Responding to the AI Challenge
Learning from Physical Industries

David Rich
drich at mathworks dot com

How can other industries inform AI in finance?

- Four learnings from outside of finance
- Three areas of exploration
- Two quick MATLAB PSAs (public service announcements)

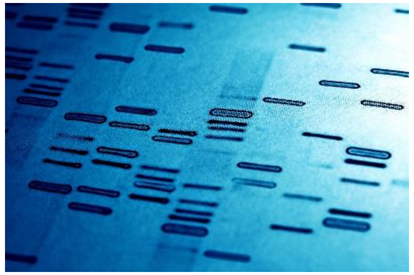
AI in this talk includes;
machine learning, deep learning, reinforcement learning…

Our Customers / Key Industries

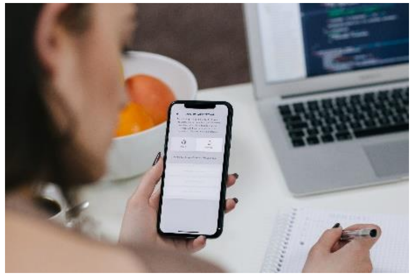

Aerospace and Defense


Automotive


Biological Sciences


Biotech and Pharmaceutical


Communications


Electronics


Energy Production


Financial Services

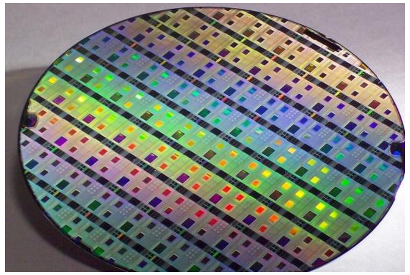

Industrial Machinery


Medical Devices


Process Industries


Neuroscience

Railway Systems


Semiconductors

Software and Internet

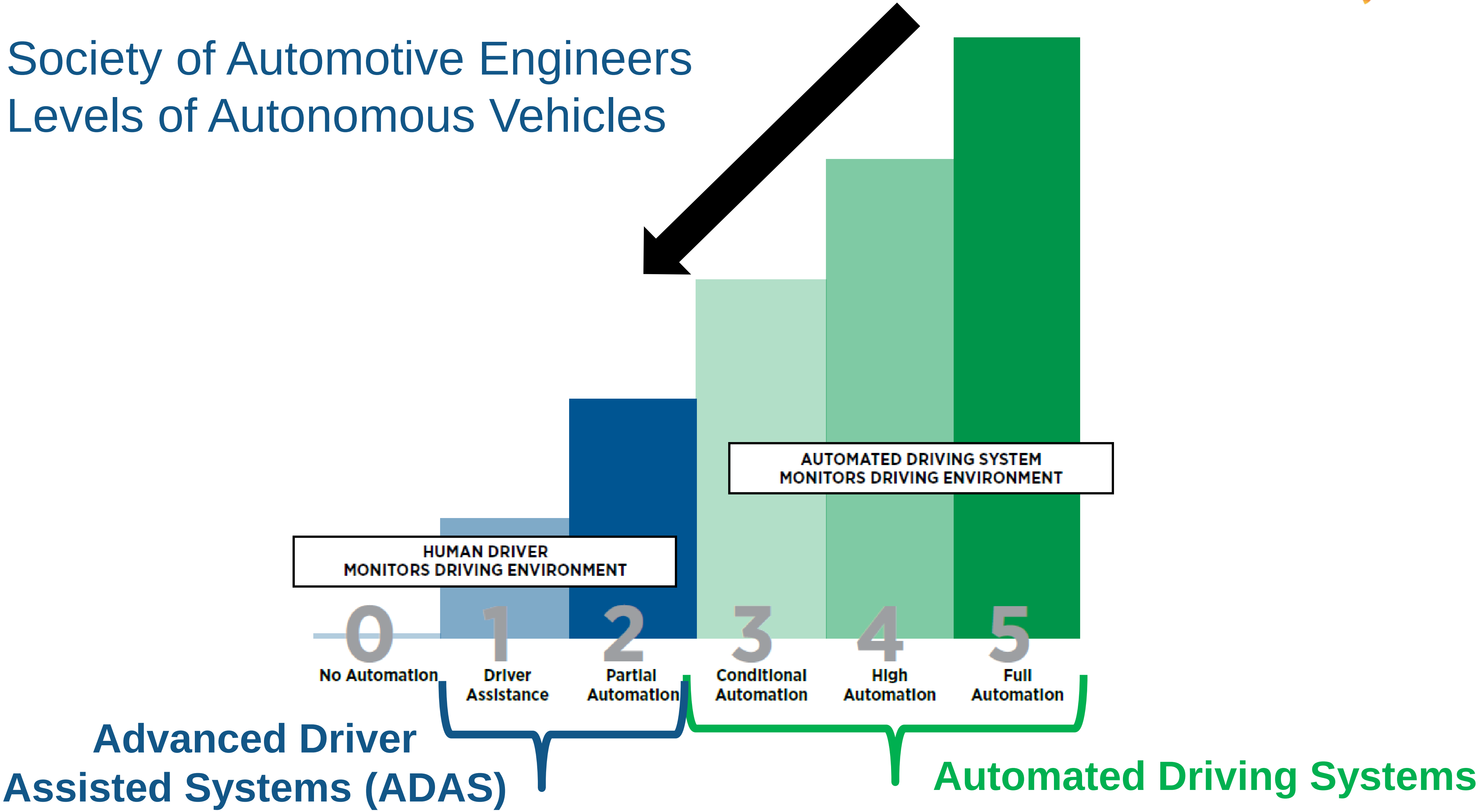
Four Learnings from Other Industries

1. Plenty of value away from the “obvious” applications
2. There’s no reason not to look for your keys under the street light:
   If you have data use it
3. Regulations can be tough – but perhaps not for advice.
4. If you don’t have data, can you create it?

Four Learnings from Other Industries

1. Plenty of value away from the “obvious” applications
2. There’s no reason not to look for your keys under the street light:
   If you have data use it
3. Regulations can be tough – but perhaps not for advice.
4. If you don’t have data, can you create it?

Society of Automotive Engineers Levels of Autonomous Vehicles



Subaru (a customer)

Advanced Driver-Assistance Systems

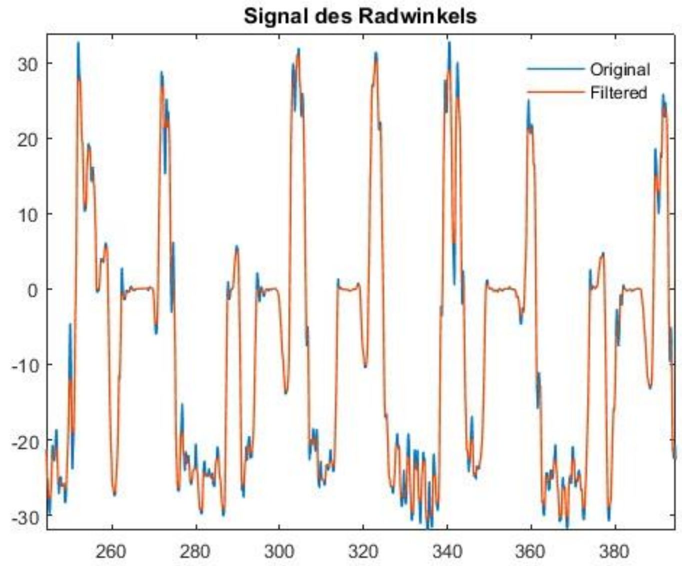
Critical safety features for everyone

Detects obstacles, applies brakes, adjusts cruise control, and stays in lane

BMW - Machine Learning to Detect Oversteering

"With little previous experience with machine learning, we completed a working ECU prototype capable of detecting oversteering **in just three weeks**." Tobias Freudling, BMW Group

- Engineers gathering and cleaned data
- Explored many machine learning approaches with Classification Learner App
- Generated code for vehicles on test track



Link to technical article

Four Learnings from Other Industries

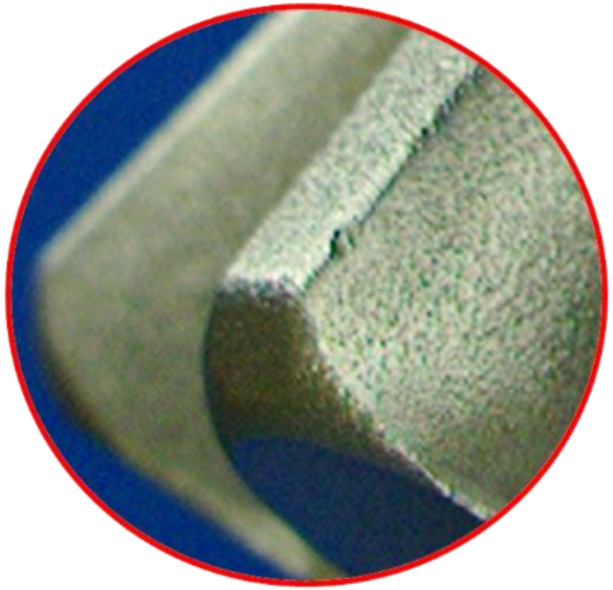
1. Plenty of value away from the “obvious” applications
2. There’s no reason not to look for your keys under the street light:
   If you have data use it
3. Regulations can be tough – but perhaps not for advice.
4. If you don’t have data, can you create it?

Four Learnings from Other Industries

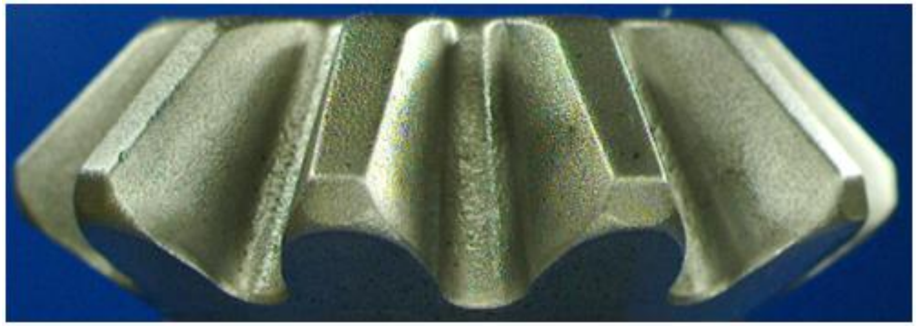
1. Plenty of value away from the “obvious” applications
2. There’s no reason not to look for your keys under the street light:
   If you have data use it
3. Regulations can be tough – but perhaps not for advice.
4. If you don’t have data, can you create it?

Musashi Seimitsu Industry Co.,Ltd.

Detect Abnormalities



Automated visual inspection of 1.3 million bevel gear per month



Manufacturers often have a trove of labelled data

Four Learnings from Other Industries

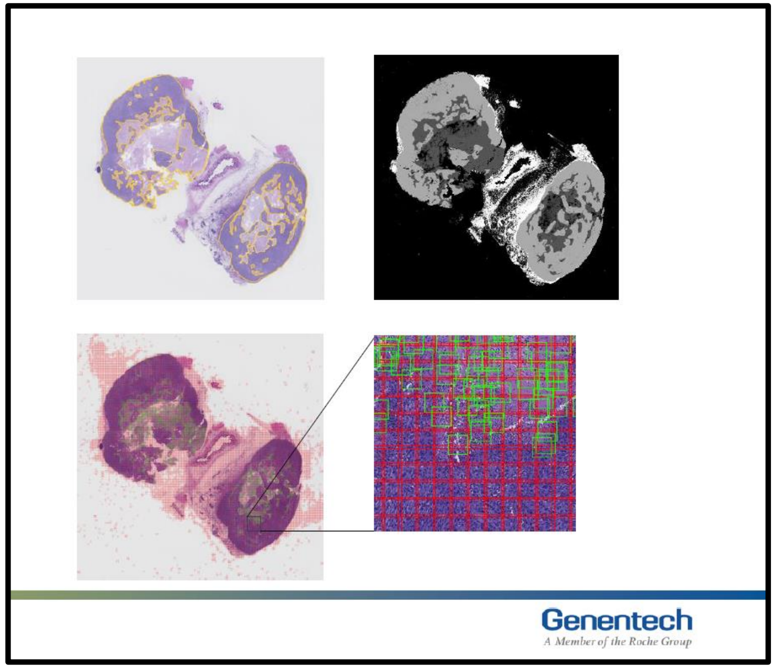
1. Plenty of value away from the “obvious” applications
2. There’s no reason not to look for your keys under the street light:
   If you have data use it
3. Regulations can be tough – but perhaps not for advice.
4. If you don’t have data, can you create it?

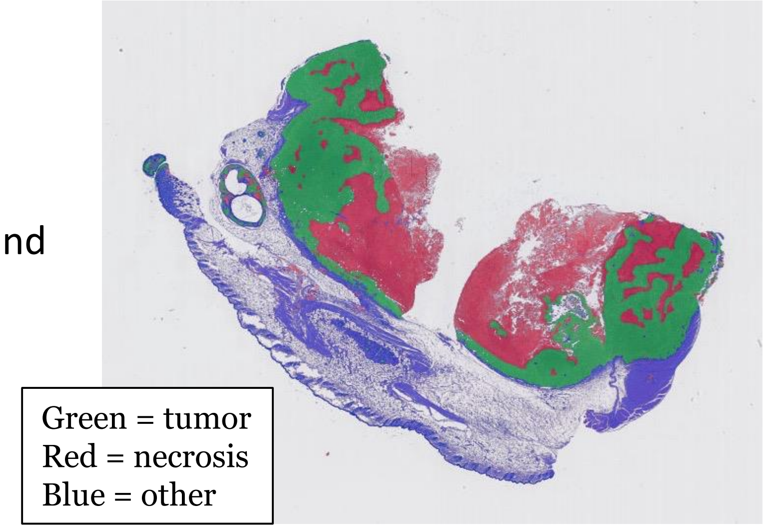
Genentech

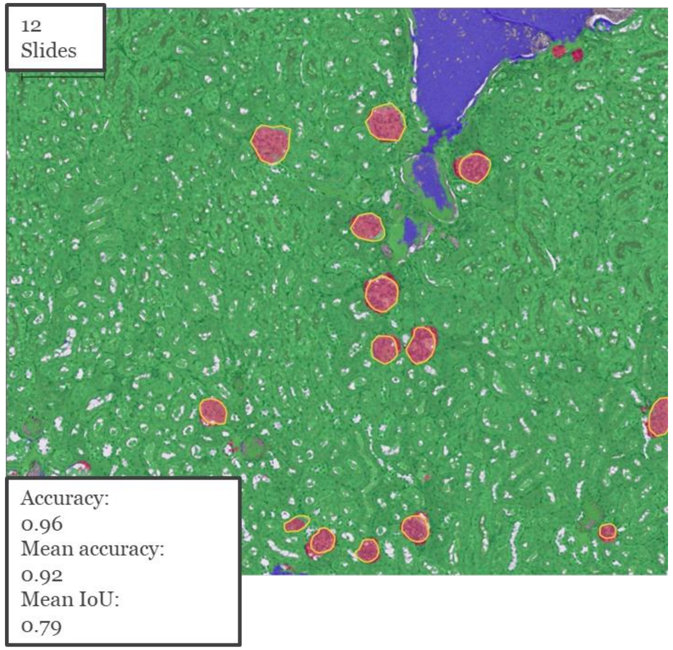
Deep Convolutional Neural Networks for Digital Pathology Analysis

Generate training data iteratively

- Model is iteratively improved by adding more data
- Removes need to annotate tumor by hand







Segment tumor tissue from necrosis

- Segmentation of massive 25k x 25k images
- Trained and deployed U-Net semantic segmentation algorithm

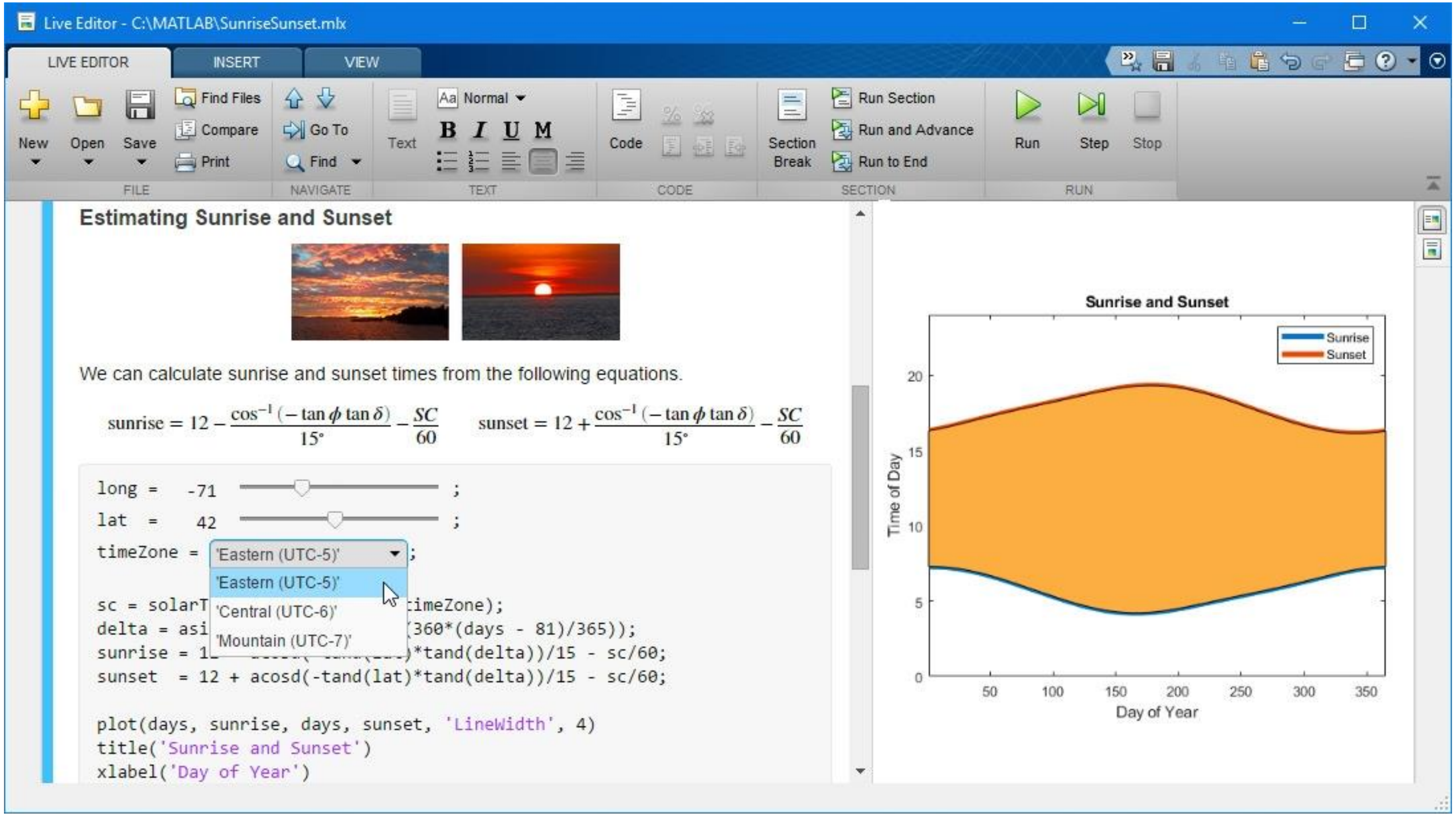
Not a diagnosis!
Assists pathologist

Presented at American Conference on Paramacometrics (7th October 2018)
Deep Convolutional Neural Networks for Digital Pathology Analysis
Jeffrey Eastham-Anderson, Kathryn Mesh, Jeff Hung, Andrea Dranberg (MathWorks)

MATLAB PSA #1

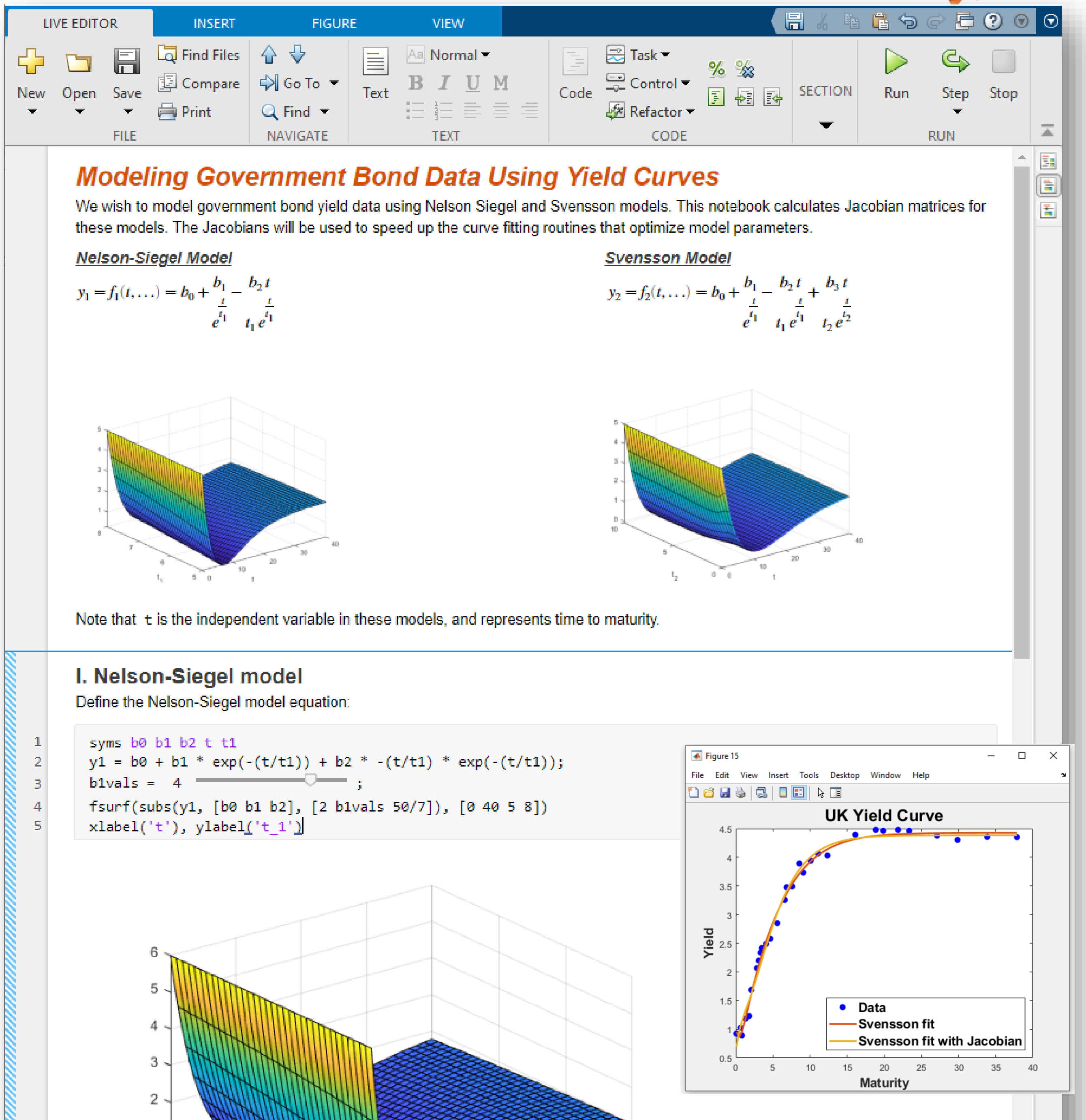
Use the **Live Editor** to create scripts that combine code, output, and formatted text in an executable notebook.

R2016a



Now integrated with Symbolic Math Toolbox

Symbolic and Numeric
in one Live Editor Notebook



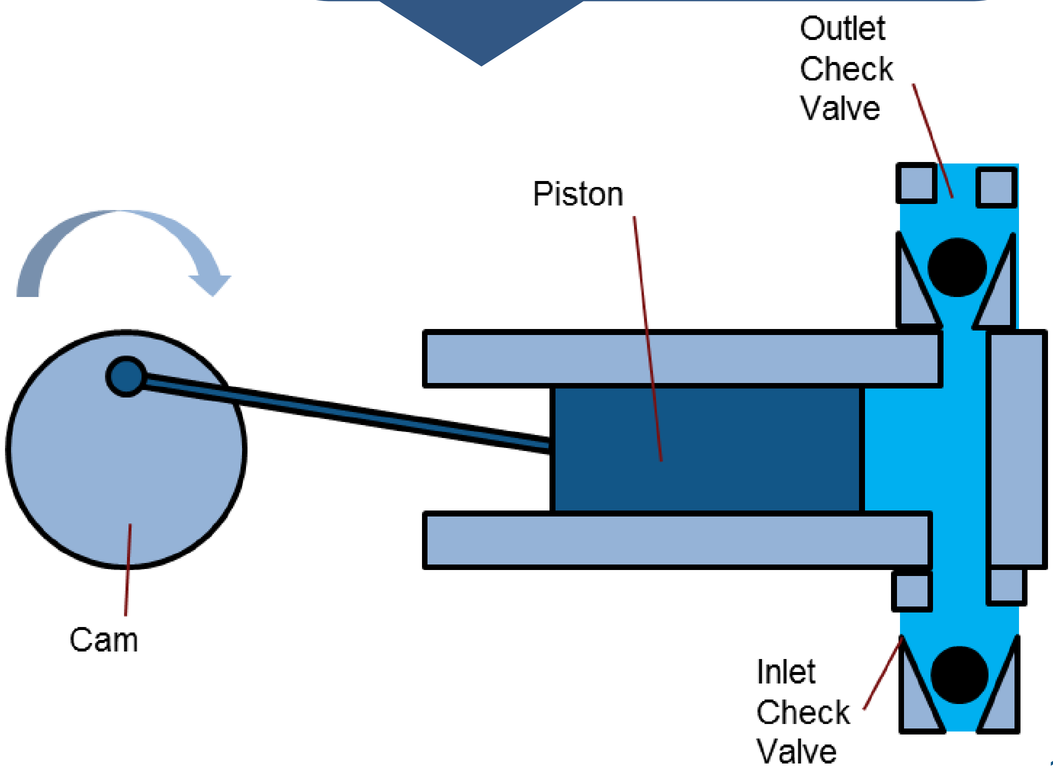
Four Learnings from Other Industries

1. Plenty of value away from the “obvious” applications
2. There’s no reason not to look for your keys under the street light:
   If you have data use it
3. Regulations can be tough – but perhaps not for advice.
4. If you don’t have data, can you create it?

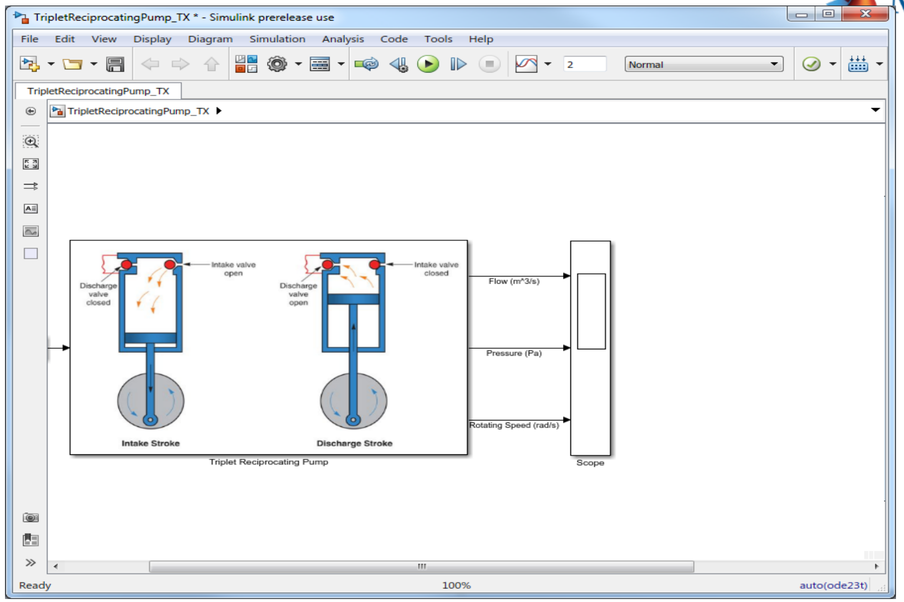
Predictive Maintenance: Reciprocating Pump

Predict pump failures in real-time using sensor data

"I keep my machines healthy and running so how do I get failure data to train a model?"



Generate Data

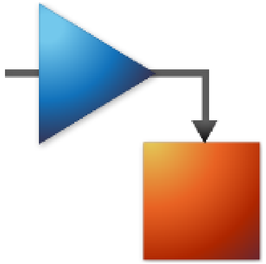


Build model

Obtain sensor data

Sensor Data

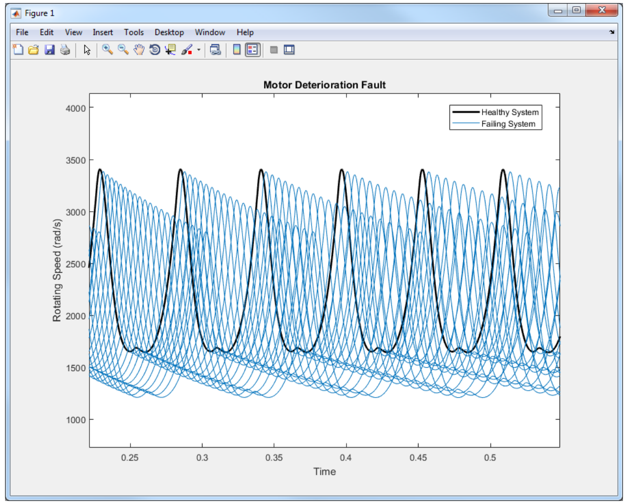
Simulink Model



Fine tune model

Run simulations

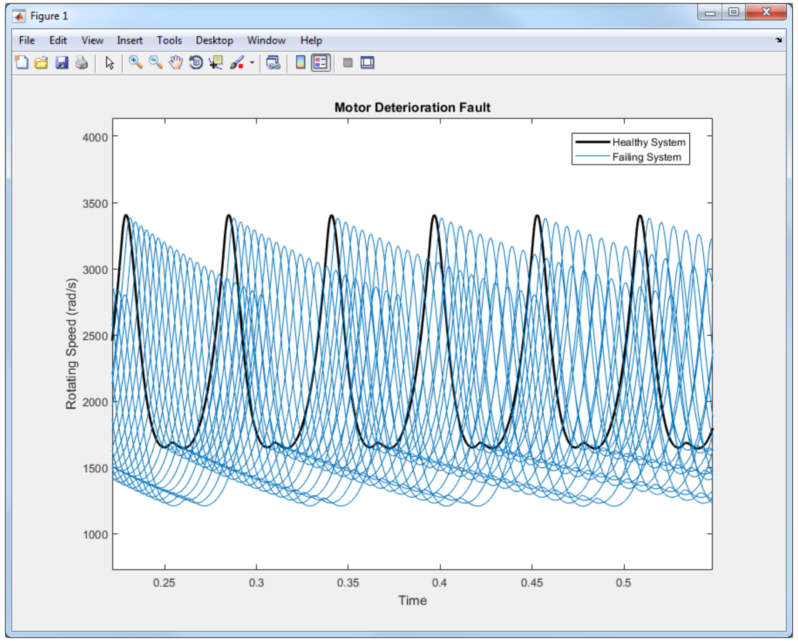
Generated Failure Data



Incorporate failure modes

Inject Failures

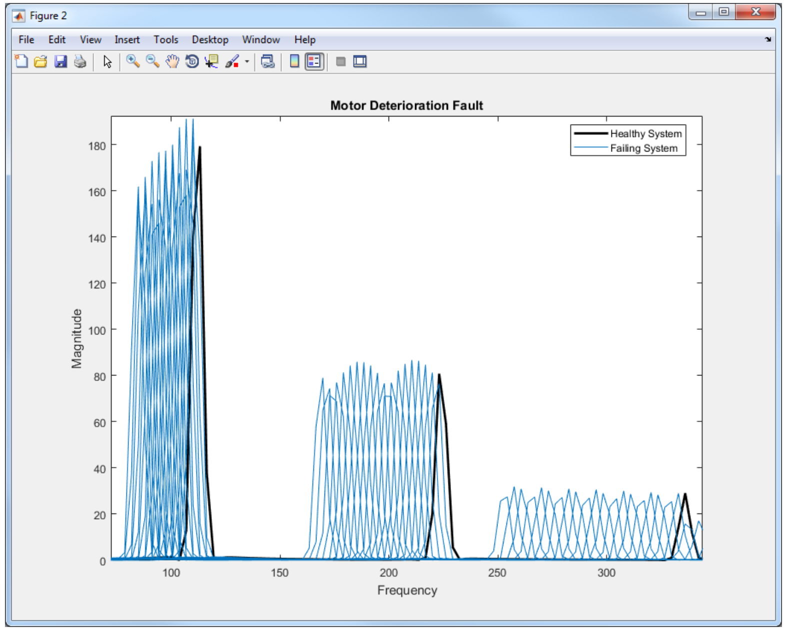
Preprocess Data



Failure Data (Sensors/Simulation)

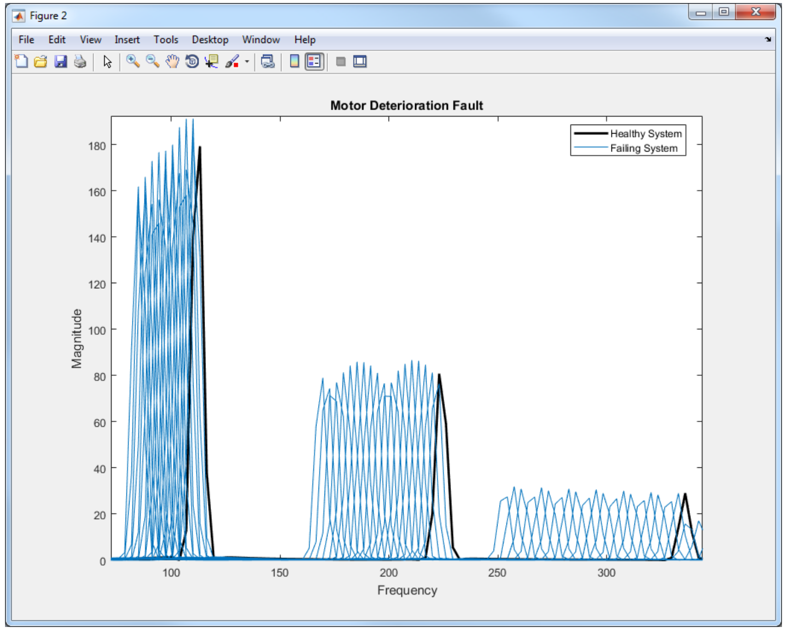
Data Preprocessing Methods

- Time Domain
- Frequency Domain
- Time-Frequency Domain



Preprocessed Data

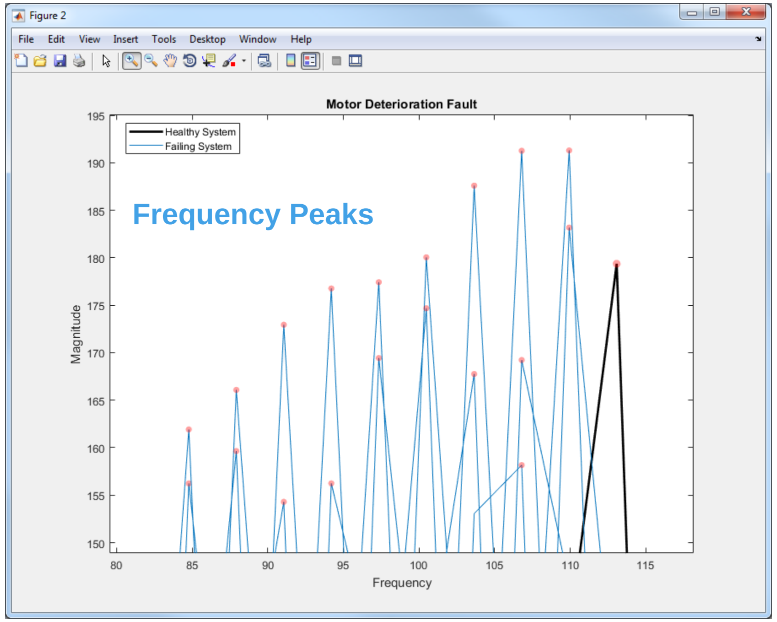
Feature Extraction & Condition Monitoring



Preprocessed Data

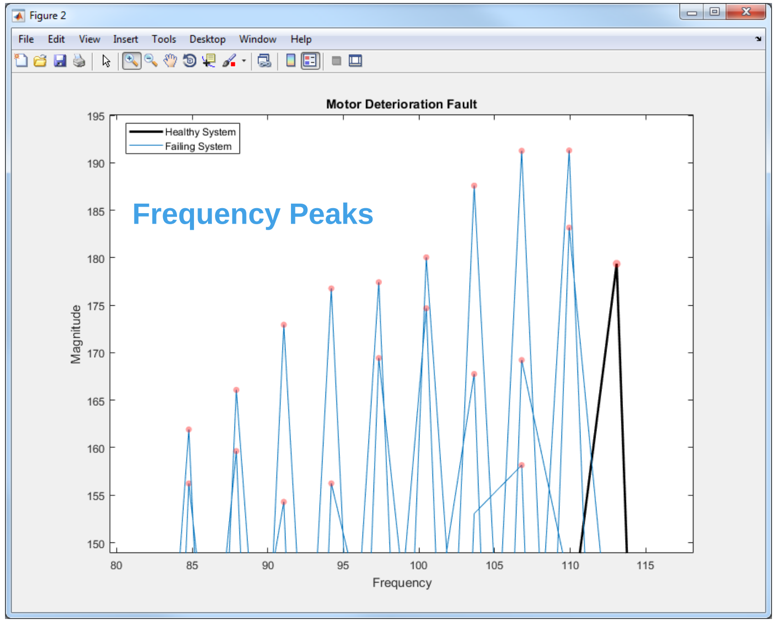
Feature Extraction Methods

- Order/Modal Analysis
- Time-Frequency Analysis
- Input-Output Models
- Model Coefficients & States
- Residual Generation
- ...



Health Indicators

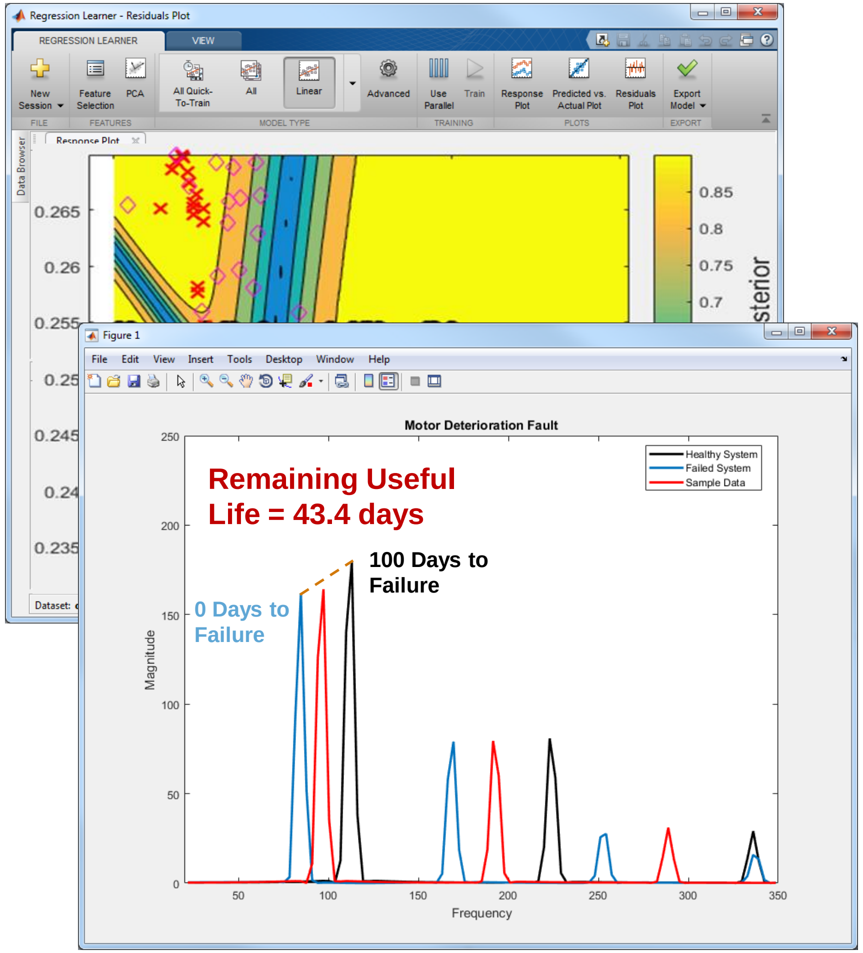
Predictive Model Training



Health Indicators

Predictive Methods

- Anomaly Detection
- Fault Classification
- Remaining Useful Life
- Trending
- Hazard Distributions
- Time series Forecasting
- ...



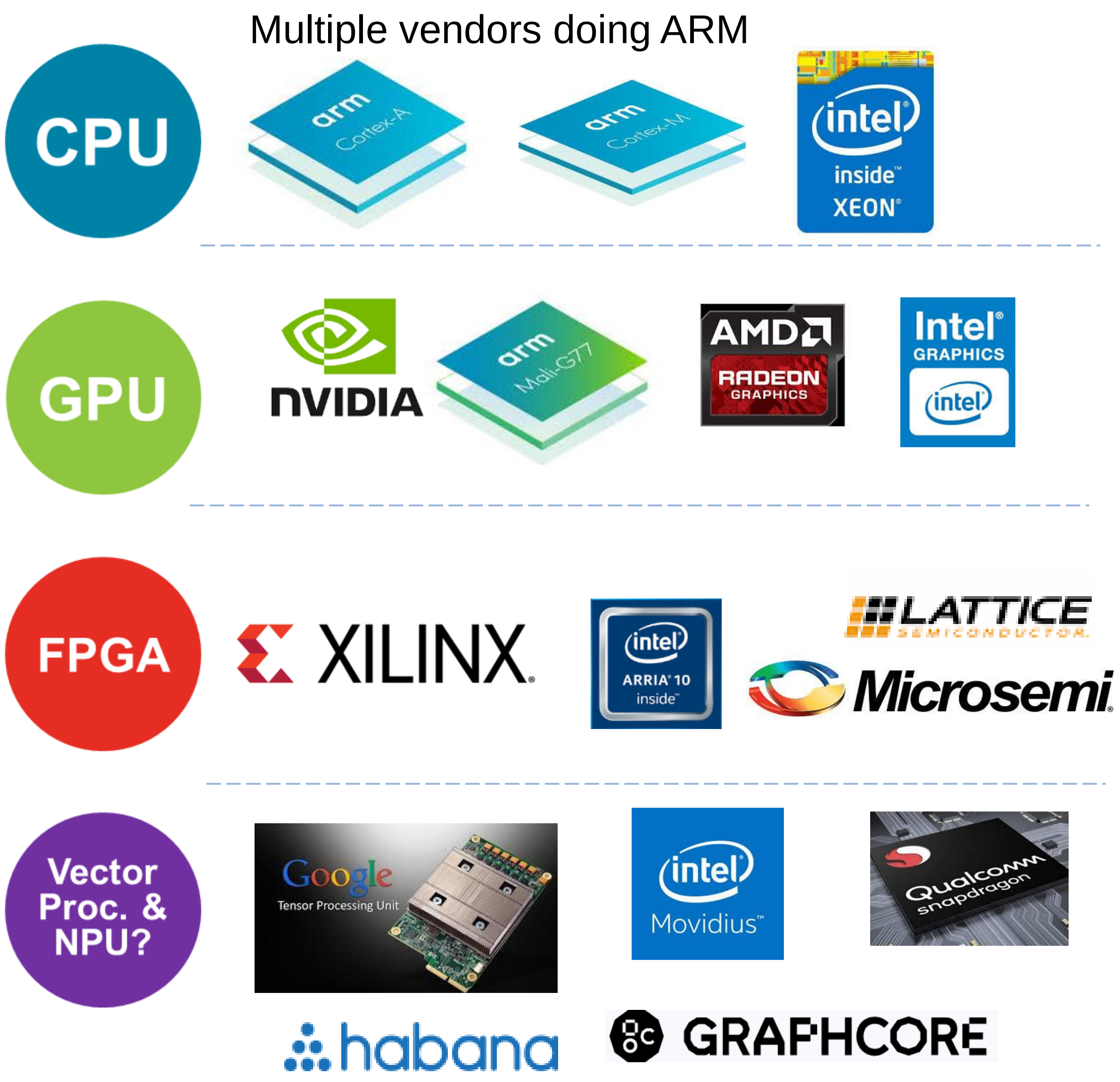
Three Topics to Watch

Three Topics to Watch

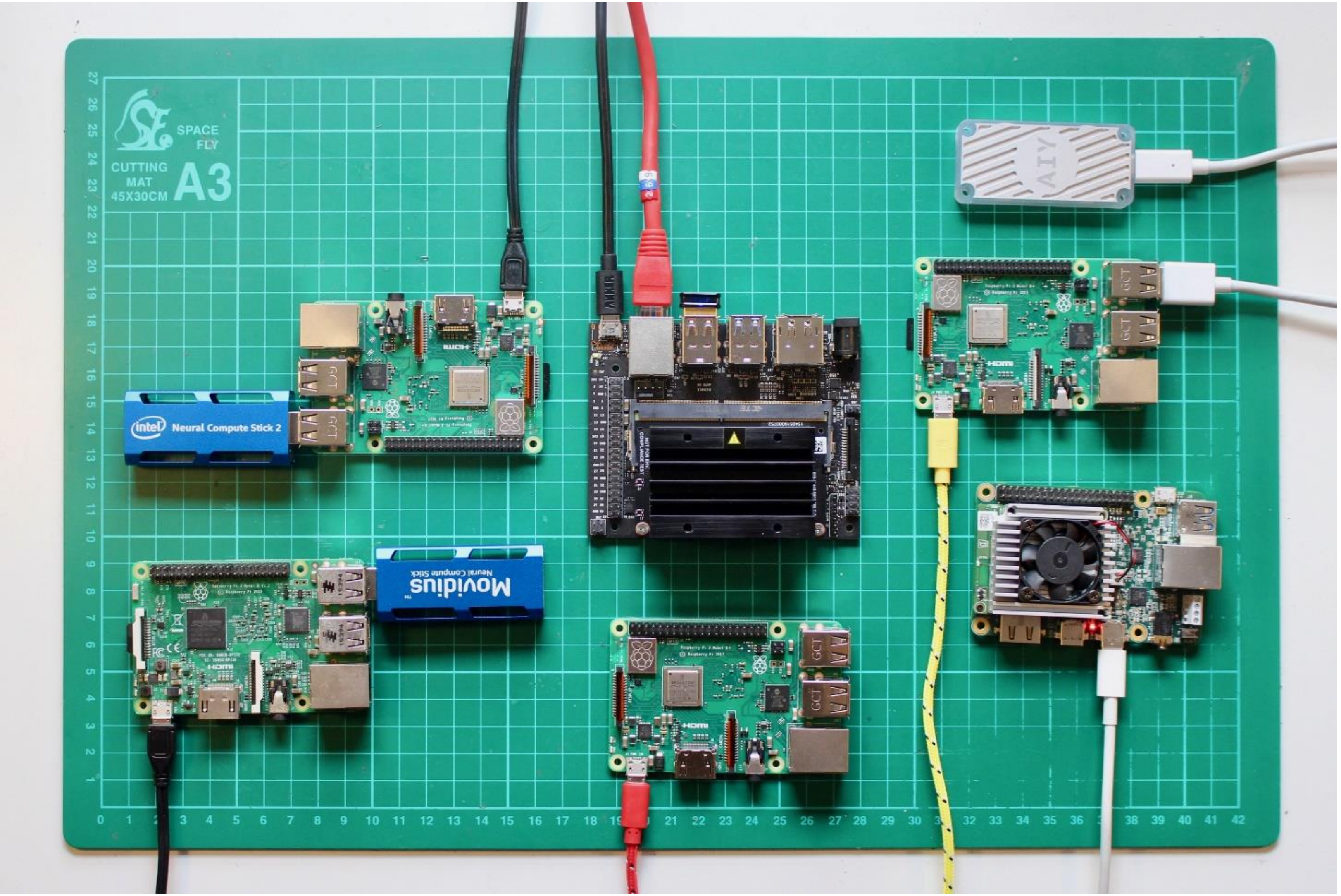
#3 Fragmentation in Hardware Architecture? (for Deep Learning)

Fragmentation in Hardware Architecture?

- NVIDIA is the standard for data-center deep learning
- But there are challengers;
  - FPGA from Xilinx, Intel, others
  - AMD Radeon
  - Google's TPU
  - Embedded processors from ST, TI, Renasas, Infineon
- Over $1B of venture investment in AI chip startups
- Cloud is an accelerator
- Training vs. Inference



GPU falling out of favor as hardware for embedded deployment?

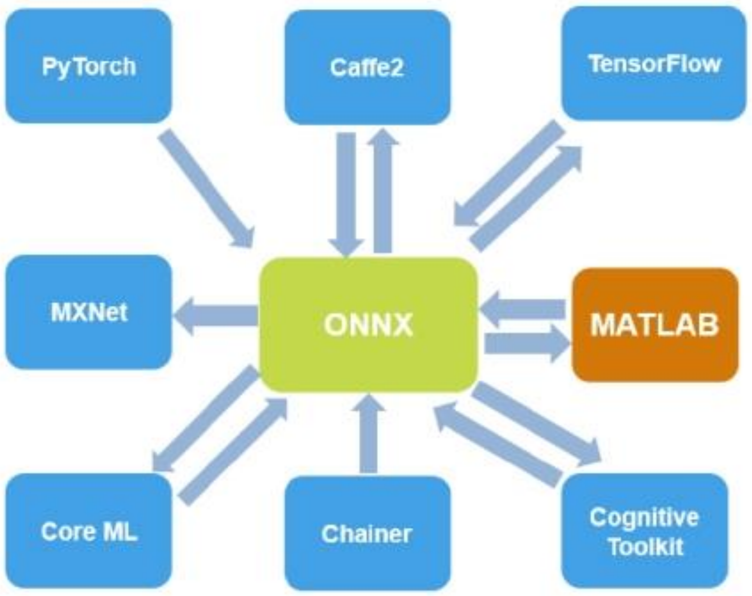


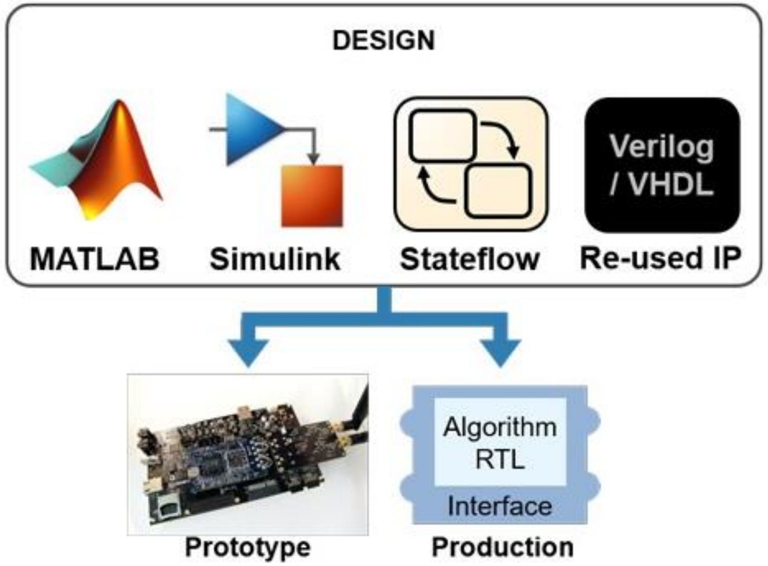
Edge computing hardware zoo:
[Intel Neural Compute Stick 2](#) (left, top)
[Movidius Neural Compute Stick](#) (left, bottom)
[NVIDIA Jetson Nano](#) (middle, top)
[Raspberry Pi 3, Model B+](#) (middle, bottom)
[Coral USB Accelerator](#) (right, top) Google TPU
[Coral Dev Board](#) (right, bottom) Google TPU

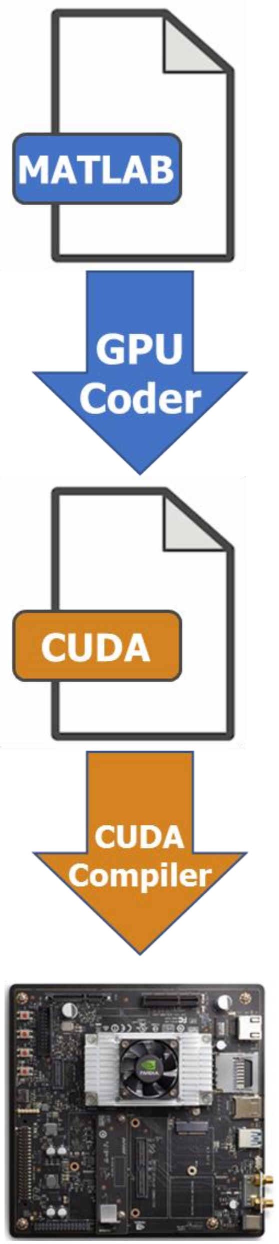
Image courtesy of Dr. Allasdair Allan
http://bit.ly/great-big-roundup

If multiple architectures become viable, then what?

- Evaluate HW for purpose – choose target
- Develop in high level language
- Transform to target executable
- A small number of finance customers doing this today (GPU, FPGA). Will this grow?







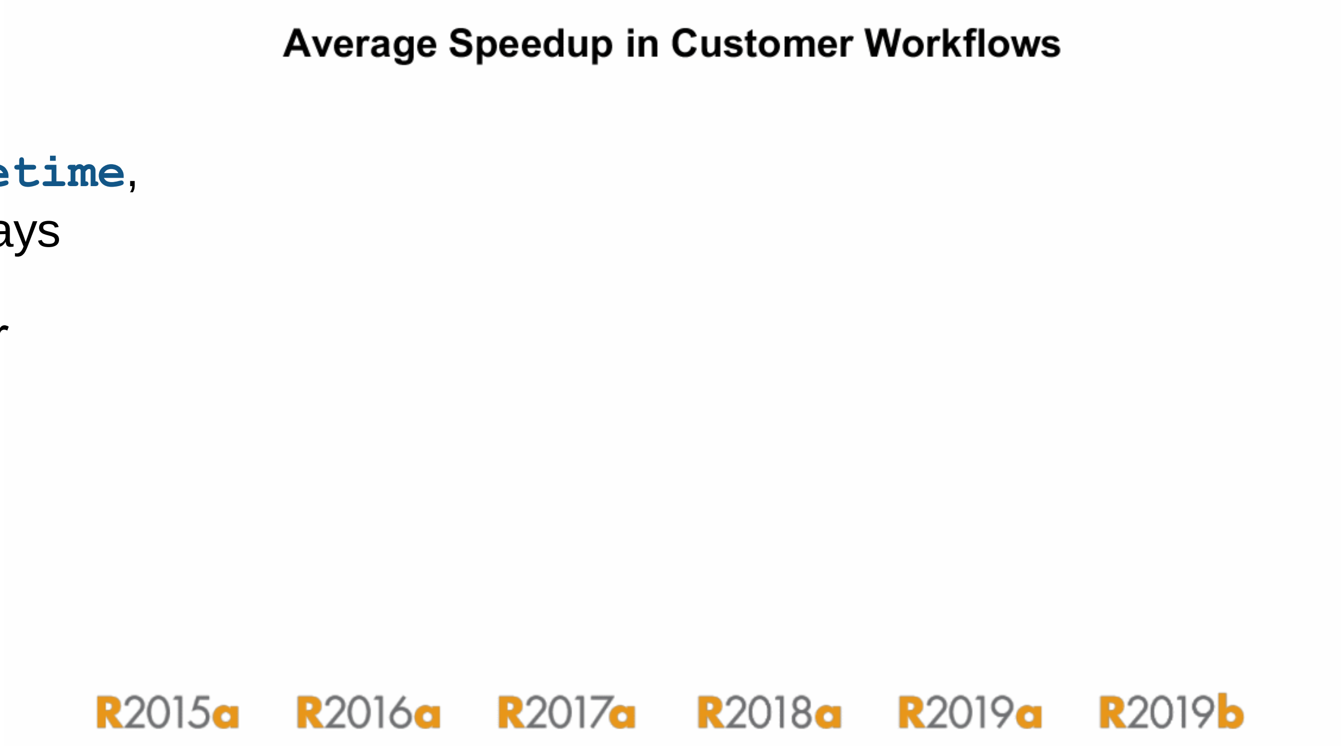
MATLAB PSA #2

Run MATLAB code faster with redesigned execution engine.

R2015b

- All MATLAB code is now JIT compiled

- Incremental improvements each release
  - Faster assignment into large `table`, `datetime`, `duration`, and `calendarDuration` arrays
  - Construct objects and set properties faster
  - Render plots with large numbers of markers faster using less memory
  - Increased speed of MATLAB startup



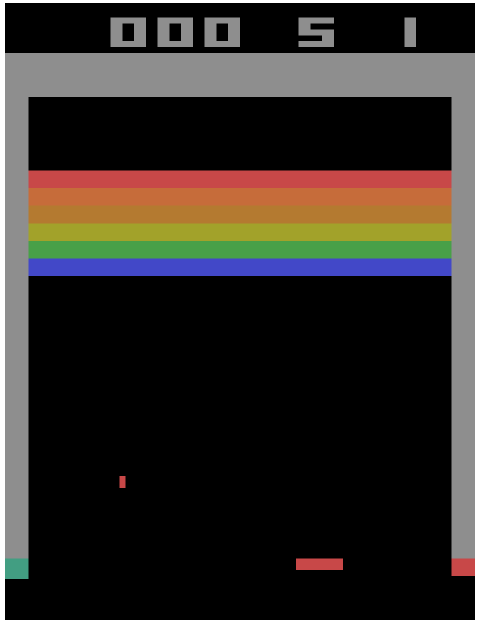
Three Topics to Watch

#3 Fragmentation in Hardware Architecture? (for Deep Learning)

#2 Reinforcement Learning

Reinforcement Learning in the News Focuses Mainly On…

- Board Games
  - Chess
  - Go
- Video Games
  - Atari
  - DoTA, Starcraft
- Recommendation Systems



NETFLIX RESEARCH

Posterior Sampling for Large Scale Reinforcement Learning

Nov 21, 2017

Posterior sampling for reinforcement learning (PSRL) is a popular algorithm for learning to control an unknown Markov decision process (MDP). PSRL maintains a distribution over MDP parameters and in an episodic fashion samples MDP parameters, computes the optimal...

…But Increasingly Being Seen In Context of Autonomous Systems

- Learn Complex Tasks
  - Manipulation
  - Planning
  - Navigation
  - Control


BBC Sign in | News | Sport | Weather | Shop | Reel | Travel | Mo

NEWS

Home | Video | World | US & Canada | UK | Business | Tech | Science | Stories | Enter

Technology

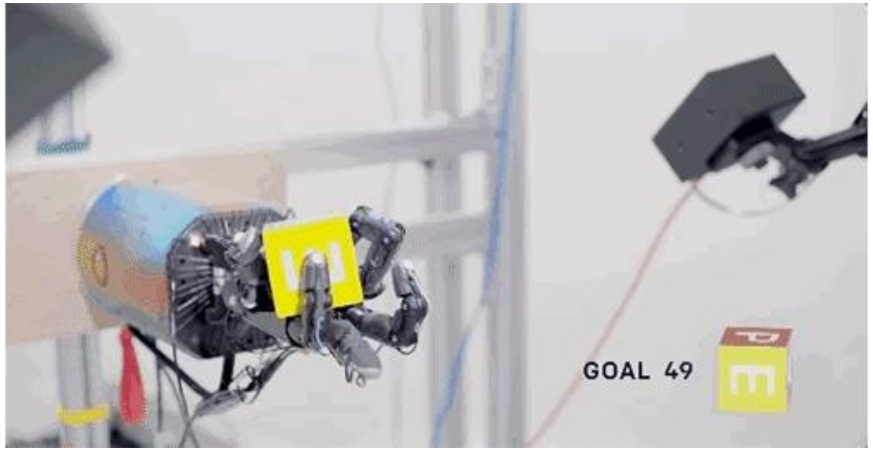
AI solves Rubik's Cube in one second

5 hours ago Share

GETTY IMAGES

The Rubik's Cube is a puzzle challenging people to create six sides of a single colour

An artificial intelligence system created by researchers at the University of California has solved the Rubik's Cube in just over a second.



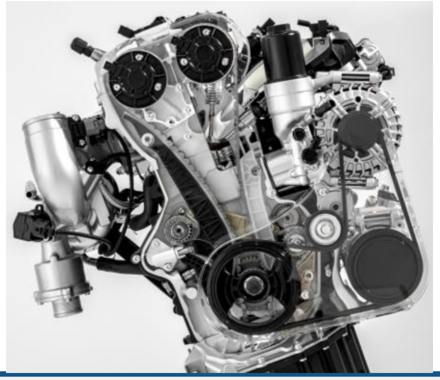
Source: OpenAI

Source: Google

Traditional “Controls” Customers Have Proactively Engaged

100+ customers have spoken to us about Reinforcement Learning since 2018

- Reinforcement learning needs a lot of data, usually generated from models
- Models can incorporate conditions hard to emulate in the real world
- **Many of them have MATLAB and Simulink models that can be reused**



Autonomous Systems (e.g. Robots)

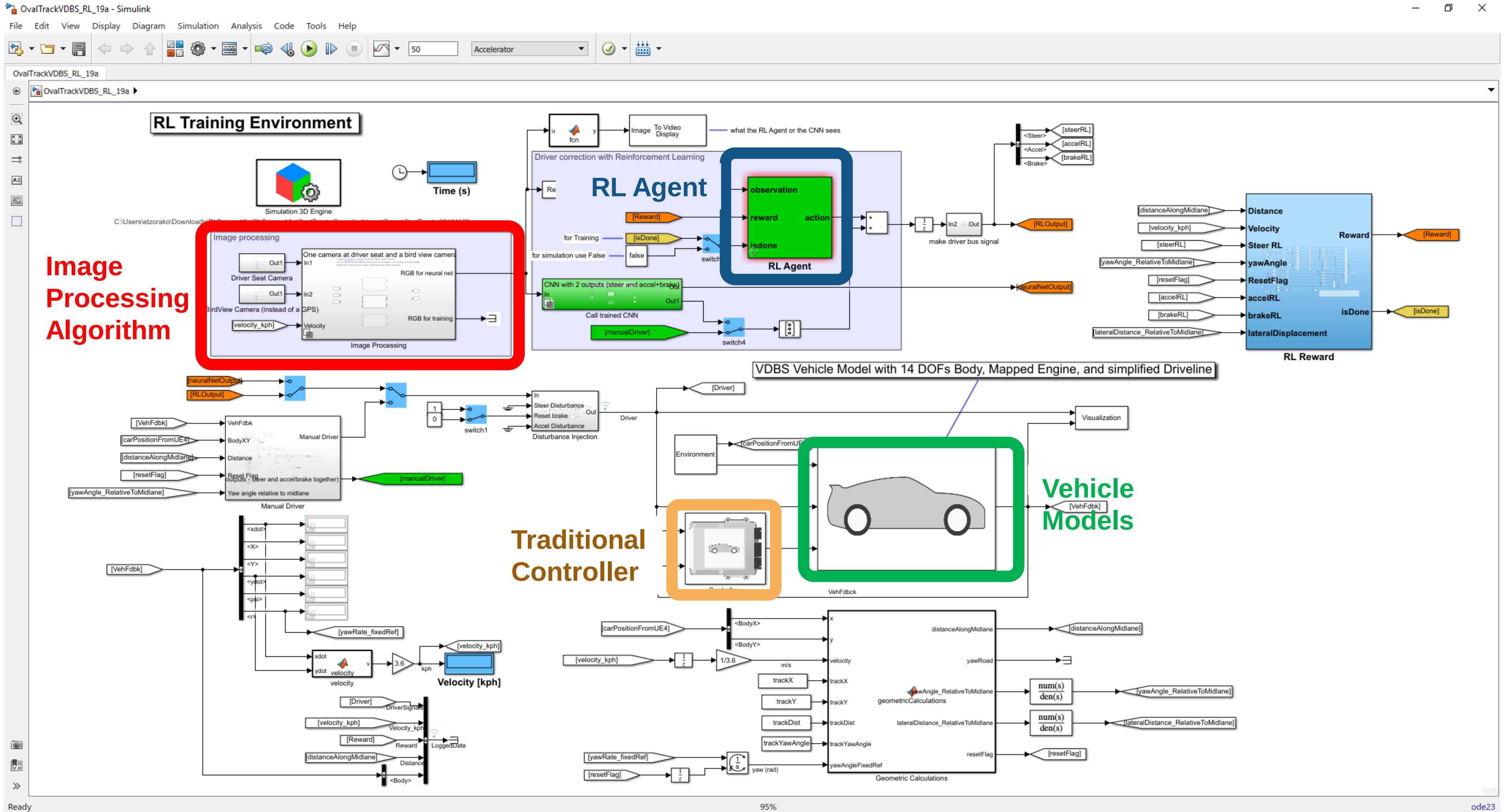
Calibration Problems (e.g. Engine Maps)

Controller Design for Nonlinear Systems

Automated Driving

Using Reinforcement Learning to Improve Driving Control

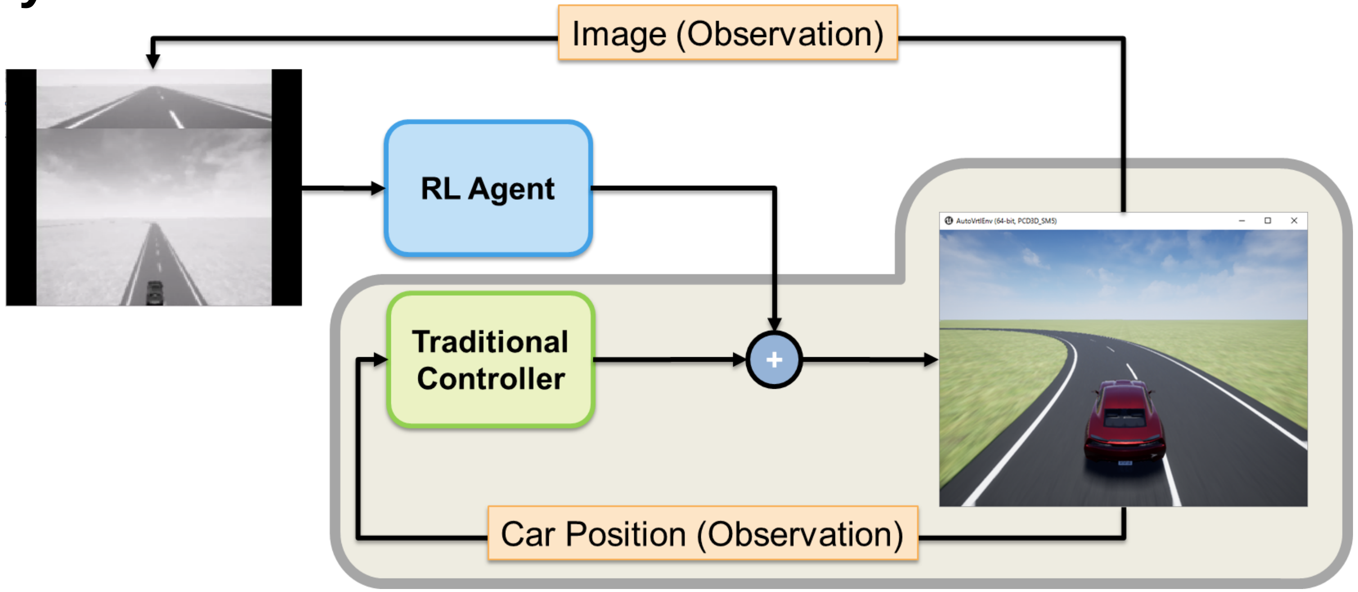
Models like this are used by our customers to develop controllers and other algorithms

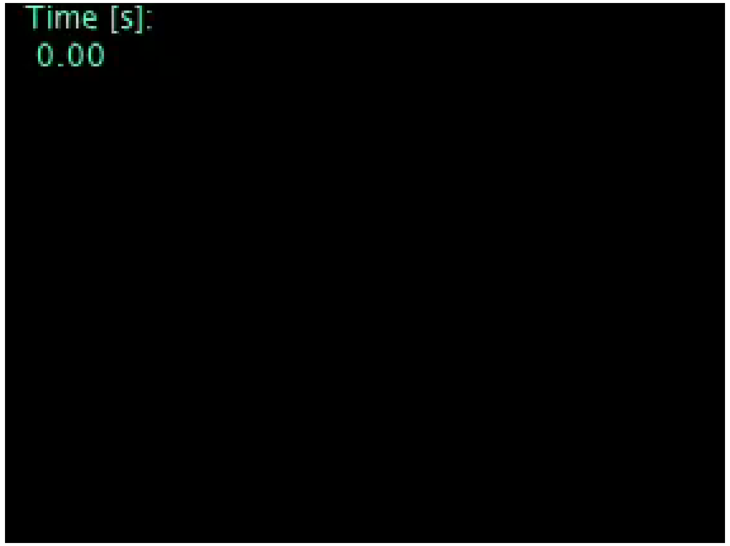


RL for Autonomous Driving – Co-simulating with Unreal Engine

Project With A Major Automotive Company

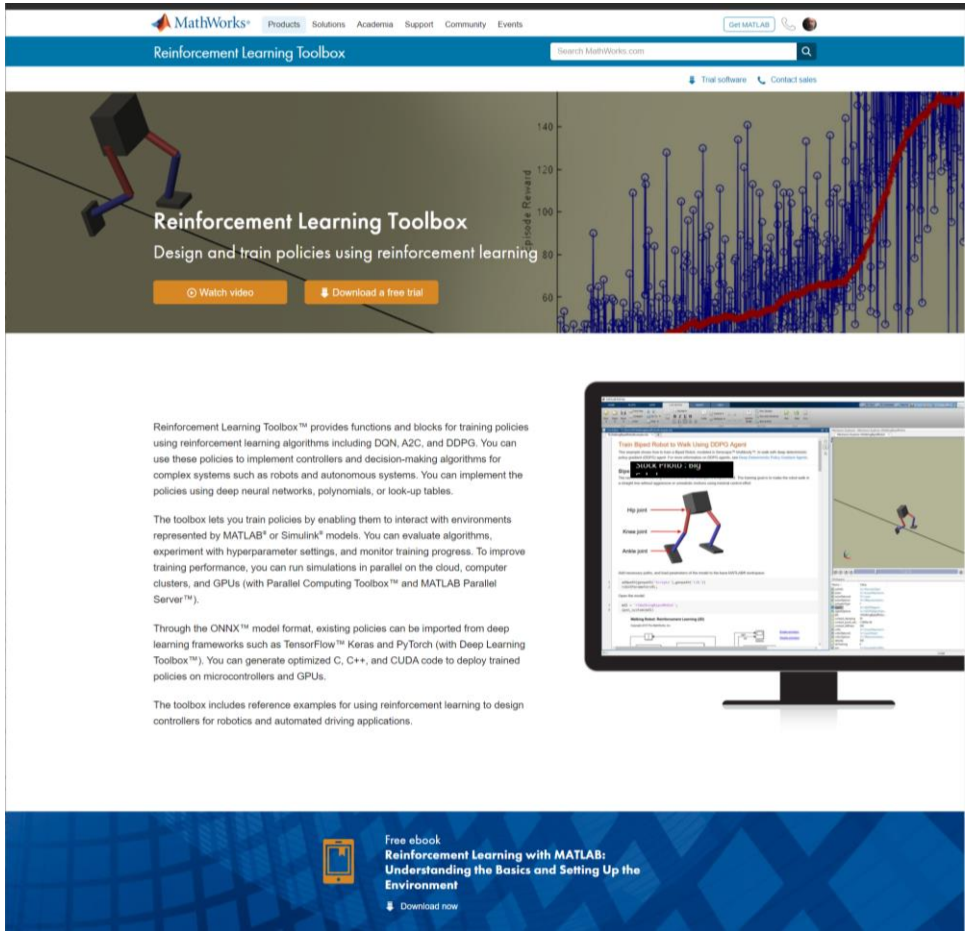
- Step 1: Trained deep neural network (DNN) based driver
- Step 2: Use RL to improve performance of DNN-based driver
- Step 3: Use improved DNN to **augment** traditional controller
- Result: 2+ sec (7%) faster than the original driver





RL Interest Growing in Finance

- Notably JP Morgan LOXM (Limit Order Execution Management)
- Positive results for our first experiments
  - Stock trading
  - Hedging
- Academic activity – mostly automated trading
- Petter Kolm from NYU speaking later;
  - *Dynamic Replication and Hedging:*
    A Reinforcement Learning Approach



Three Topics to Watch

#3 Fragmentation in Hardware Architecture? (for Deep Learning)

#2 Reinforcement Learning

#1 Explainability and V&V for AI

Explainability in Finance: Principles of fairness require being able to explain why the model is making decisions

2. Use of personal attributes as input factors for AIDA-driven decisions is **justified**.

4. AIDA-driven **decisions** are regularly **reviewed** so that models behave as designed and intended.

8. Firms using AIDA are **accountable** for both internally developed and externally sourced AIDA models.

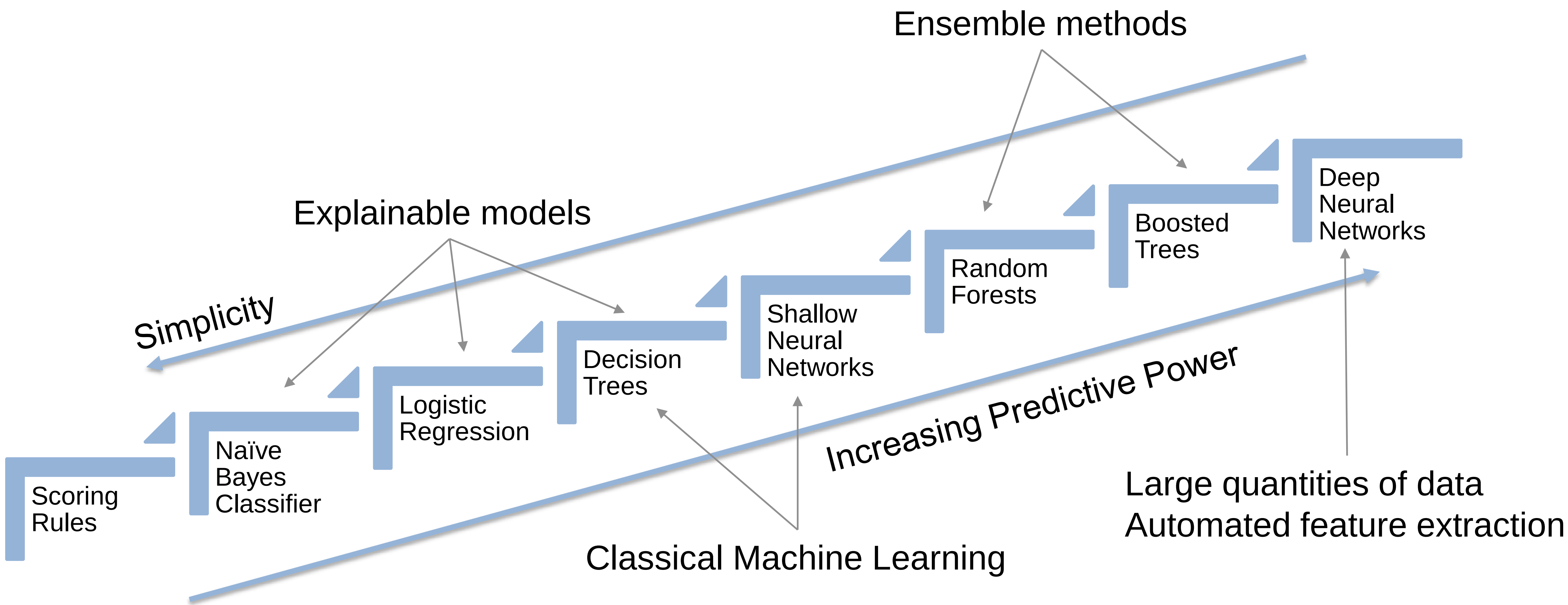
13. Data subjects are provided, upon request, clear **explanations** on what data is used to make AIDA-driven decisions about the data subject and how the data affects the decision.


Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector

MAS Monetary Authority of Singapore

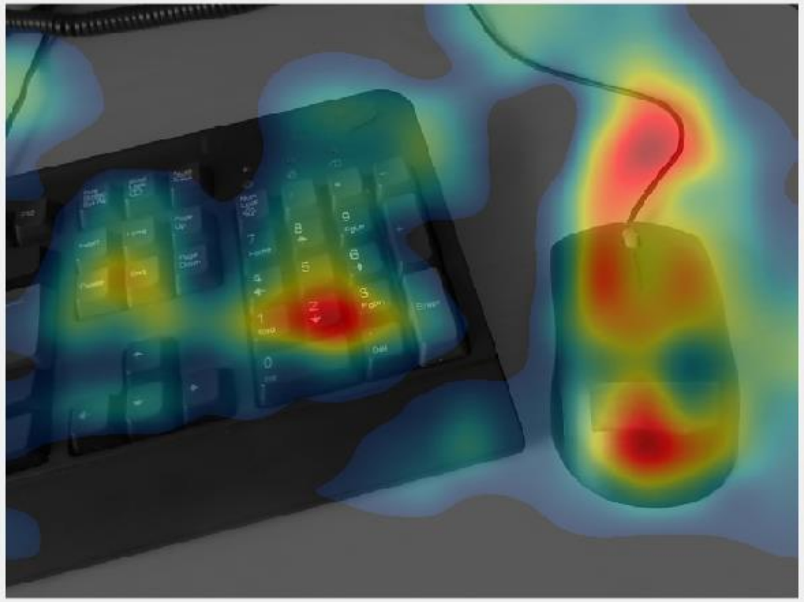
"AIDA" refers to artificial intelligence or data analytics, which are defined as technologies that assist or replace human decision-making.
https://www.mas.gov.sg/~/media/MAS/News%20and%20Publications/Monographs%20and%20Information%20Papers/FEAT%20Principles%20Final.pdf

Trade-off between predictive power and explainability



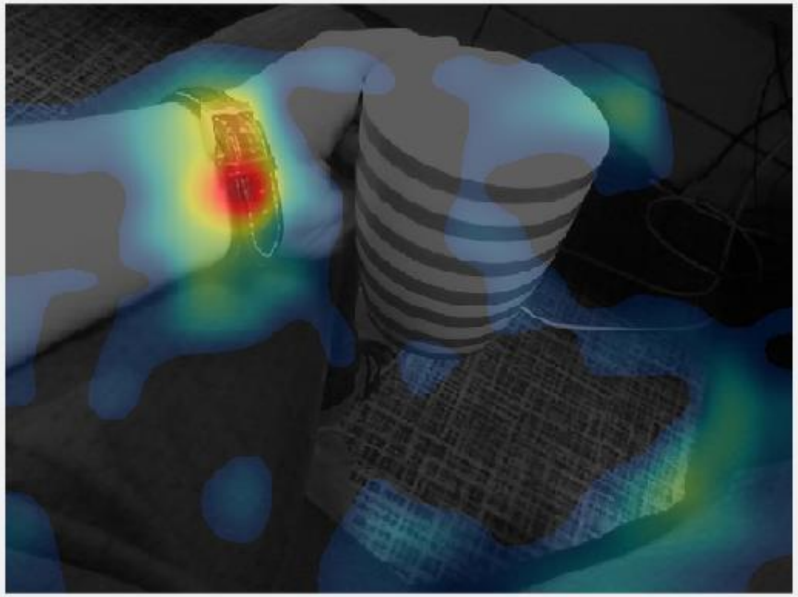
Attribution Reveals the Why Behind Deep Learning Decisions

mouse, 0.46095
remote control, 0.24144
computer keyboard, 0.12748



Classified as “keyboard” due in part to the presence of the mouse

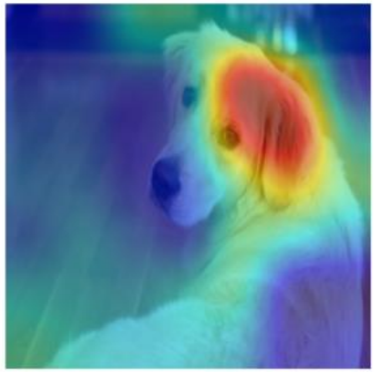
buckle, 0.14911
sock, 0.087194
mailbag, 0.056052



Incorrectly classified “coffee mug” as “buckle” due to the watch

golden retriever (0.55)

Grad-CAM



https://www.mathworks.com/help/deeplearning/examples/investigate-network-predictions-using-class-activation-mapping.html
https://www.mathworks.com/help/deeplearning/ug/gradcam-explains-why.html

Across Industries there are different meanings for…

Verification & Validation for AI

EXPLAINABILITY

Can you explain the working of AI model in human-understandable terms?

INTERPRETABILITY

Can you observe and trace cause and effect in an AI model and explain the rationale of the decision?

ROBUSTNESS

Is AI system immune from spoofing and other common attacks ?

DATA PRIVACY

Can an attacker deduce sensitive training data from output of AI model or system?

RIGOR & TRUST

Has AI system been developed with defined, traceable and rigorous process?

SAFETY CERTIFICATION

Has AI system been developed with safety lifecycle as key component.

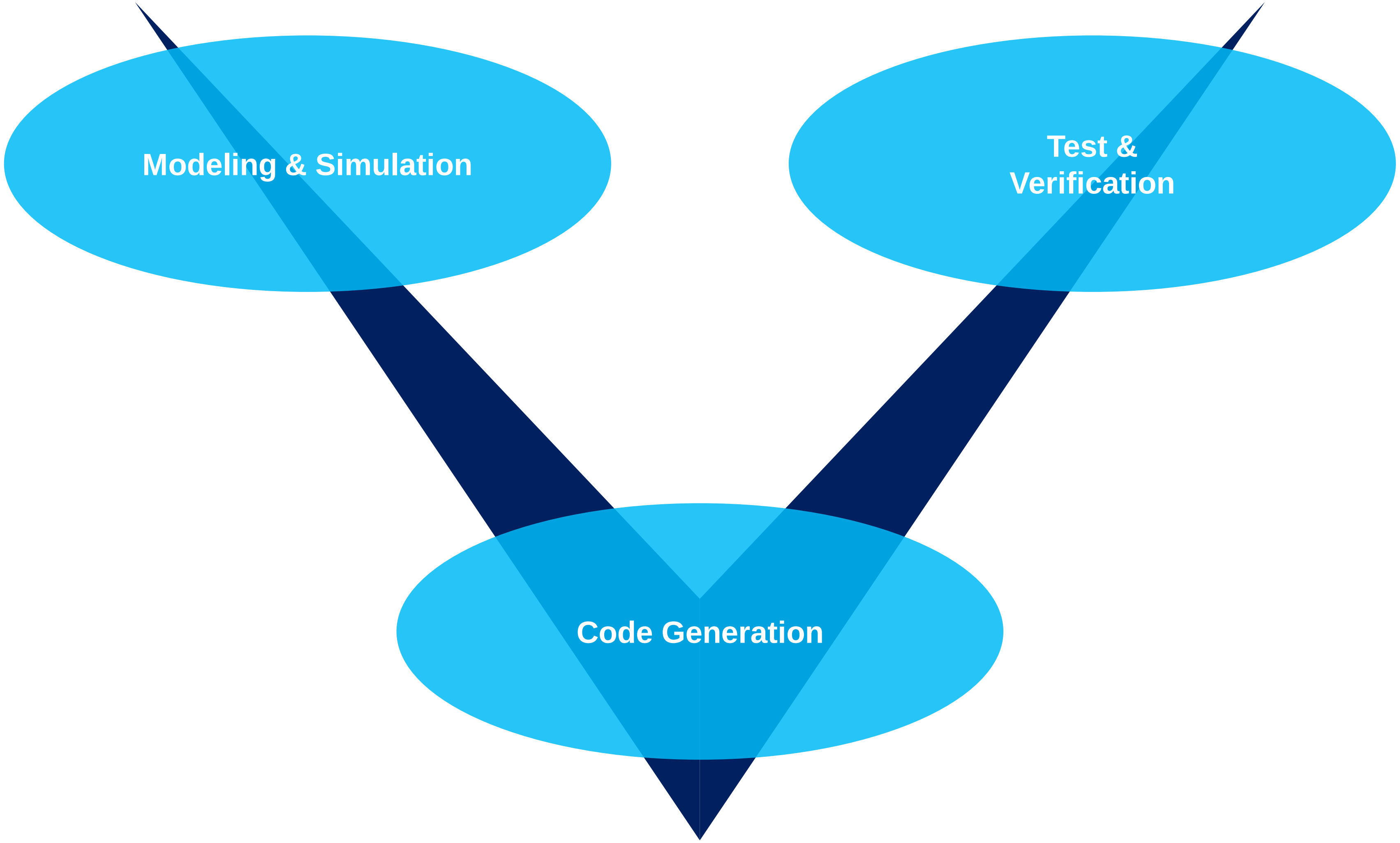
Common safety practices

- **Redundancy**: Redundant implementation with voting
- **FMEA**: Failure mode analysis to ensure safe behavior in anticipated scenarios
- **Monitoring**: Monitoring and logging to record decisions for post-mortem analysis
- **Auditing**: Audit trails of development activities
- **Data Separation**: Firewalls between training and test/validation data to ensure appropriate accuracy metrics are computed
- **Best Practices**: Documentation of best practices being followed to improve repeatability

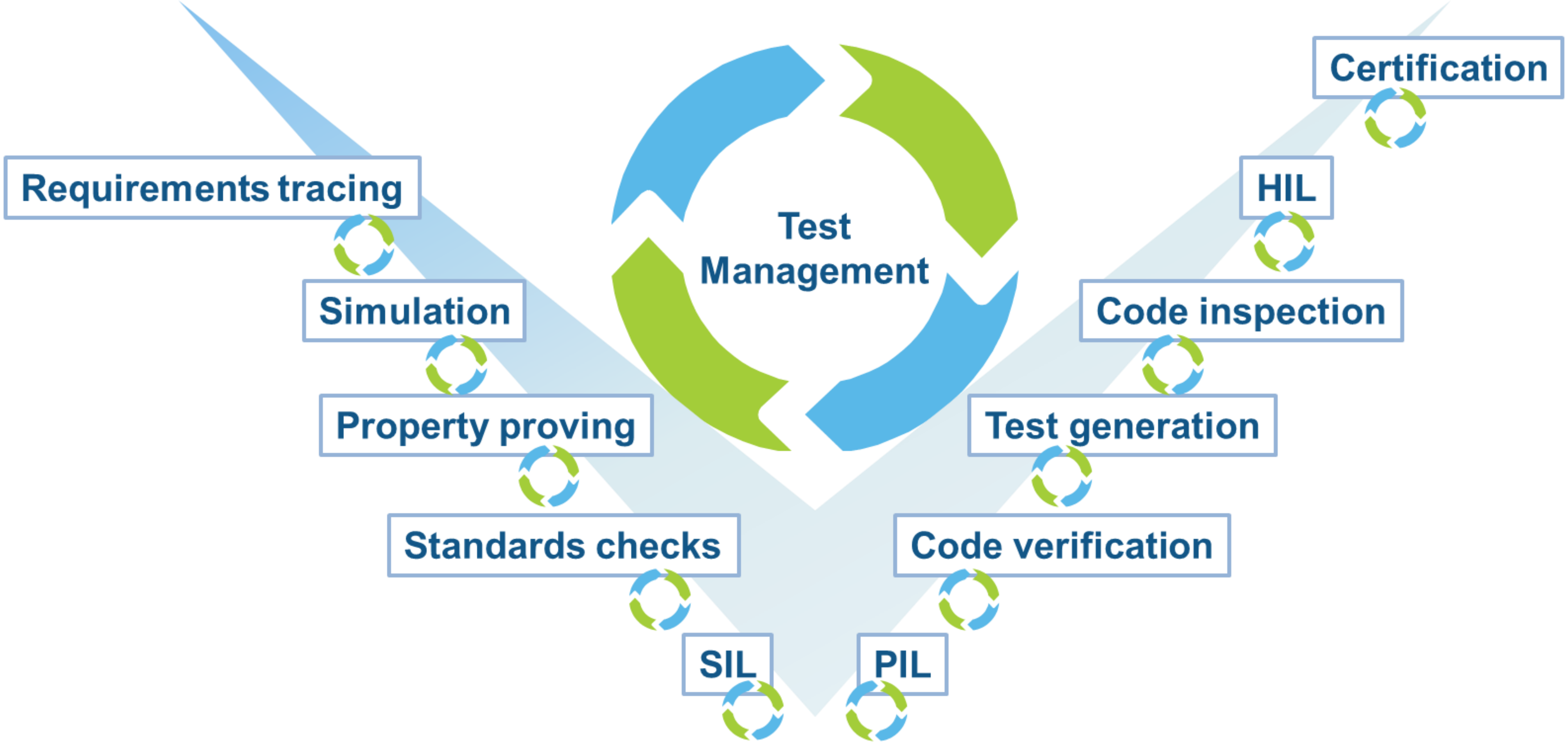
Safety Standards Updates:
Very Early Phase

- TUV SUD
  - Open GENESIS
  - Started in May 2019
- SAE and EUROCAE
  - Joint Working Group WG-114
  - Kick-off in August 2019
- RTCA (Aerospace, US)
  - Still evaluating member's interests
- ISO JTC 1/SC 42
  - Standardization program on Artificial Intelligence
  - In "Preparatory" phase – work not yet started

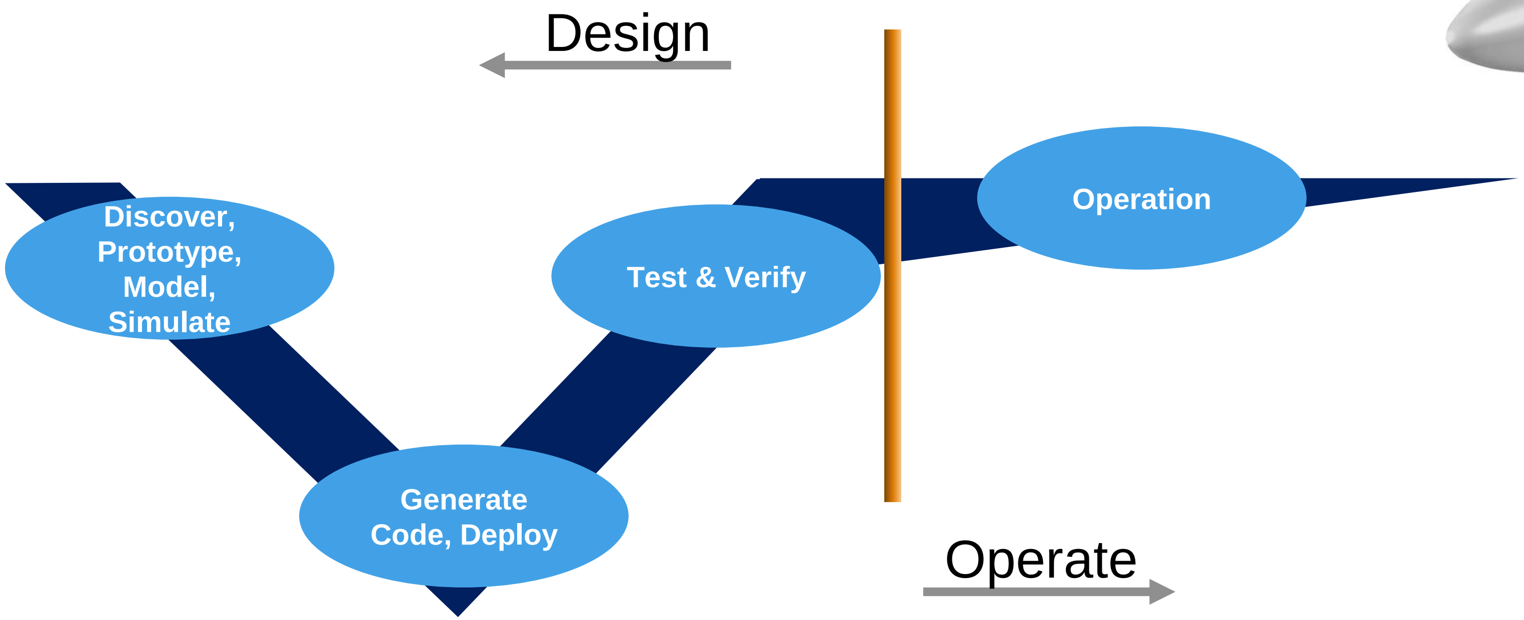
When designing physical products

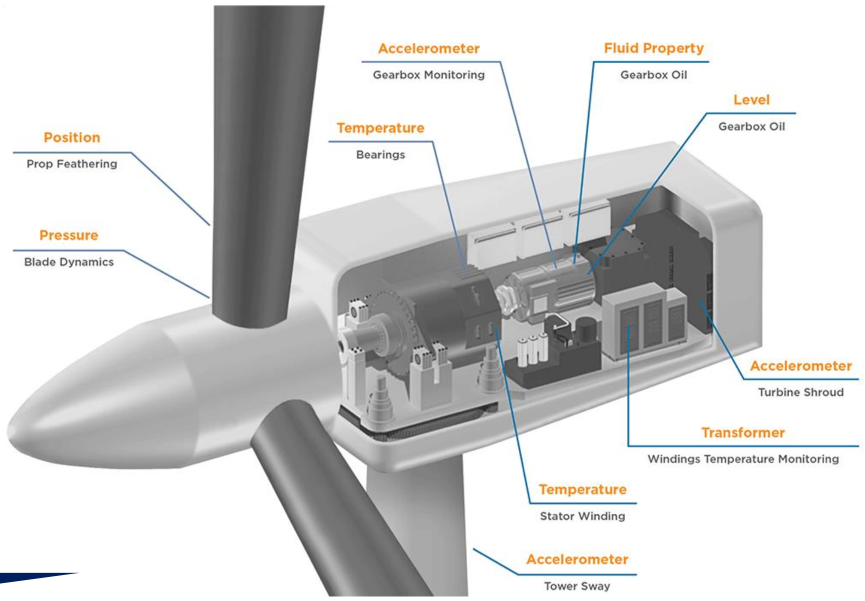


Much of the process may be regulated

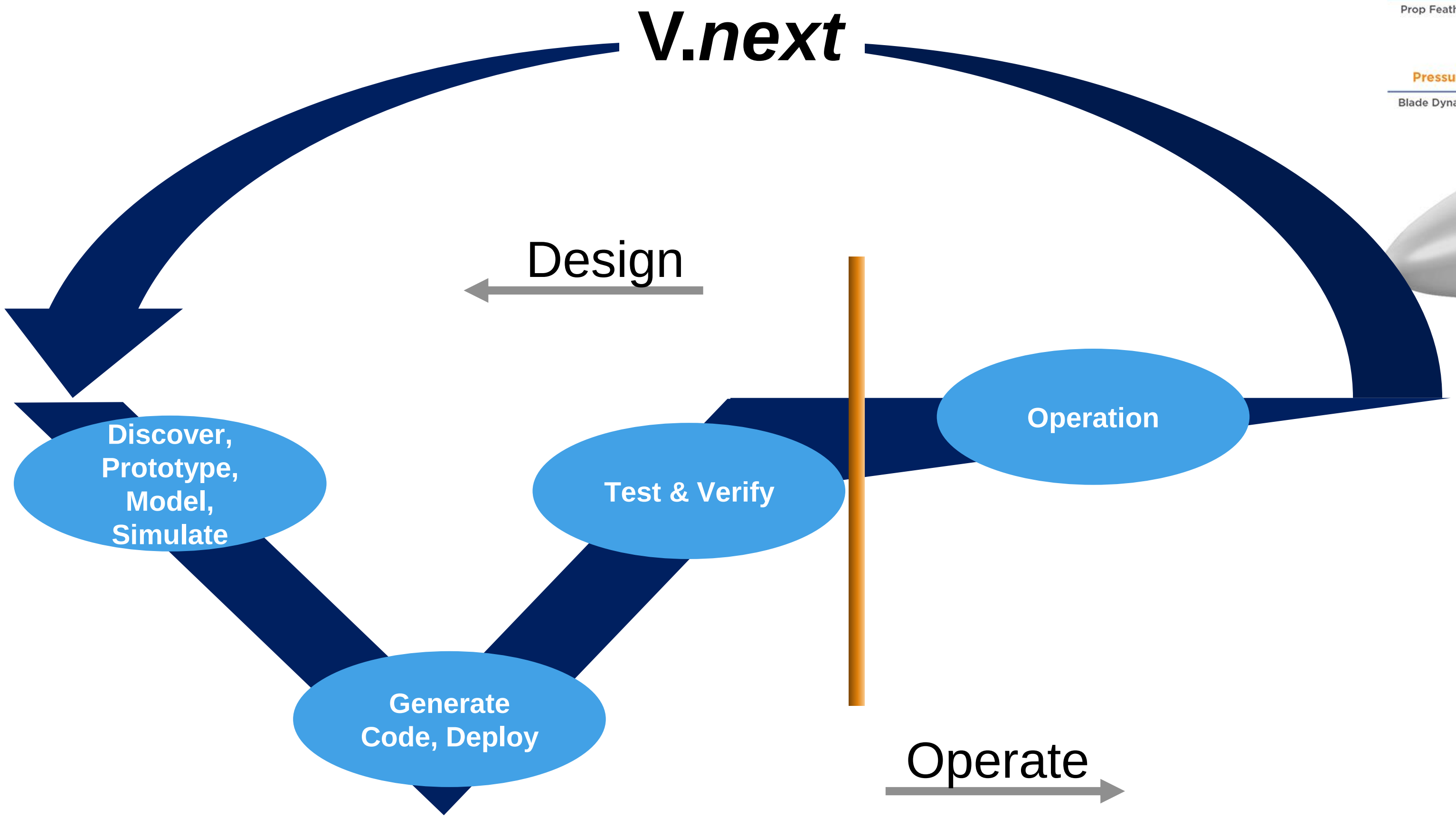


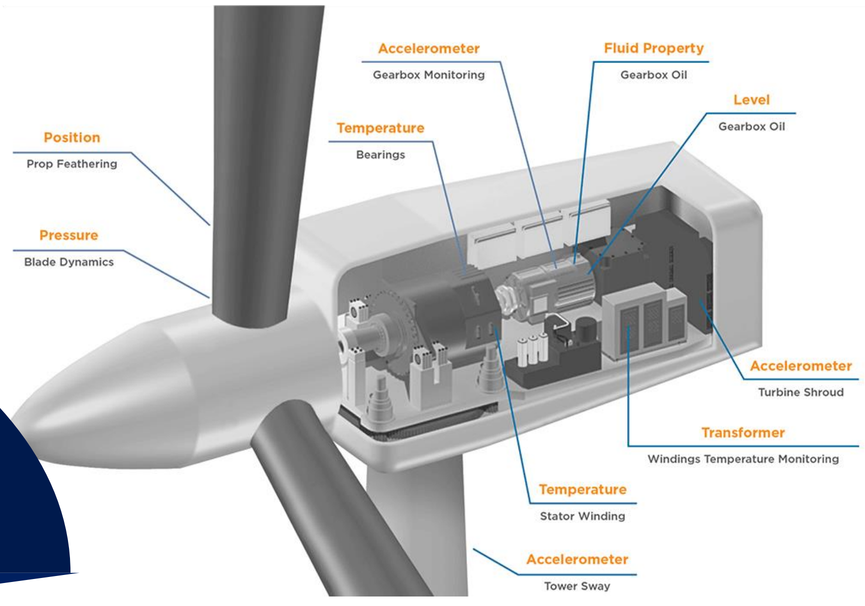
Connected Systems Means…





Positive improvements from field data





Model Risk Management – The Model is the Product

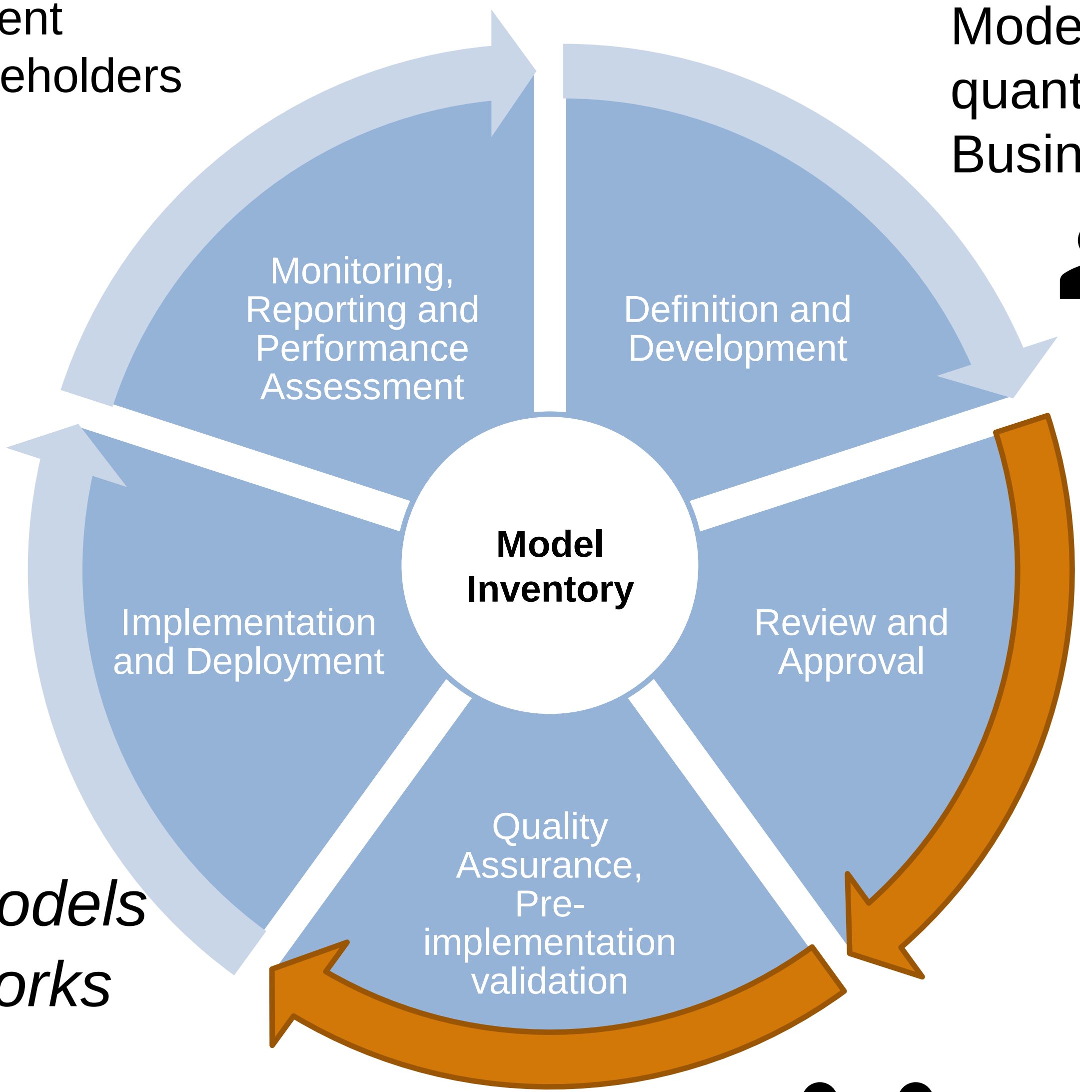
- Emphasis on
  - Process
  - People
  - Inventory of Assets
  - Execution Phase
  - Reporting
- Learn more
  - *A Platform for Risk Models*
    Paul Peeling, MathWorks

Risk management
Board and stakeholders
Regulator
Model Owners

IT, front-office
Model Users

FLOD

SLOD



Model developers, quants, analysts. Business lines

Independent model review and audit
Regulator

Model Validation, IT

“Explainability” More than SHAPley and LIME

- …and Partial Dependency Plots and … this is an active area of research
- The basics apply:
  - What is the quality and relevance of the data used to train the model?
  - Has the process to develop the model been recorded properly? (How was the data cleaned? What were the parameters used for training?)
  - How will the model be monitored in use?
- Anecdotally, customers have been able to explain models and methods sufficiently to allow use, when they have followed good practices.
  - Talk with our consultants if you need help with this

Four Learnings from Other Industries

1. Plenty of value away from the “obvious” applications
2. There’s no reason not to look for your keys under the street light:
   If you have data use it
3. Regulations can be tough – but perhaps not for advice.
4. If you don’t have data, can you create it?

Three Areas to Watch

1. If your application is performance dependent; Hardware Options
2. Reinforcement Learning is developing quickly, time to investigate?
3. AI regulations are here and coming; good practices are important