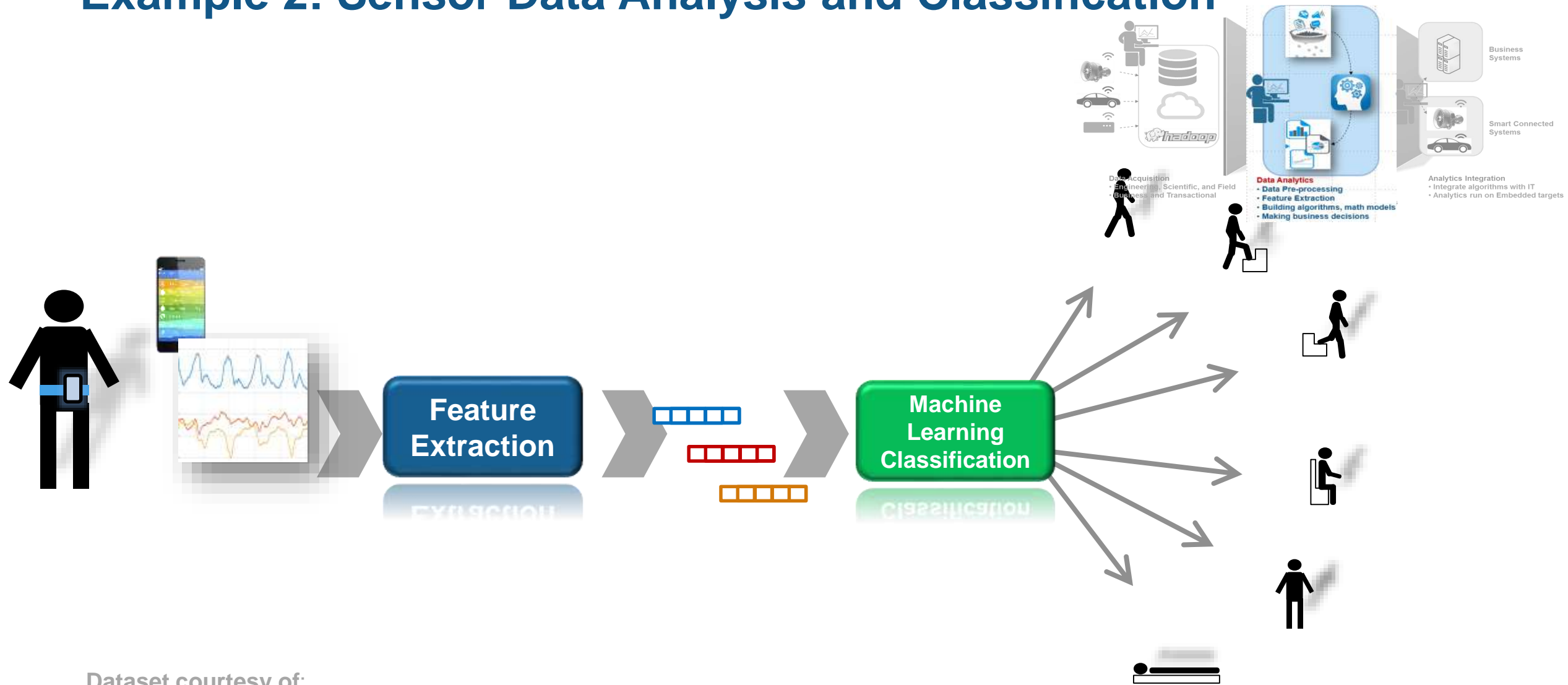


*“Yes, these engines indeed needed maintenance”*

**Maintenance engineer**



# Example 2: Sensor Data Analysis and Classification



Dataset courtesy of:

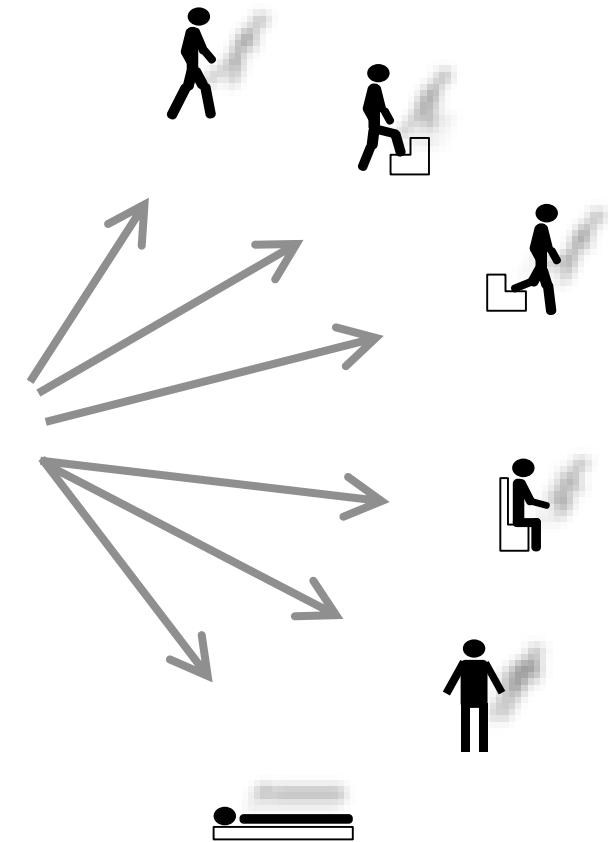
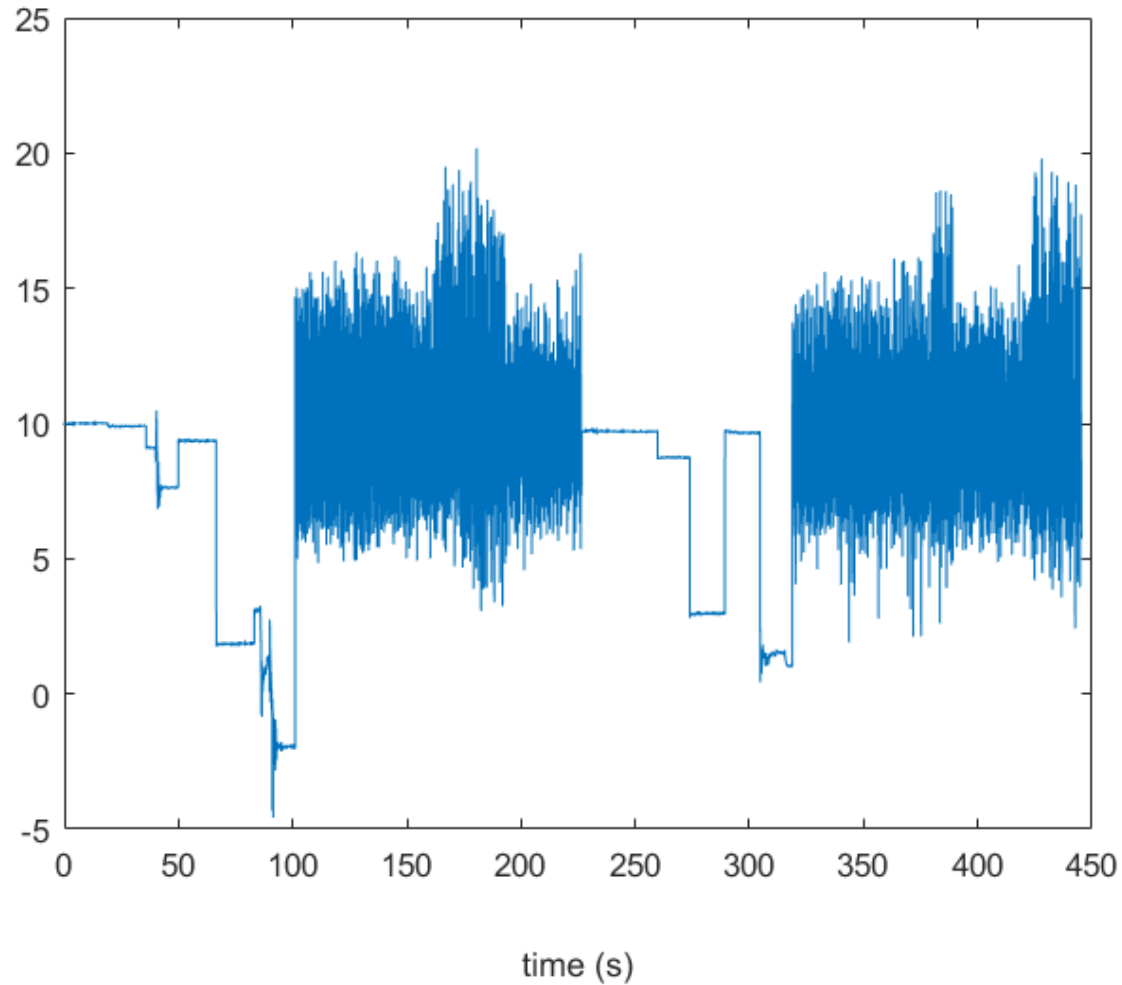
Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz.

*Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine.*

International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012

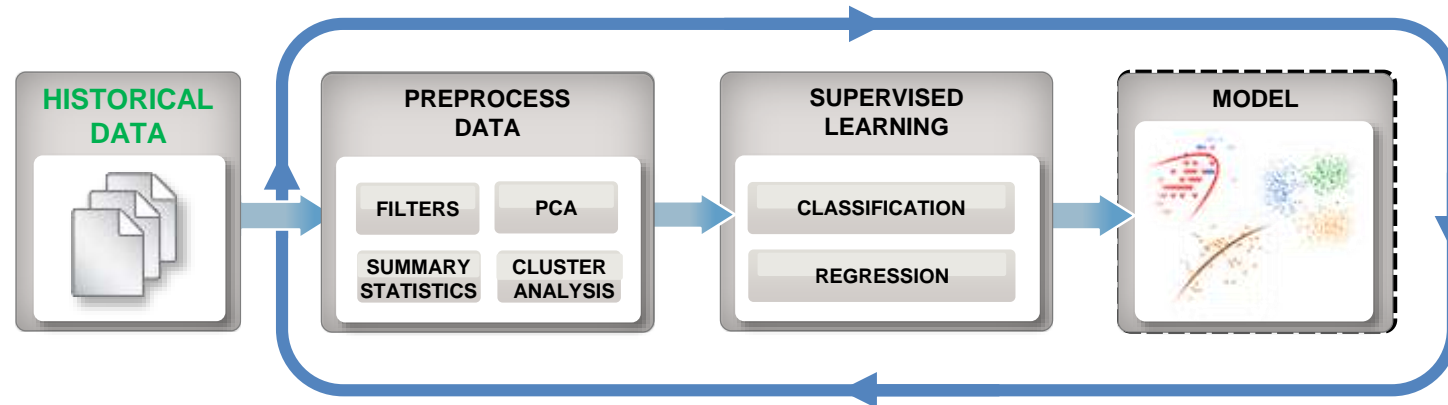
<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

# Why Machine Learning?

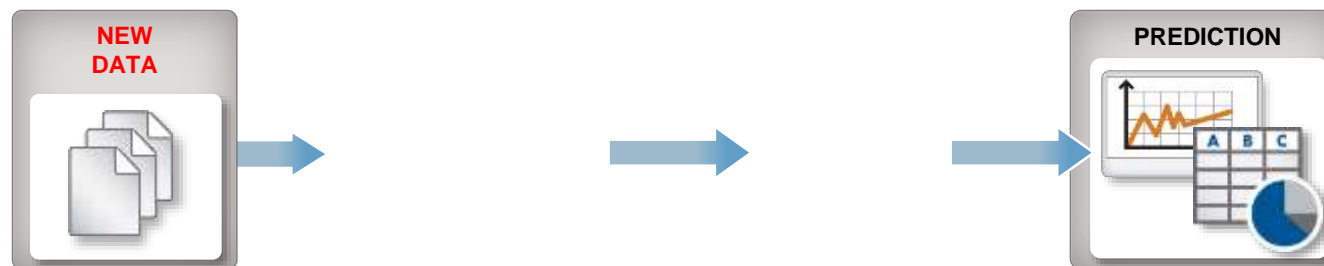


# Machine Learning Workflow

**Train:** Iterate till you find the best model using historical data

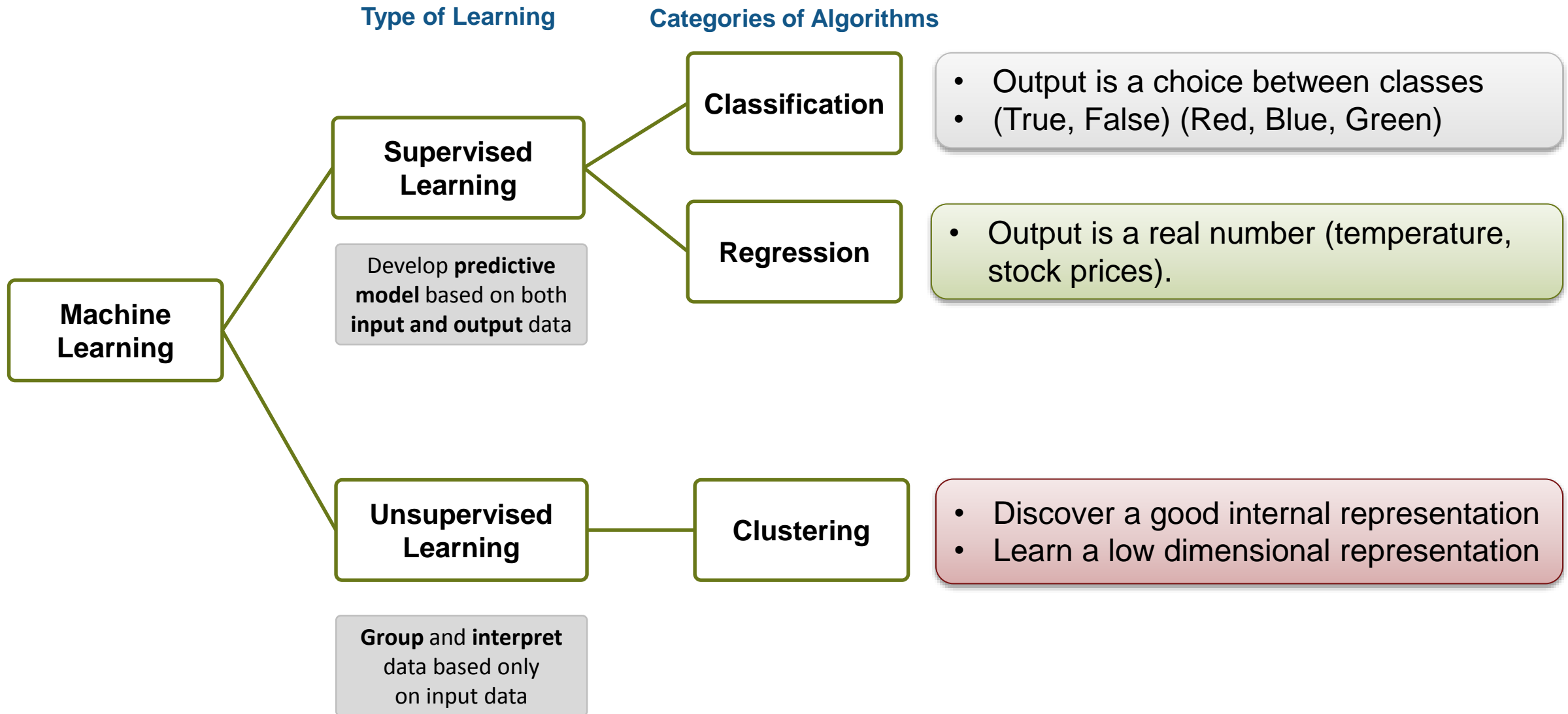


**Predict:** Integrate trained models into applications

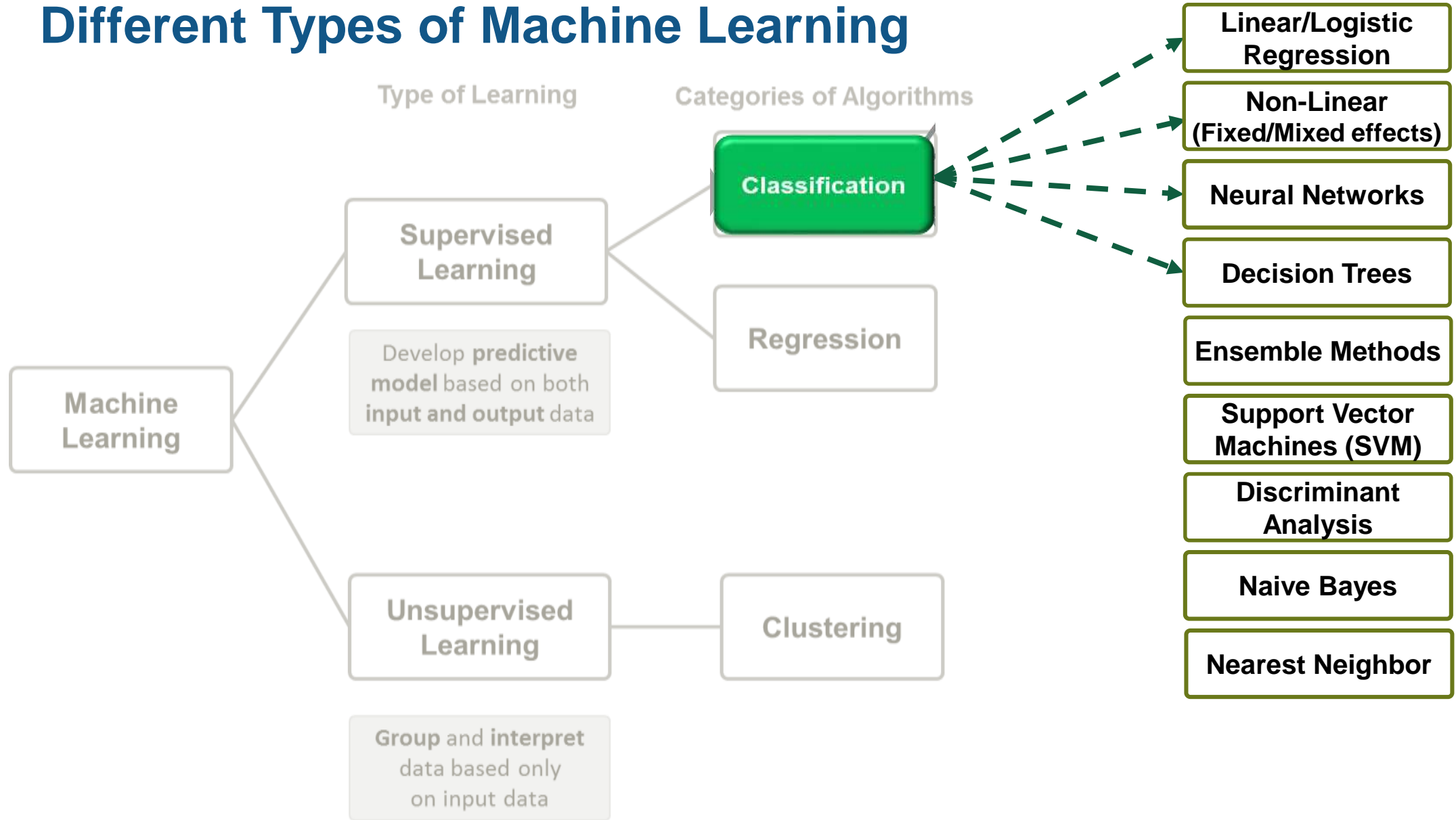




# Different Types of Machine Learning



# Different Types of Machine Learning



# Summary: Machine learning for Sensor Data Classification



- *K-Mean clustering*
- Naïve Bayes
- SVM
- Classification Trees
- KNN
- Neural Networks
- Evaluation metrics

- $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

- ROC

# Learn Further: MATLAB for Machine Learning

## Classification

Build models to classify data into different categories.



**Algorithms:** support vector machine (SVM), boosted and bagged decision trees, k-nearest neighbor, Naïve Bayes, discriminant analysis, neural networks, and more

» Get started with introductory examples

**Applications:** credit scoring, tumor detection, image recognition

## Regression

Build models to predict continuous data.



**Algorithms:** linear model, nonlinear model, regularization, stepwise regression, boosted and bagged decision trees, neural networks, adaptive neuro-fuzzy learning, and more

» Get started with introductory examples

**Applications:** electricity load forecasting, algorithmic trading

## Clustering

Find natural groupings and patterns in data.



**Algorithms:** k-means, hierarchical clustering, Gaussian mixture models, hidden Markov models, self-organizing maps, fuzzy c-means clustering, subtractive clustering, and more

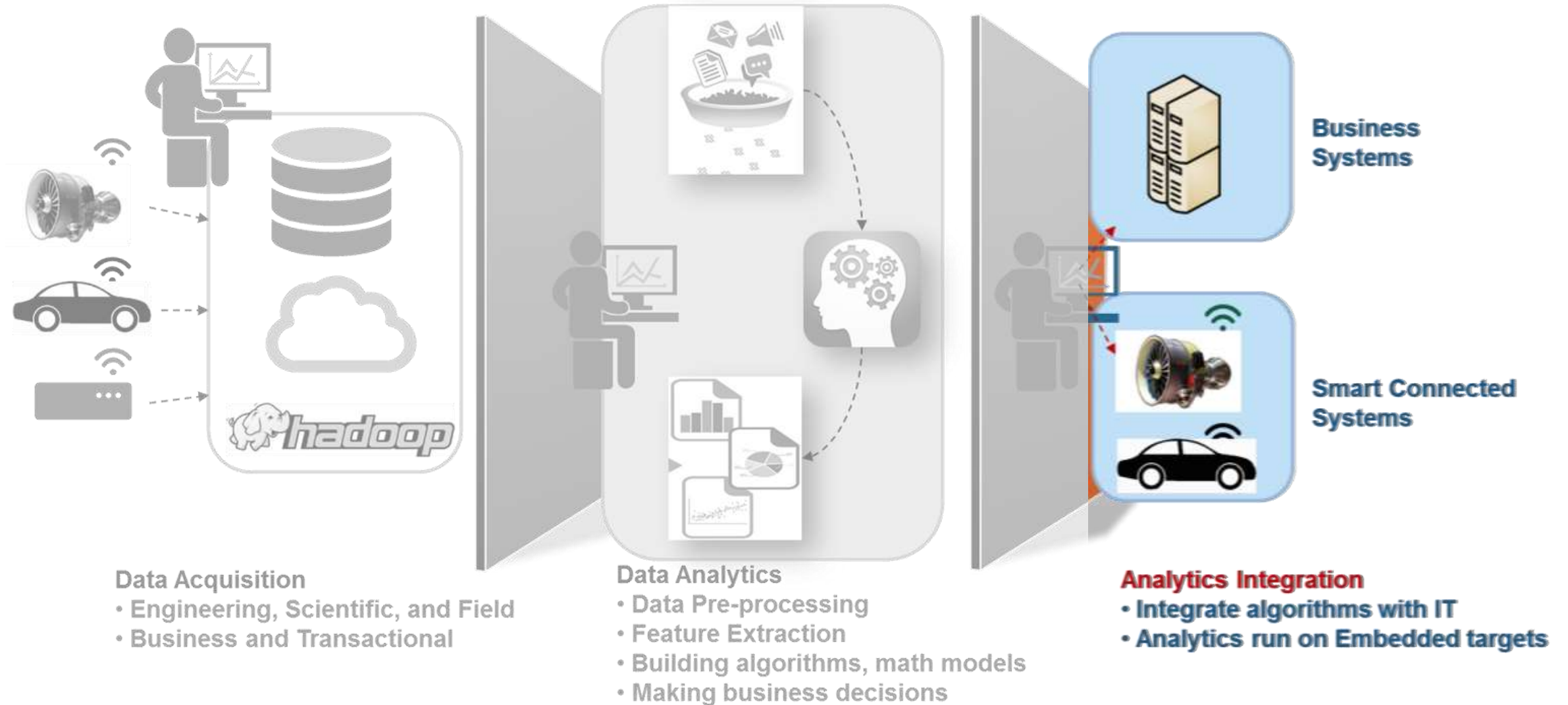
» Get started with introductory examples

**Applications:** pattern mining, medical imaging, object recognition

Go to MATLAB Help → 

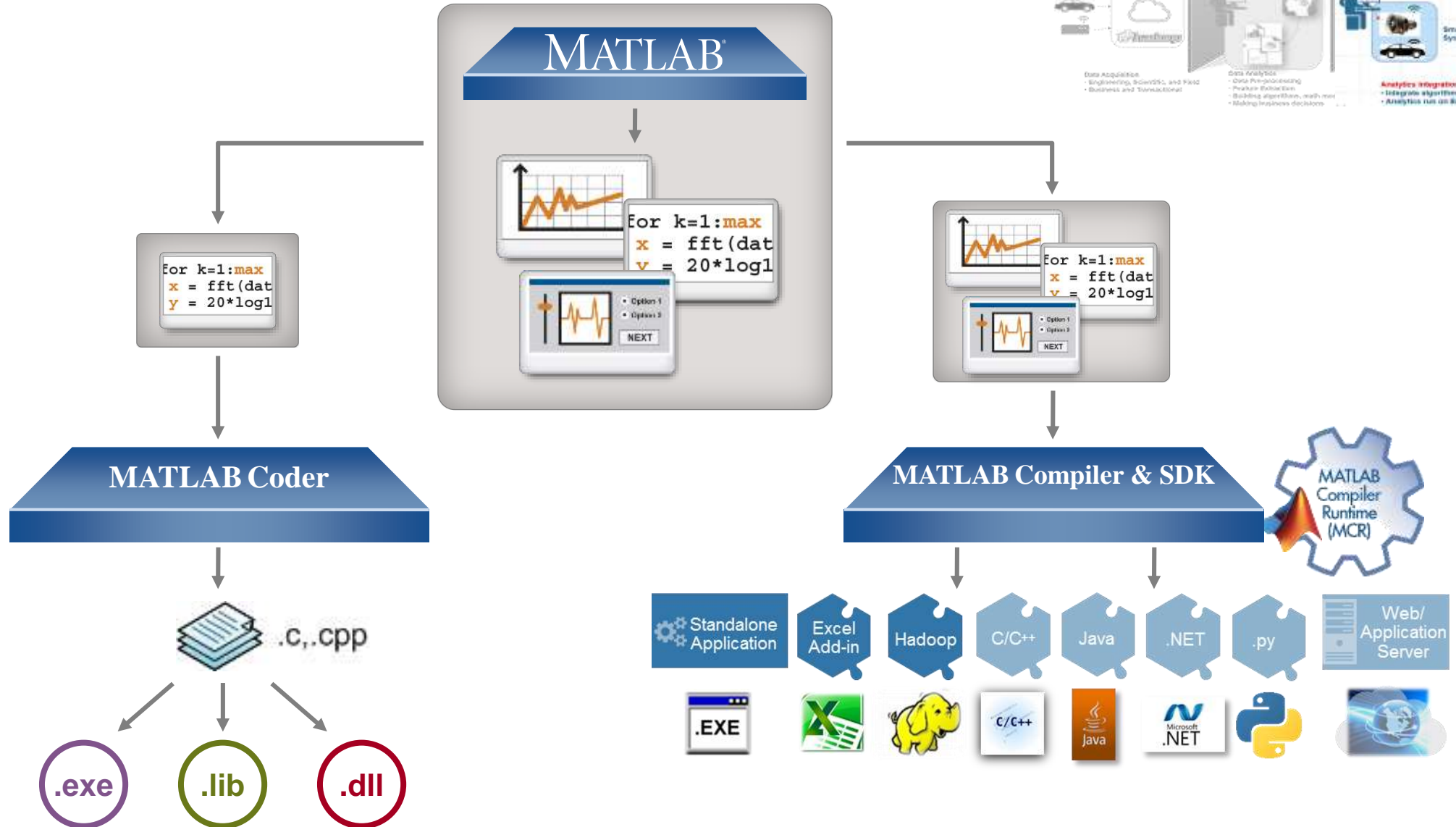
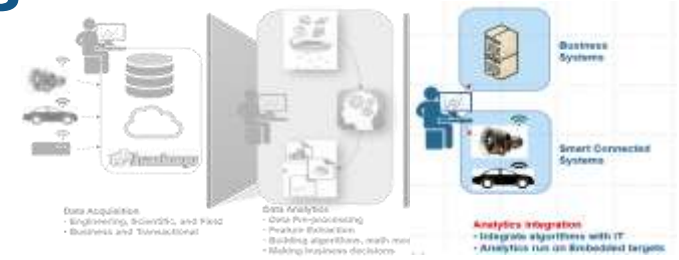
- Functions
- Classes
- Examples and How-To
- Concepts

# Data Analytics Workflow: Analytics Integration

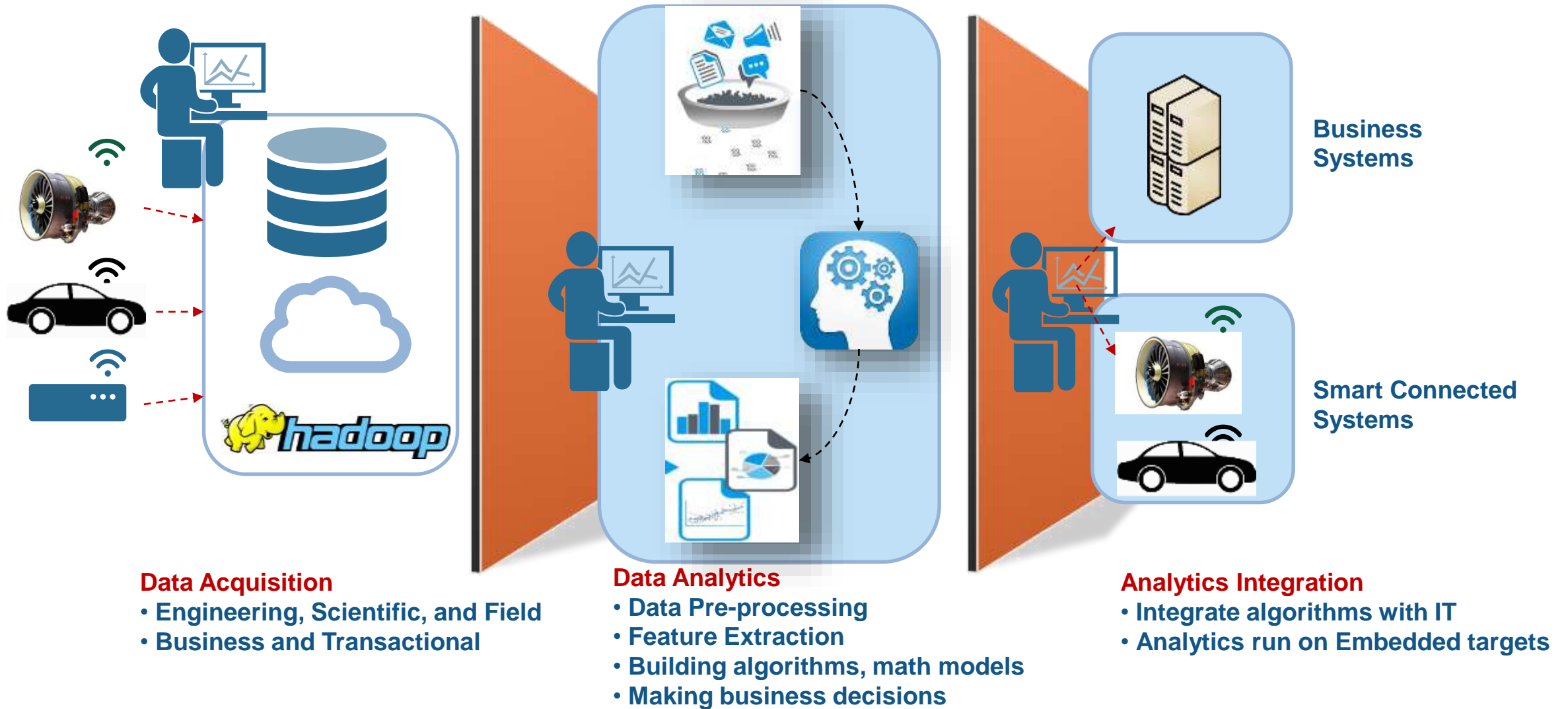


# Integrate analytics with your enterprise systems

*MATLAB Compiler and MATLAB Coder*



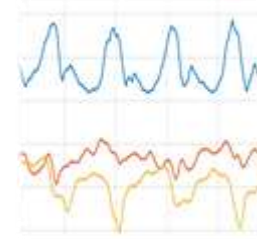
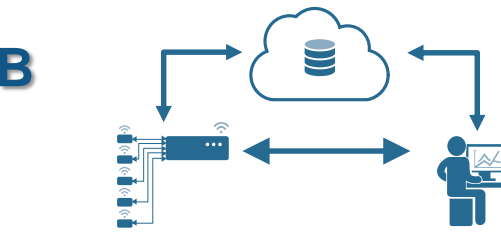
# Summary: Data Analytics Workflow



**MATLAB: Single Platform**

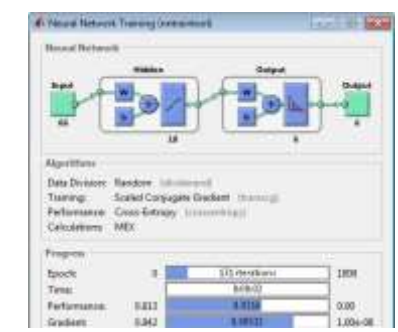
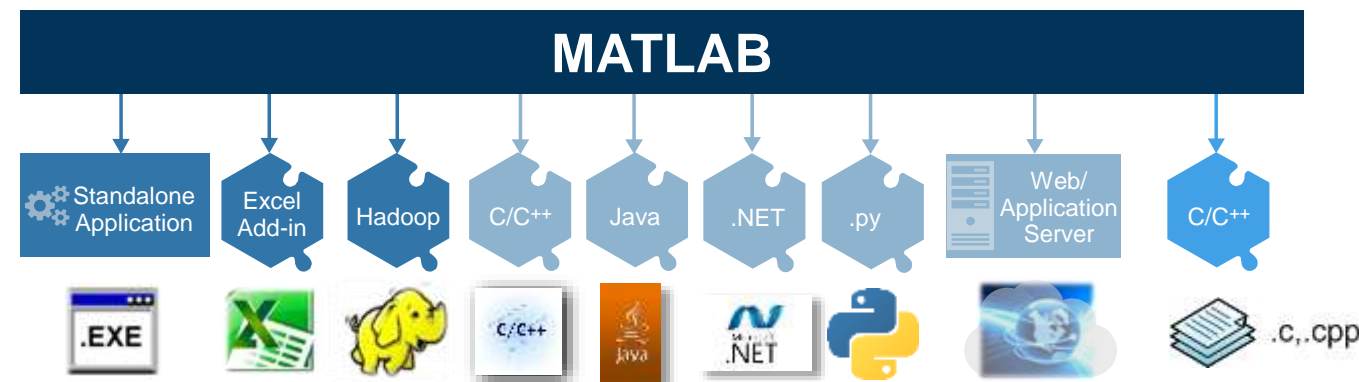
# Key Takeaways: Data Analytics with MATLAB

- No need to be an expert in everything...  
... and if you can still develop & test faster!
- Direct access to sensors/HW (& aggregators)
- Integrated workflow from a single environment
  - Access → Rapid/Iterative Analysis → Deployment
- Leverage parallel computing to scale-up your analytics to large datasets
- Eliminate need to recode by deploying/embedding algorithms into sensors or production



ellip  
filter  
rms  
periodogram  
xcov  
findpeaks  
...

parfor





# MathWorks Services

- Consulting

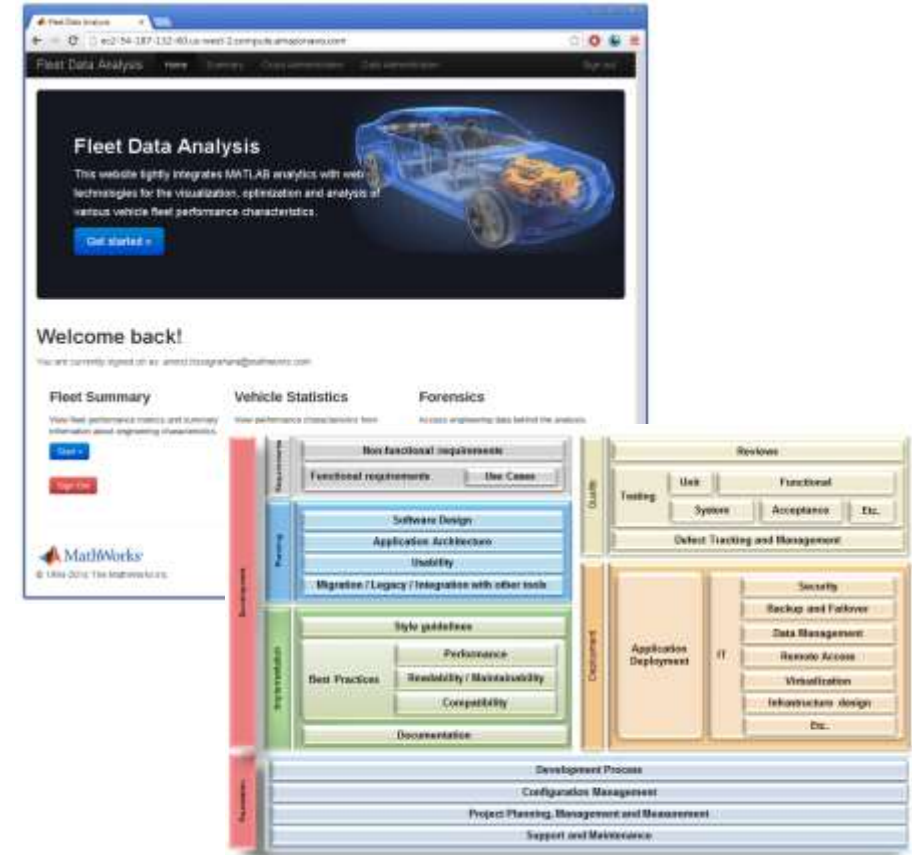
- Integration
- Data analysis/visualization
- Unify workflows, models, data

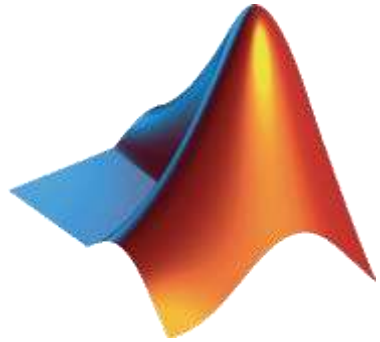
[www.mathworks.com/services/consulting/](http://www.mathworks.com/services/consulting/)

- Training

- Classroom, online, on-site
- Data Processing, Visualization, Deployment, Parallel Computing

[www.mathworks.com/services/training/](http://www.mathworks.com/services/training/)





Questions?