



Data Science Project from End to End: A Sucker-Rod Pump Example

Patrick Bangert
algorithmica technologies Inc.



SPE Webinars are Sponsored by:

SPE
SPE Foundation



Speaker



Dr. Patrick Bangert



3

Agenda



1. Digital Oilfield
2. Sucker-Rod Pumps
3. Machine Learning
4. Application in the Field
5. Results





Part 1: Digital Oilfield



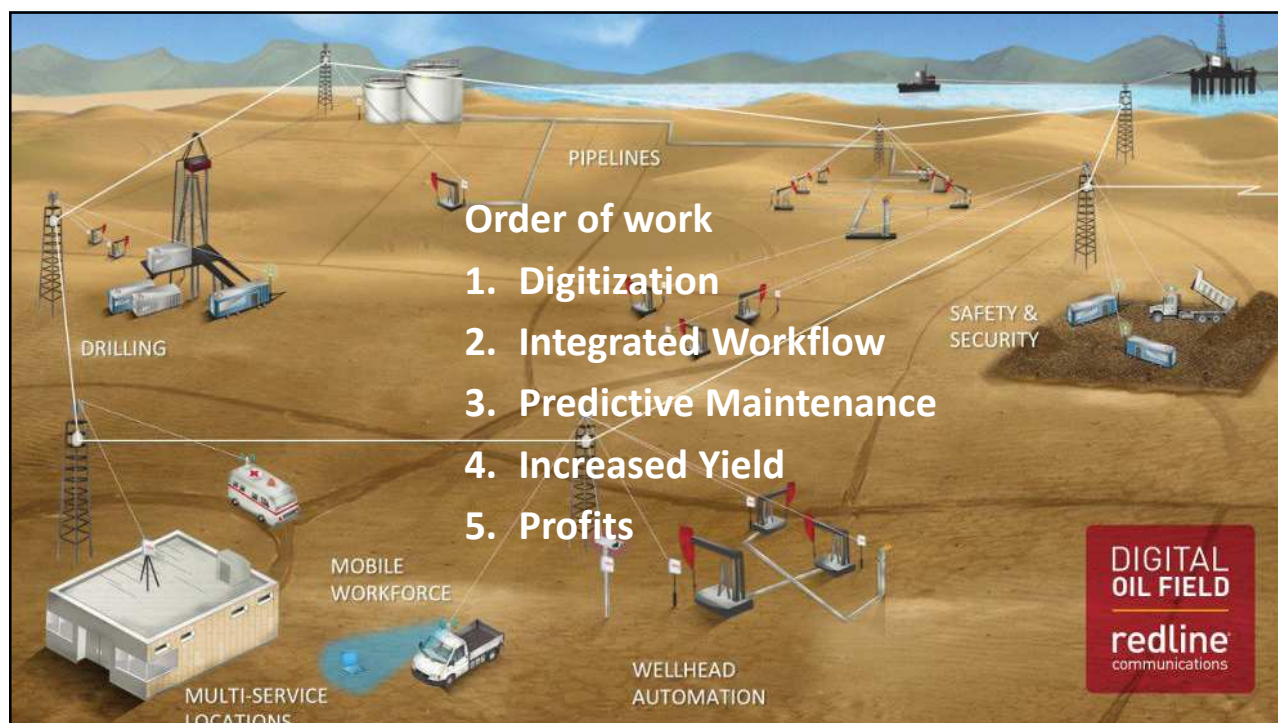
“

The purpose of the digital oilfield is to maximize **oilfield recovery**, eliminate **non-productive time**, and increase **profitability** through the design and deployment of **integrated workflows**.

”

PetroWiki, SPE





Digital Oilfield

Key differences to “normal” IoT

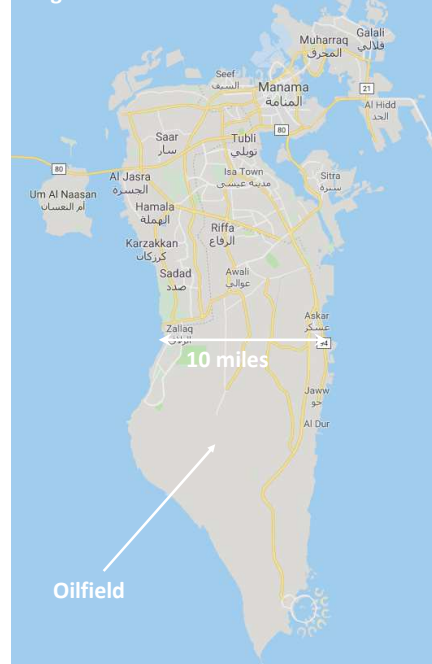
- Harsh environment
 - Expensive sensors
- Large area
 - Data transmission issues
- Inaccessible locations
 - Long response times

In Bahrain

- 100 square miles of oilfield
- Thousands of wells
- Dedicated electricity grid



Kingdom of Bahrain



IoT & Integrated Workflow



Measure & Transmit

- Sensors chosen and installed
- Analog to digital converters
- Data acquisition at well
- Digital data transmission to canopy
- Digital canopy for transmission to central server



Record & Visualize

- Establish database
- Meta-data
- Align timestamps
- Store data in database
- Dashboards
- Historical data access and plots

Interpret & Diagnose

- Look at data
- Is data reliable & significant?
- Does it indicate a problem?
- If so, which problem?
- Must be done regularly & systematically

Actionize & Plan

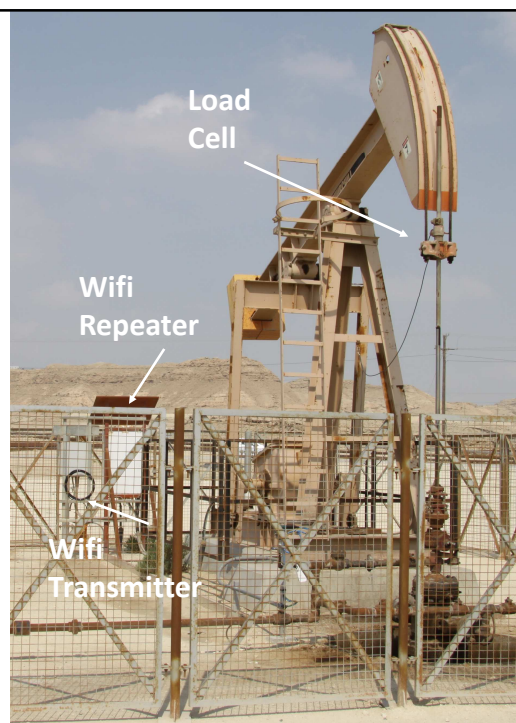
- Is interpretation actionable?
- If so, what action?
- Who should do it and when?
- What material, spare part, qualifications, etc. do they need?
- Enter in planning system
- Notify teams

Track & Trace

- What is the status?
- What was done?
- What is the result?
- Use result to learn & train for future
- Record resources used to track effort, cost, materials, parts

Measure & Transmit

- Focus on Sucker-Rod Pumps
- Oilfield is illuminated by wifi repeater network
- Each rod pump gets a wifi transmitter
- All measurements are collected at the transmitter and regularly sent
- For example, the load cell that measures the weight of the oil on the traveling valve is clearly seen in the picture



Record & Visualize



- Control system and historian
- Meta-data and diagrams
- Set-points transmitted back to the field



Record & Visualize



Exploration

- Data Combination
- Seismic Inversion
- Drill site selection
- Bidding on leases
- Reservoir Model

Drilling

- Variable Speed
- Layer Detection
- Maintenance
- Reduce Drill Time
- Safety

Completion

- Project Timetabling
- Labor Organization
- Logistics Optimization
- Risk Management
- Advanced Sourcing
- Training

Production

- Optimization
- Advanced Process Control
- Yield Forecast
- Slugging Prevention
- HSE
- Engineering Optimization

Maintenance

- Predictive
- Prescriptive
- Diagnostic
- Drone Surveillance
- Parts Orders



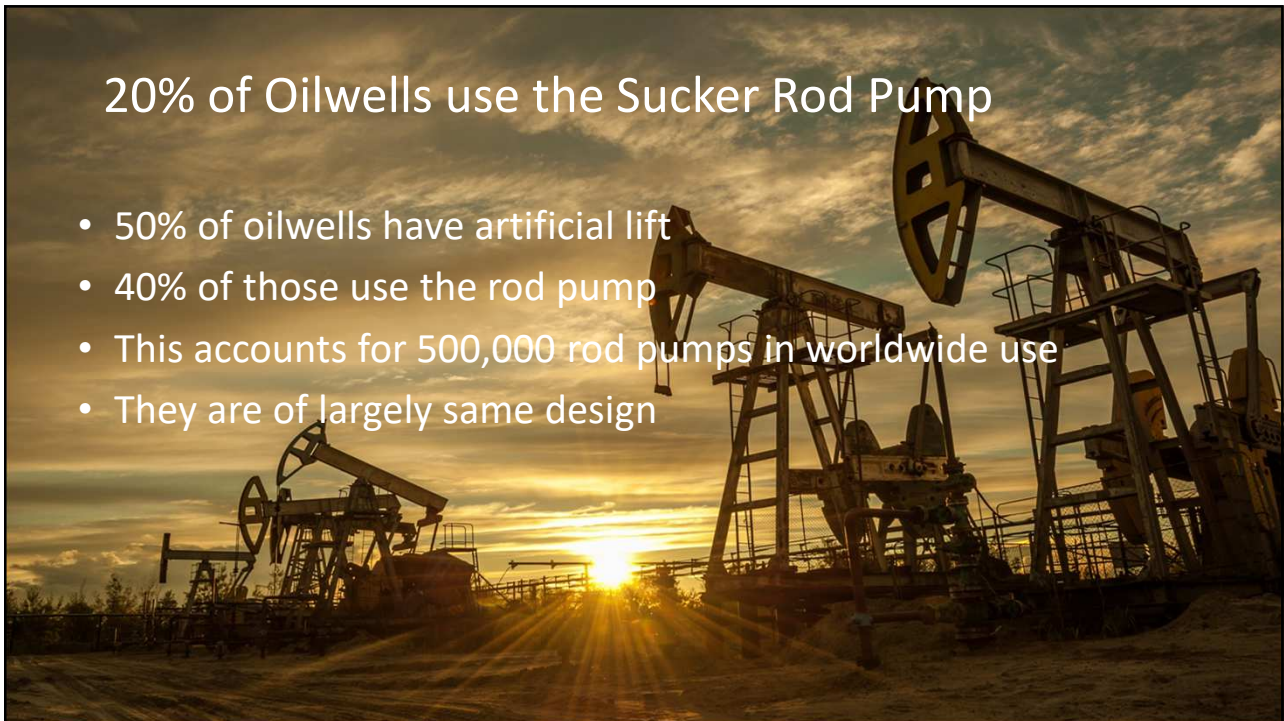


Part 2: Sucker-Rod Pumps

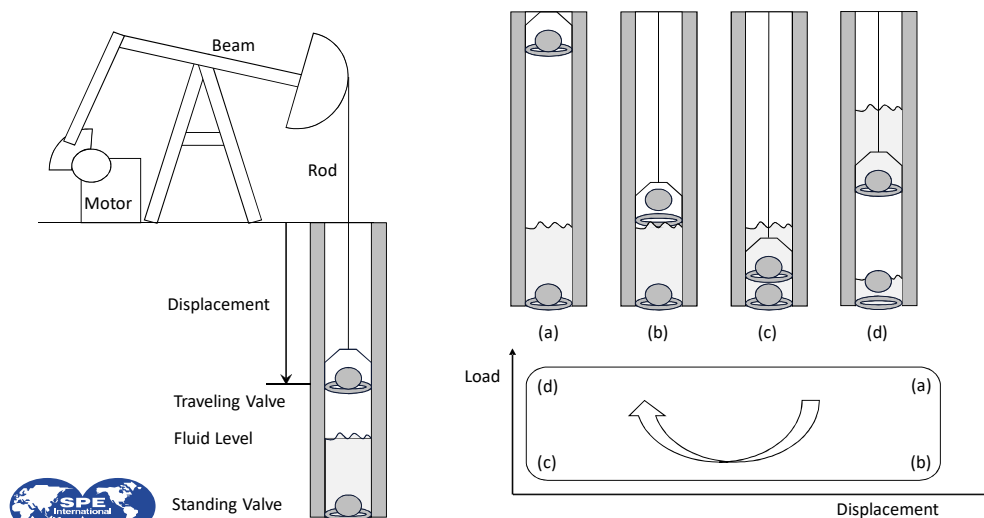


20% of Oilwells use the Sucker Rod Pump

- 50% of oilwells have artificial lift
- 40% of those use the rod pump
- This accounts for 500,000 rod pumps in worldwide use
- They are of largely same design



Rod Pumps



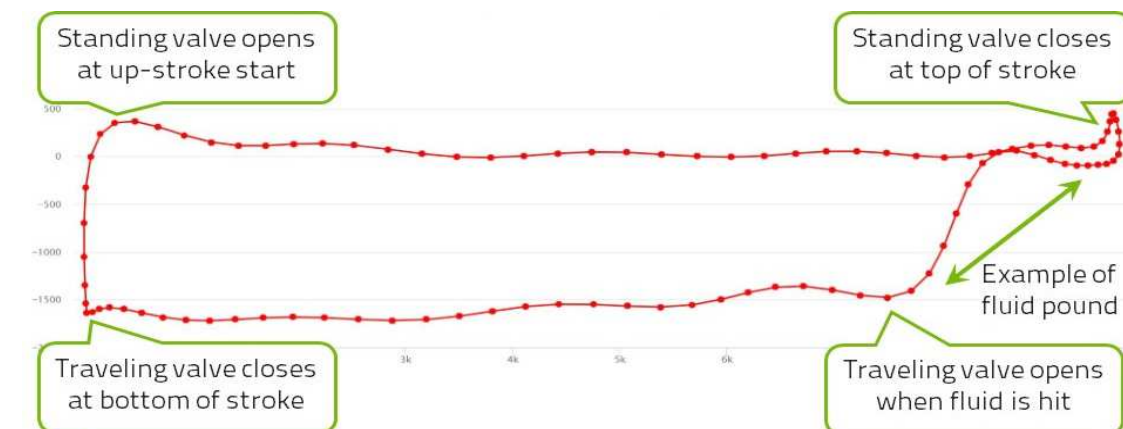
Damage Mechanisms of Rod Pumps



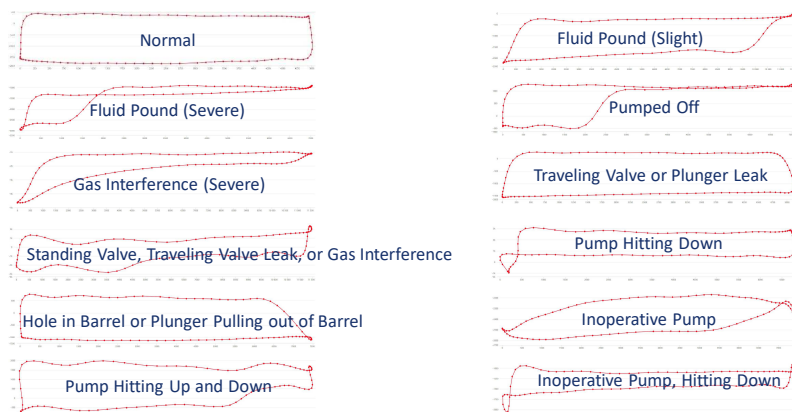
- Diverse damage mechanisms
- Can be diagnosed by experts
- Diagnosis can be automated with higher accuracy
- Works in real-time, several times per second on thousands of pumps
- Challenge: One pump produces about 250,000 cards per day



Dynamometer Card Shows the Damage



Different Cases have Unique Signatures

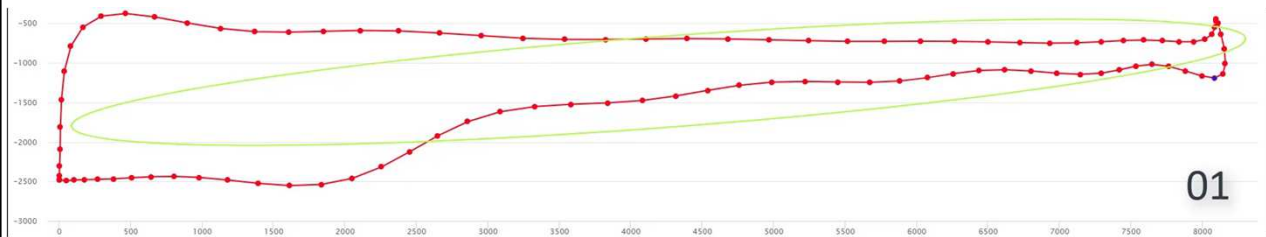


Fourier Series describe Dynamometer Cards



$$u(s) = x(s) + iy(s)$$

$$u(s) = \sum_{n=-M}^M a_n \exp\left(\frac{2\pi i n s}{S}\right) \quad a_n = \frac{1}{S} \int_0^S u(s) \exp\left(\frac{-2\pi i n s}{S}\right) ds$$



Part 3: Basic Ideas of Machine Learning



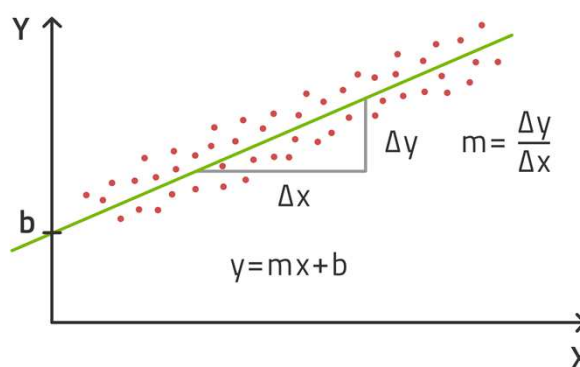


“ Machine Learning is about **prediction**:
Using information you have to generate
information you don't have. ”

Prediction Machines
by A. Agrawal, J. Gans, A. Goldfarb
Harvard Business Review Press

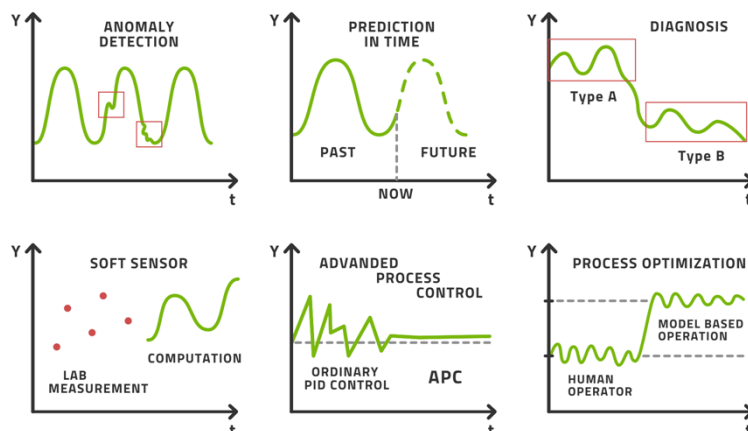


Machine Learning produces
a summary of data by
adjusting parameters in an
equation until that equation
fits the data in the best
possible way.

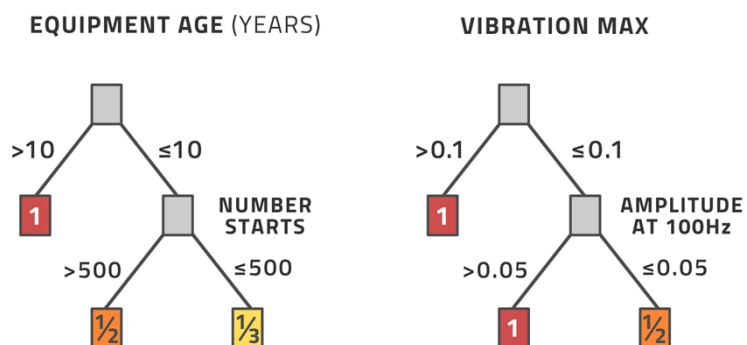




Different model types provide different kinds of output targeted to particular use cases in the application

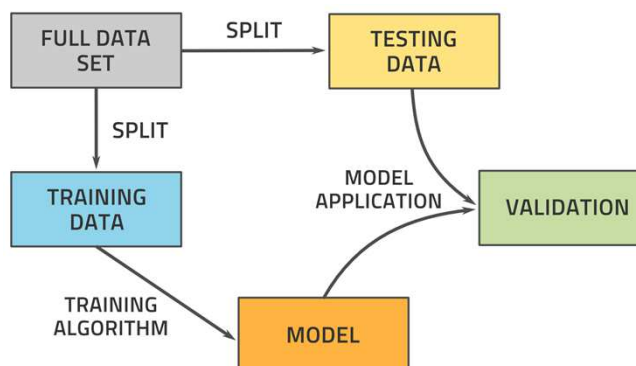


Decision trees or random forests have proven to be the most successful classification models.





Data of known cases must be split into training and testing data sets so that we can verify the model on novel data.



Machine Learning Results



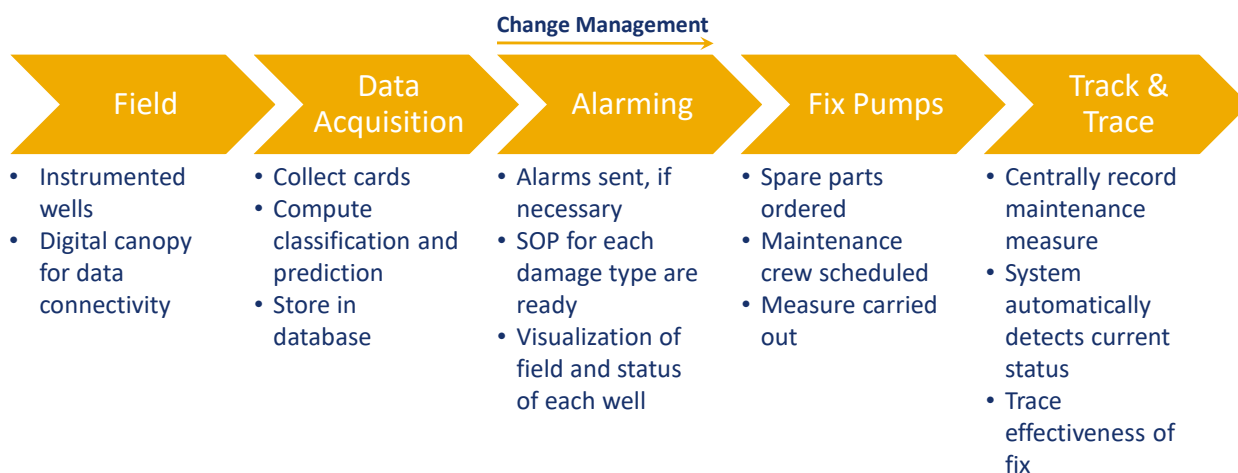
Category	Training Samples	Testing Samples	Testing Errors
Normal	8557	1529	0
Fluid Pound (Slight)	5347	955	0
Fluid Pound (Severe)	93	15	0
Inoperative Pump	1981	379	0
Pump Hitting Down	1740	303	2
Pump Hitting Up and Down	2258	407	2
Inoperative Pump, Hitting Down	9045	1626	1
Traveling Valve or Plunger Leak	98	15	1
Standing Valve, Traveling Valve Leak, or Gas Interference	345	62	0
Pumped Off	234	39	1
Hole in Barrel or Plunger Pulling out of Barrel	132	20	0
Gas Interference (Severe)	101	11	0
Total	29931	5361	7



Part 4: Application in the Field



Full Workflow



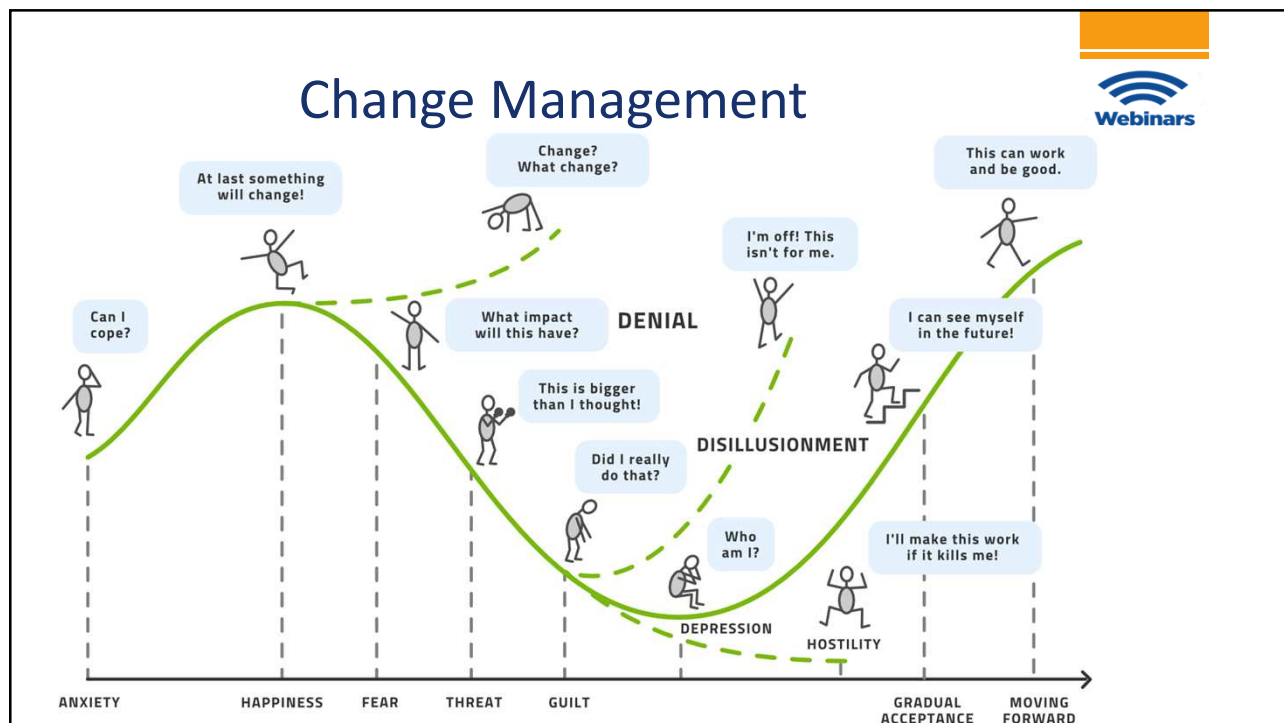
Technical Challenges



- Data interfaces
- IT Regulations
- Data security
- Data interpretation
- Card labeling
- Result validation



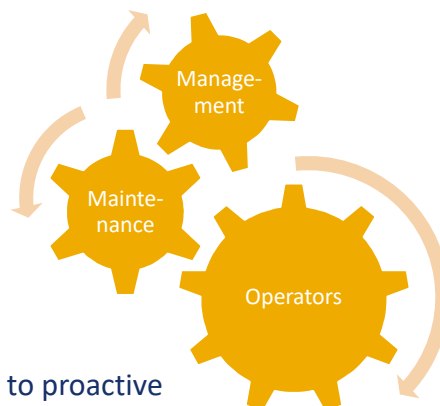
Change Management



Change Management



- Operators
 - accept the quality of the predictions
 - plan the actions
- Maintenance
 - carries out the actions
- Management
 - tracks the benefits
- All stakeholders
 - participate in the journey
 - ultimately change their behavior from reactive to proactive
 - agree to make these changes and see the benefits



Part 5: Results



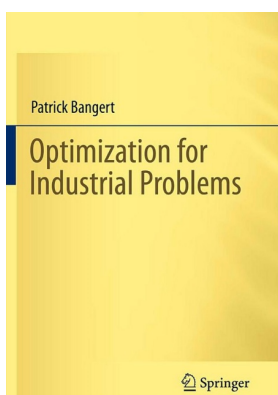
Lessons Learned



- Investment
 - In sensors is significant
 - In software is small
 - Software creates ROI
- People
 - Everyone must support and agree
 - Change management is important
- Integration
 - Effort is spent integrating IT systems
 - Departments must be integrated in expectations
- Effort
 - Mostly integration/organization
 - Some domain knowledge
 - Little data science
- Conclusions
 - There is much more to a data science project than data science!
 - It is not necessary for operators to have dedicated data science software developers – rather knowledgeable persons to organize, and manage the project and domain knowledge



Thank you for your attention!



Patrick Bangert

Office

algorithmica technologies Inc.
Cupertino, CA 95014, USA

Contact

p.bangert@algorithmica-technologies.com
+1 408 707 0956
www.algorithmica-technologies.com

