

MATLAB EXPO 2016

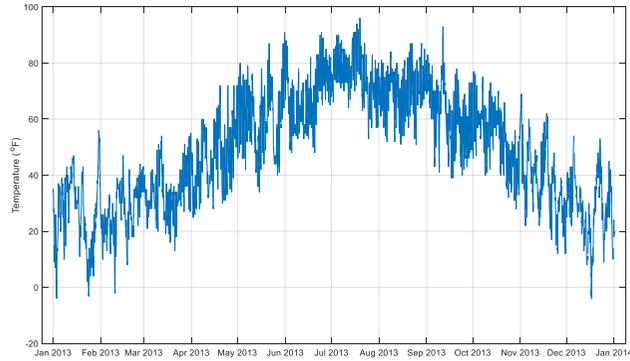
Modelle für die Zukunft dank
prädiktiver Datenanalyse

Jérémy Huard, MathWorks

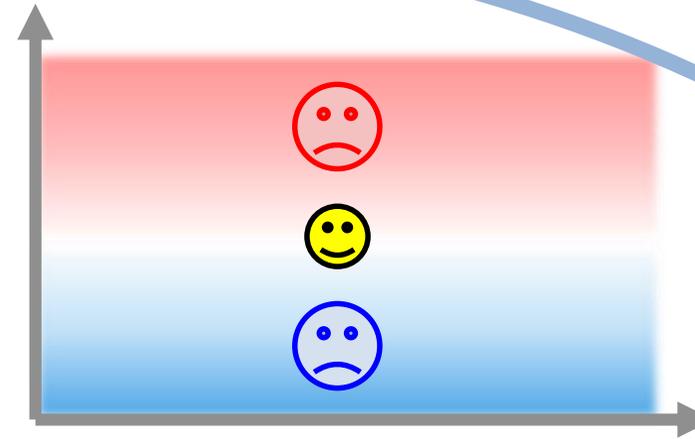


An aerial photograph of a dense city skyline, likely New York City, during the golden hour of sunset. The sky is a mix of orange, yellow, and light blue. The buildings are silhouetted against the bright light, with some windows reflecting the sun. The overall atmosphere is warm and dramatic. In the center of the image, the text "30%" is overlaid in a large, white, sans-serif font with a subtle drop shadow.

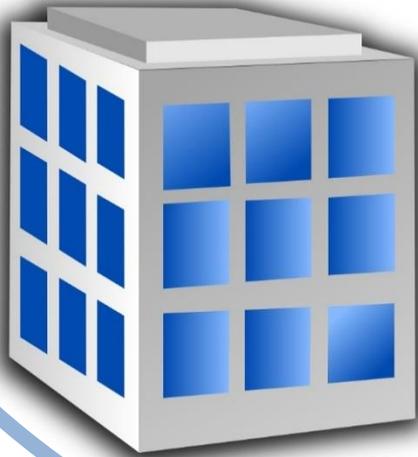
30%



Temperatures change

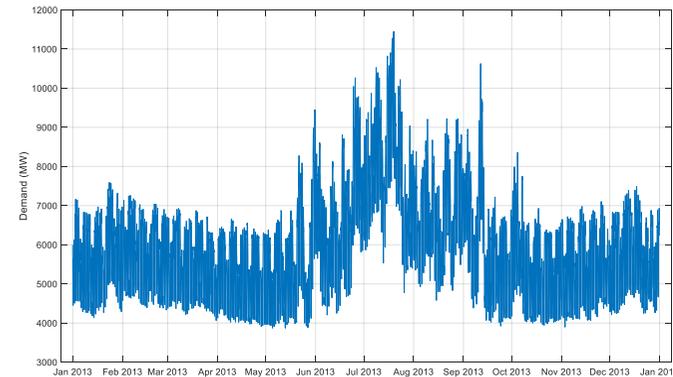


Humans have comfort bounds



$$\frac{\partial u}{\partial t} - \alpha \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = 0$$

Thermodynamic properties



Electricity demand varies

BuildingIQ Develops Proactive Algorithms for HVAC Energy Optimization in Large-Scale Buildings



Large-scale commercial buildings can reduce energy costs by 10–25% with BuildingIQ's energy optimization system.

Office buildings, hospitals, and other large-scale commercial buildings account for about 30% of the energy consumed worldwide. The heating, ventilation, and air-conditioning (HVAC) systems in these buildings are often inefficient because they do not take into account changing weather patterns, variable energy costs, or the building's thermal properties.

BuildingIQ has developed Predictive Energy Optimization™ (PEO), a cloud-based software platform that reduces HVAC energy consumption by 10–25% during normal operation. PEO was developed in cooperation with the Commonwealth Scientific and Industrial Research Organisation (CSIRO),

HVAC pressure sensors, as well as weather and energy cost data. A single building often produces billions of data points, and the scientists and engineers needed tools for efficiently filtering, processing, and visualizing this data.

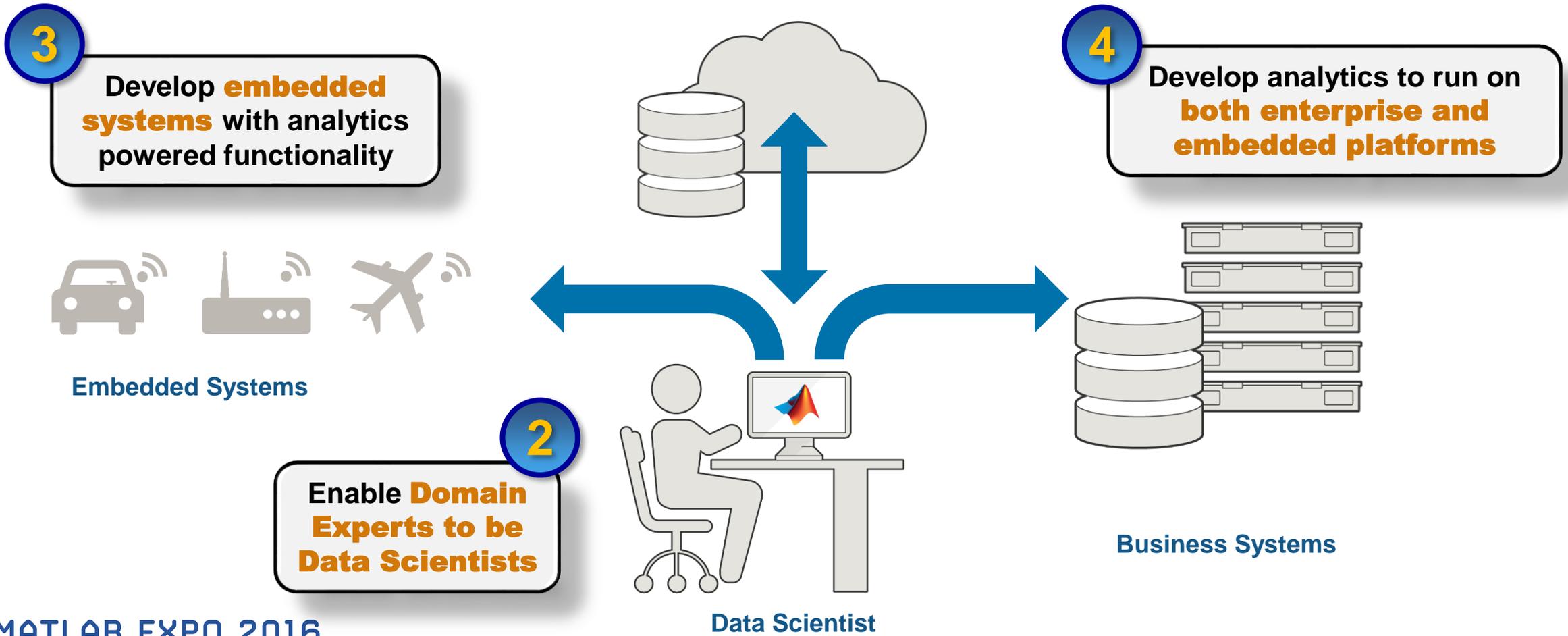
To run their optimization algorithms, the scientists and engineers had to create an accurate mathematical model of a building's thermal and power dynamics. The algorithms would use this calculated model to run constrained optimizations that maintained occupant comfort while minimizing energy costs.

BuildingIQ needed a way to rapidly develop mathematical models, test optimization a

Traits of Data Analytics applications

- 1. Diverse and/or Big Data**
- 2. Advanced Algorithms**
- 3. Deployment**

Why MATLAB?



Why MATLAB?

1 Analytics that increasingly require **both business and engineering data**

- DATA
- Engineering, Scientific, and Field
 - Business and Transactional

3 Develop **embedded systems** with analytics powered functionality



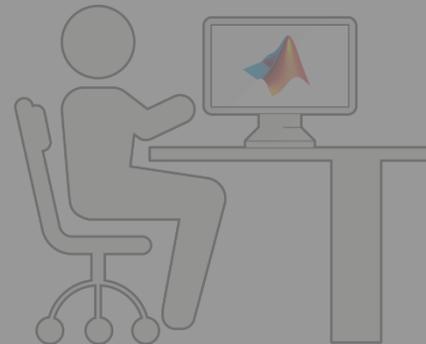
Smarter Embedded Systems

4 Develop analytics to run on **both enterprise and embedded platforms**

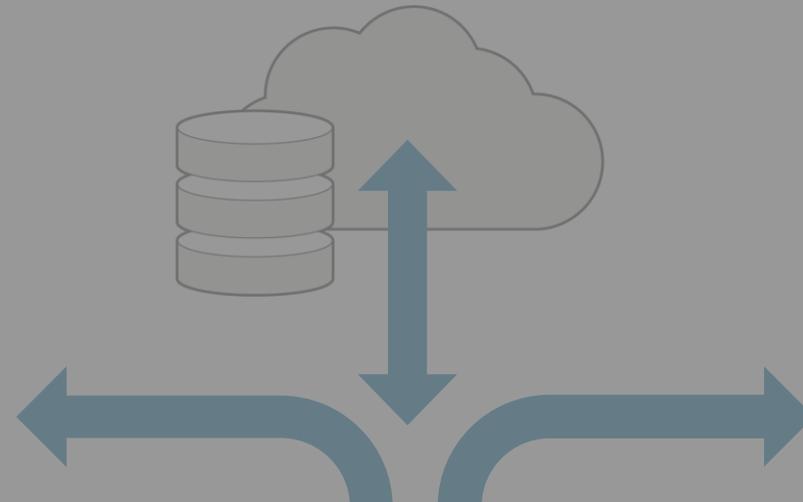


Business Systems

2 Enable **Domain Experts to be Data Scientists**



Data Scientist



Business and Engineering Data

Business and Transactional Data

Repositories

- Databases
- **Hadoop**

File I/O

- Text
- Spreadsheet
- XML

Web Sources

- HTML
- Mapping
- Financial datafeeds
- **RESTful**
- **JSON**

15:15 [Analyse von operationellen Flugdaten aus einem Hadoop System unter Verwendung von MapReduce und dem MATLAB Distributed Computing Server](#)
Lukas Höhndorf, TU München

Engineering, Scientific, and Field Data

File I/O

- Text
- Spreadsheet
- XML
- CDF/HDF
- Image
- Audio
- Video
- Geospatial

Communication Protocols

- CAN (Controller Area Network)
- DDS (Data Distribution Service)
- OPC (OLE for Process Control)
- XCP (eXplicit Control Protocol)

Real-Time Sources

- Sensors
- GPS
- Instrumentation
- Cameras
- Communication systems
- Machines (embedded systems)

*“No matter what industry our client is in, and **no matter what data they ask us to analyze—text, audio, images, or video**—MATLAB enables us to provide clear results faster.”*

Data handling and visualization

The screenshot shows the MATLAB R2016a environment. The Command Window on the left contains the prompt `fx >>`. The Variables window on the right displays a table named `pal` with 3910 rows and 3 columns. The table data is as follows:

	1	2	3	4	5	6
	TimeStamp	Name	Load			
1	01/01/04 00:00:00	'CAPITL'	1015			
2	01/01/04 00:00:00	'CENTRL'	1651			
3	01/01/04 00:00:00	'DUNW...	618			
4	01/01/04 00:00:00	'GENESE'	972			
5	01/01/04 00:00:00	'HUD VL'	1120			
6	01/01/04 00:00:00	'MHK VL'	645			
7	01/01/04 00:00:00	'MILLW...	223			
8	01/01/04 00:00:00	'N.Y.C._...	7267			
9	01/01/04 00:00:00	'NORTH'	622			
10	01/01/04 00:00:00	'WEST'	1591			
11	01/01/04 00:00:47	'CAPITL'	1016			
12	01/01/04 00:00:47	'CENTRL'	1653			
13	01/01/04 00:00:47	'DUNW...	615			
14	01/01/04 00:00:47	'GENESE'	985			
15	01/01/04 00:00:47	'HUD VL'	1106			
16	01/01/04 00:00:47	'MHK VL'	641			

The Workspace window at the bottom shows the variable `pal` as a `3910x3 table` class, occupying 583907 bytes.

High-quality domain-specific libraries

Data type

Common Techniques for Deriving Features

Sensor data

Signal Processing

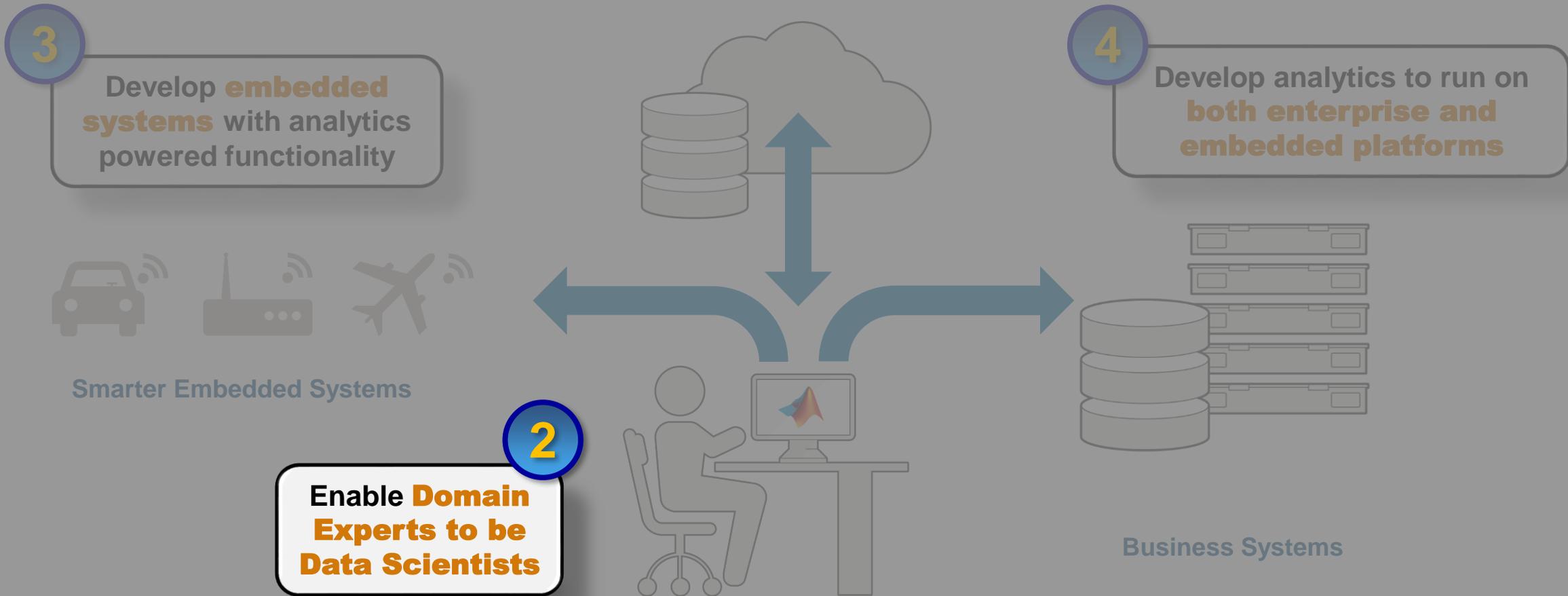
Image and video data

**Image Processing
Computer Vision**

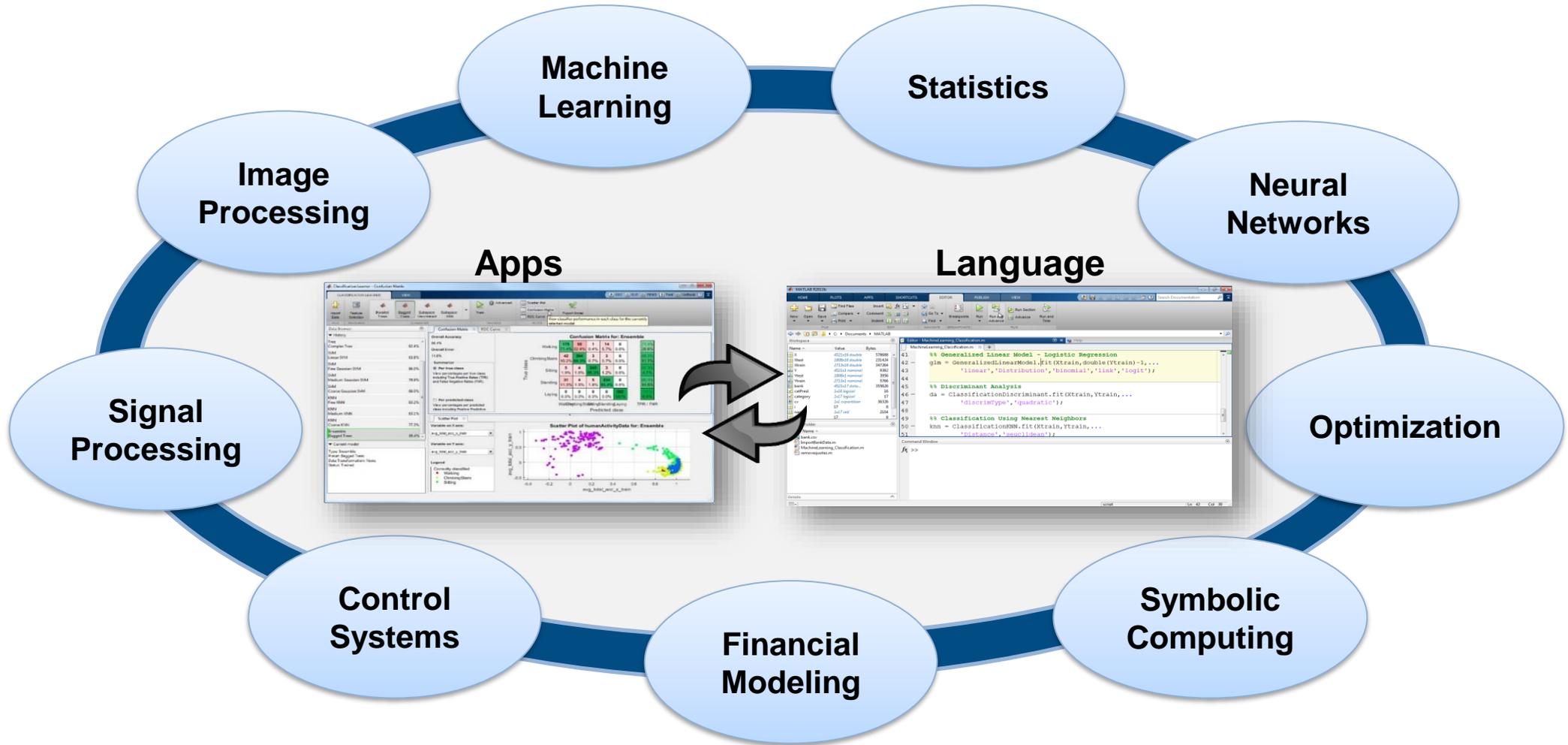
Transactional data

Statistics

Why MATLAB?



Enabling Domain Experts to be Data Scientists



Built-in algorithms

Clustering

Hierarchical Clustering

Produce nested sets of clusters

k-Means and k-Medoids Clustering

Cluster by minimizing mean or medoid distance, calculate Mahalan

Gaussian Mixture Models

Cluster based on Gaussian mixture models using the EM algorithm

Nearest Neighbors

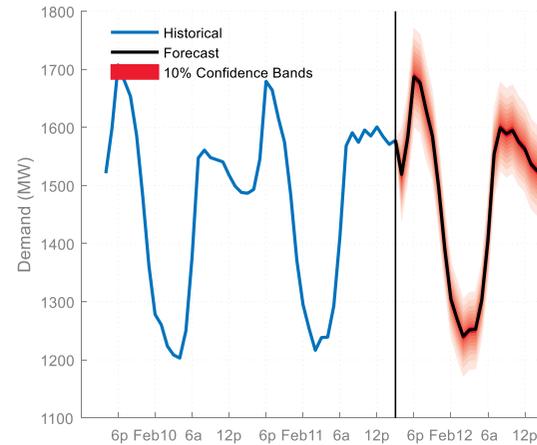
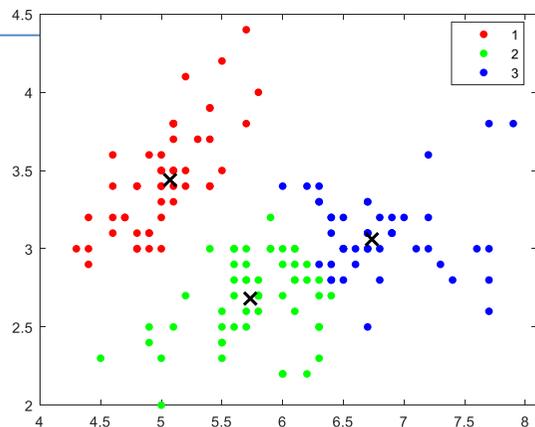
Find nearest neighbors using exhaustive search or *kd*-tree s

Hidden Markov Models

Markov models for data generation

Cluster Visualization and Evaluation

Plot clusters of data and evaluate optimal number of clusters



Linear Regression

Multiple, stepwise, multivariate regression models, and m

Generalized Linear Models

Logistic regression, multinomial regression, Poisson regre

Nonlinear Regression

Nonlinear fixed- and mixed-effects regression models

Support Vector Machine Regression

Support vector machines for regression models

Gaussian Process Regression

Gaussian process regression models (kriging)

Regression Trees

Binary decision trees for regression

Regression Tree Ensembles

Random forests, boosted and bagged regression trees

Regression

Classification

Classification Trees

Binary decision trees for multiclass learning

Discriminant Analysis

Regularized linear and quadratic discriminant analysis

Naive Bayes

Naive Bayes model with Gaussian, multinomial, or kernel predictors

Nearest Neighbors

k nearest neighbors classification using *Kd*-tree search

Support Vector Machine Classification

Support vector machines for binary or multiclass classification

Classification Ensembles

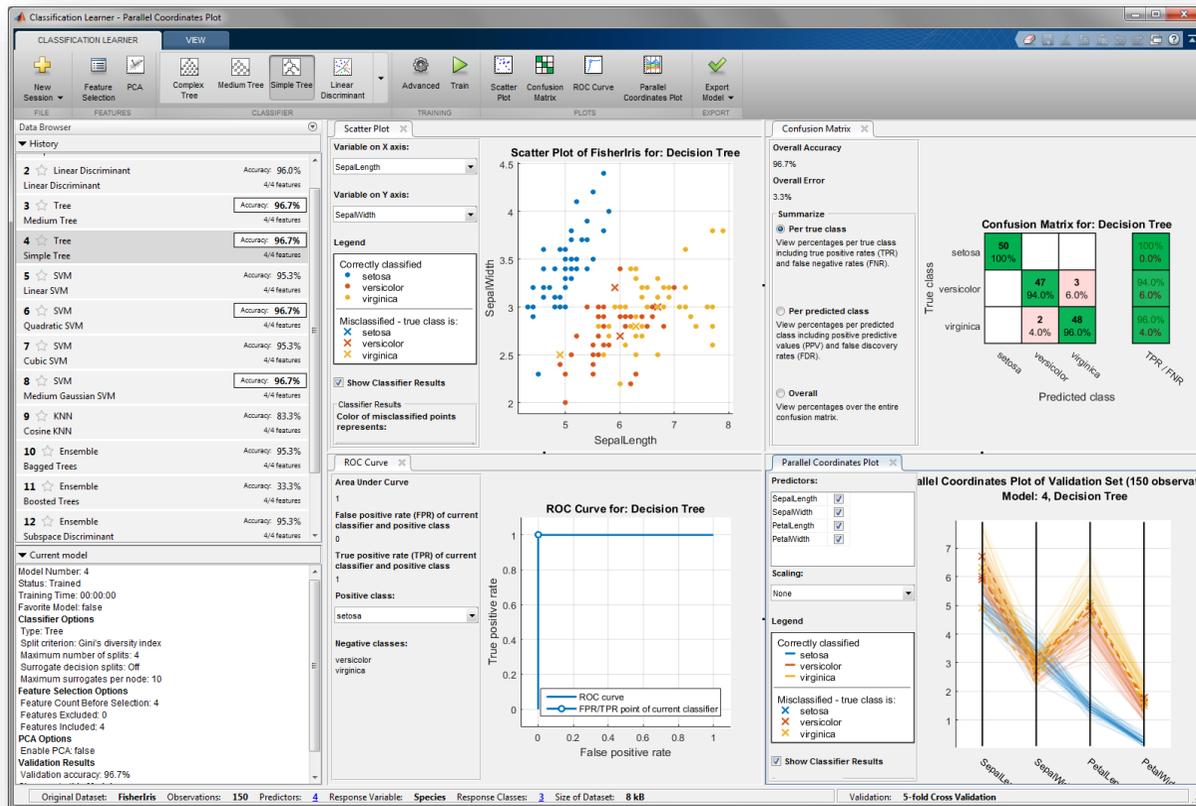
Boosting, random forest, bagging, random subspace, and ECOC ensembles for multiclass learning

Model Building and Assessment

Feature selection, cross validation, predictive performance evaluation, classification accuracy comparison tests

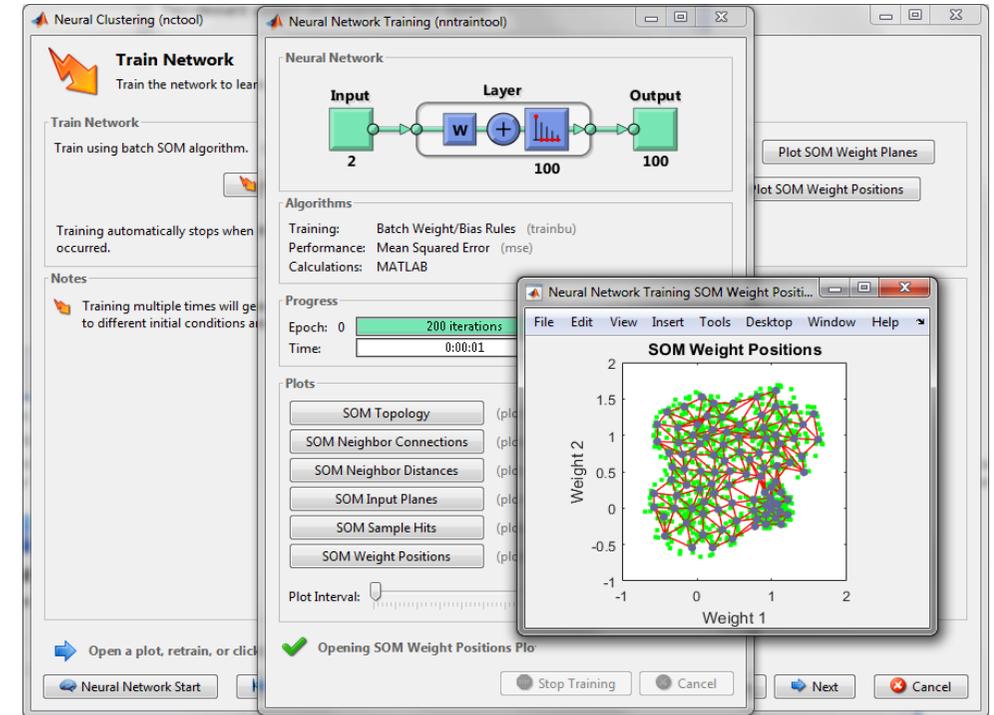
		Confusion Matrix			
		1	2	3	
Output Class	1	10 37.0%	0 0.0%	0 0.0%	100% 0.0%
	2	1 3.7%	8 29.6%	1 3.7%	80.0% 20.0%
	3	0 0.0%	0 0.0%	7 25.9%	100% 0.0%
		90.9% 9.1%	100% 0.0%	87.5% 12.5%	92.6% 7.4%
		1	2	3	
		Target Class			

Interactive Apps to focus on machine learning, not programming



Classification Learner App

MATLAB EXPO 2016



Neural network Apps

Features

- Train models
- Assess results
- Export models to the MATLAB or generate MATLAB code

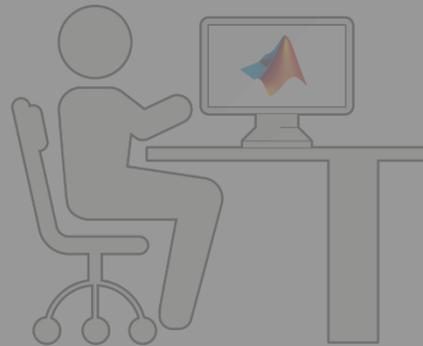
Why MATLAB?

3 Develop **embedded systems** with analytics powered functionality



Smarter Embedded Systems

2 Enable **Domain Experts to be Data Scientists**



Data Scientist

1 Analytics that increasingly require **both business and engineering data**

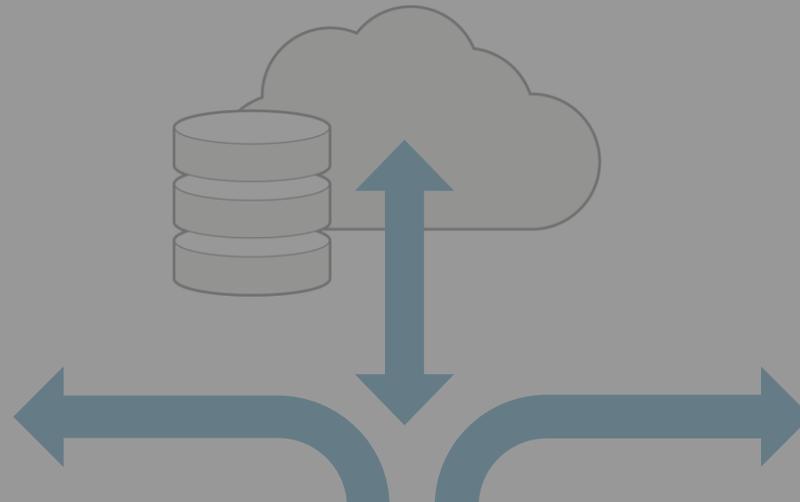
DATA

- Engineering, Scientific, and Field
- Business and Transactional

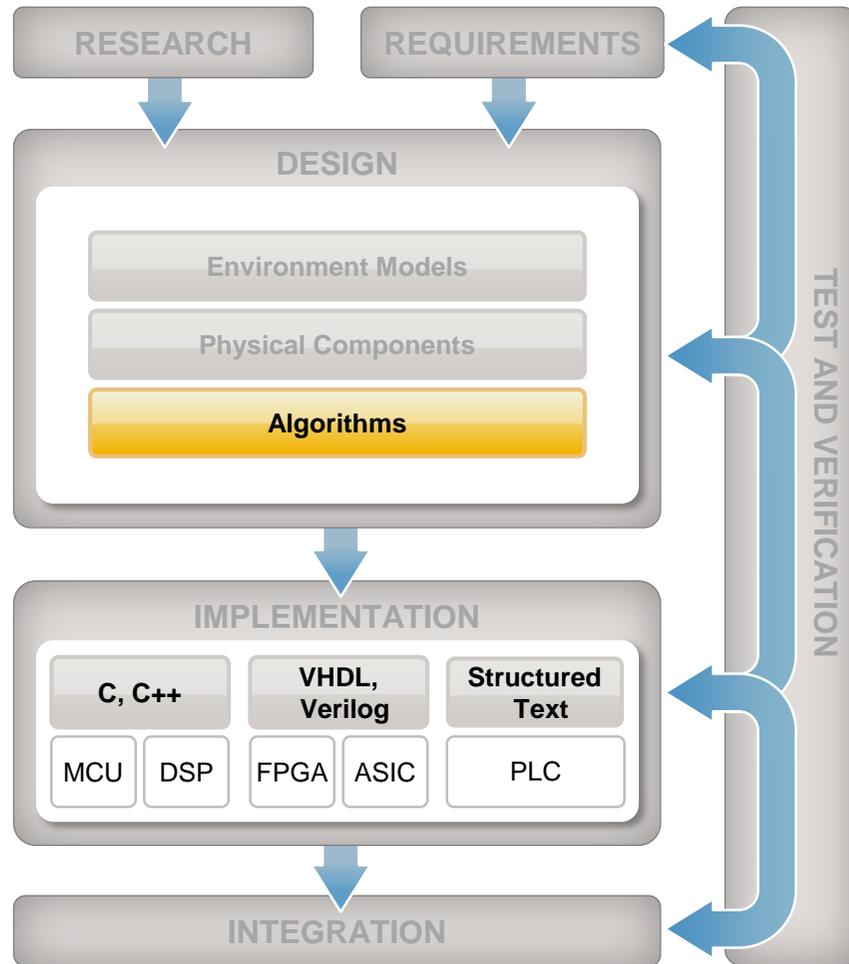
4 Develop analytics to run on **both enterprise and embedded platforms**



Business Systems



Smarter Embedded Systems



Airbus
Battery management



GM
Climate control



Festo
Industrial robots



Sonova
Hearing implants



Weinmann
Transport ventilator



ABB
Smart Grid controller



manroland
Printing presses

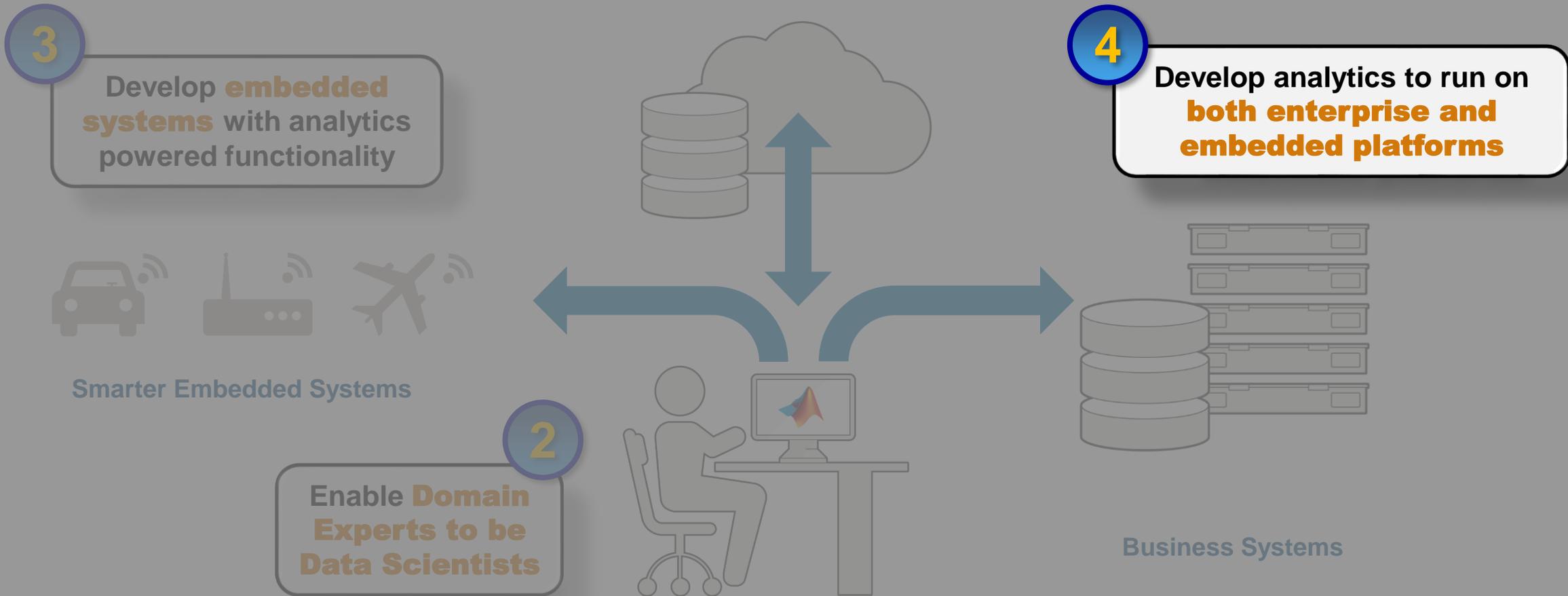


FLIR
Thermal imaging

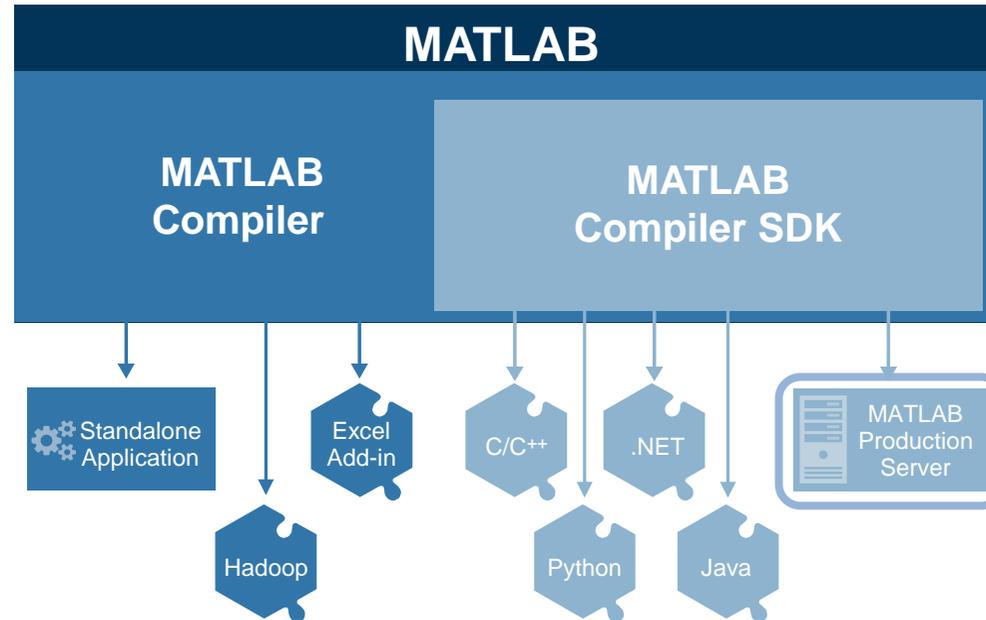


Daimler
Cruise controller

Why MATLAB?



Deploying Algorithms to Enterprise Systems



MATLAB Compiler enables sharing MATLAB programs without integration programming

MATLAB Compiler SDK provides implementation and platform flexibility for software developers

MATLAB Production Server provides the most efficient development path for secure and scalable web and enterprise applications

Enterprise Integration – Forecasting Model

Electricity Demand Foreca x

54.165.201.58:8080/DemandForecastWeb/demandForecast.jsp

Predictive Data Analytics Home Demand Forecasting Web Service Description Documentation

Select Zone

Zone Generate Forecast Model Diagnostics Report

Karte Satellit

udbury Québec

Réserve faunique La Vérendrye

Algonquin Provincial Park

Ottawa Montreal Sherbrooke

Kingston

Toronto Mississauga Hamilton Buffalo

MAINE VERMONT NEW HAMPSHIRE

Portland

NEW YORK MASSACHUSETTS RHODE ISLAND

Albany Providence

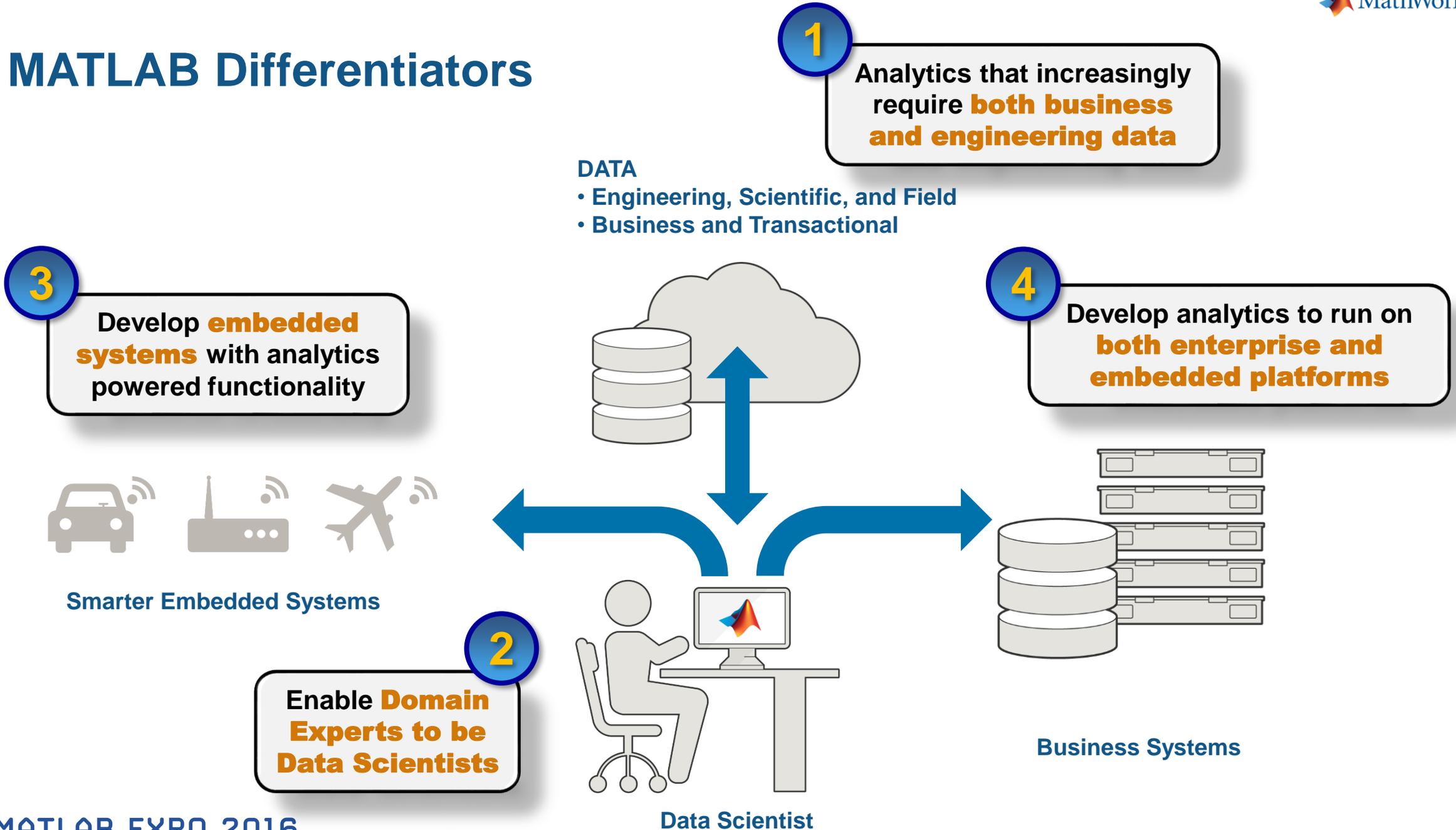
Pittsburgh PENNSYLVANIA New York City

Forecast

Select zone & generate forecast

Comparison

MATLAB Differentiators



Learn More

Presentations

1.2. Data Science mit MATLAB

Session Chair: Dr. Alexander Diethert

15:15 Analyse von operationellen Flugdaten aus einem Hadoop System unter Verwendung von MapReduce und dem MATLAB Distributed Computing Server
Lukas Höhndorf, TU München

15:45 Mensch-Maschine-Interface zur multisensorischen Prozessüberwachung in der Polymerindustrie
Michael Kohlert, Mondi Gronau
Dr. Sarah Drewes und Elmar Tarajan, MathWorks

16:15 Algorithmen für Predictive Maintenance effizient entwickelt mit MATLAB
Dr. Sarah Drewes, MathWorks

Machine Learning

mathworks.com/machine-learning

The screenshot shows the MathWorks website with the following visible content:

- Header: MathWorks logo, navigation links (Contact sales, Trial offer), and a search bar.
- Section: "Machine Learning with MATLAB" with a sub-header "Machine Learning Consulting Services".
- Highlighted Section (red box): "MathWorks Consulting Services" featuring a video thumbnail of a man and the text "Get up and running fast. MathWorks Consulting Services - industry experience and MATLAB and Simulink expertise." with a "Watch video" button.
- Section: "Why Choose MathWorks Consulting? Consulting" with the text "Working with MathWorks Consulting gives you the advantage of their years of project work, industry backgrounds, and deep MATLAB and Simulink know-how." and a link "Explore the value of MathWorks Consulting Services".
- Section: "Classification" with the text "Build models to classify... into different categories".
- Section: "Algorithms" with sub-sections for "support vector machine (SVM), boosted and bagged decision trees." and "regularization".
- Section: "unüberwachten Lernens zum Erkennen von Merkmalen in großen Datensätzen, s... Vorhersagemodellen vorgestellt. Beispiele und Übungen verdeutlichen Verfahren Themen sind unter anderem:"

Training