



Data Quality and It's Impact On Analytics

David Johnson
VP Innovation & Marketing
Petrolink
Sept 14th, 2016








David C. Johnson
Vice-President of
Research and
Innovation, Petrolink





Poll: Who are you?

What is your job title (or role) within your organization?



Everyday Analytics

- Objective:
 - Houston to Orlando

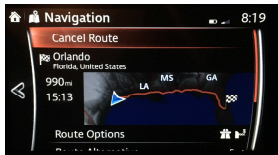
Everyday Analytics

- Search / Closest Match
 - Orlando, FL or Orland, CA

Everyday Analytics

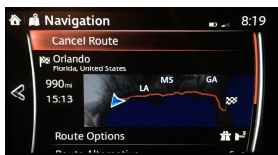
- Search / Closest Match
 - Orlando, FL or Orland, CA
- Route Planning
 - Fastest, Direct, Tolls

Everyday Analytics



- Search / Closest Match
 - Orlando, FL or Orland, CA
- Route Planning
 - Fastest, Direct, Tolls
- Prediction
 - Analysis of distance
 - Estimation of time

Everyday Analytics



- Search / Closest Match
 - Orlando, FL or Orland, CA
- Route Planning
 - Fastest, Direct, Tolls
- Prediction
 - Analysis of distance
 - Estimation of time
- KPIs (Waypoints)
 - Turns, Roads, Bridges

Everyday Analytics



What could go wrong?



Everyday Analytics



What could go wrong?

- Recalculating....

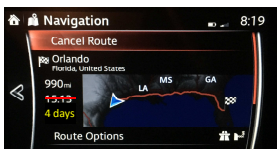
Everyday Analytics



What could go wrong?

- Recalculating....
- Unknown Hazards

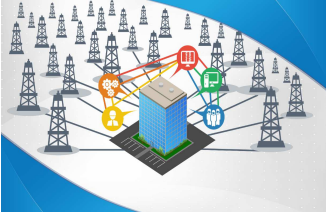
Everyday Analytics



What could go wrong?

- Recalculating....
- Unknown Hazards
- Traffic
- Car troubles
- "Are we there yet"

Today's complex Drilling & Operations environment

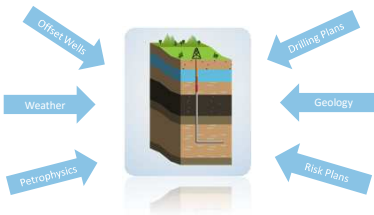


Operating Companies today:

- Scale
 - 10-100's of wells in construction
 - 100's of wells in production
 - 100's – 1000's of users
- Complexity
 - Multiple service providers
 - Many technologies
 - Incompatible formats
 - Aging experience pool
 - Data ownership
 - Financial pressures

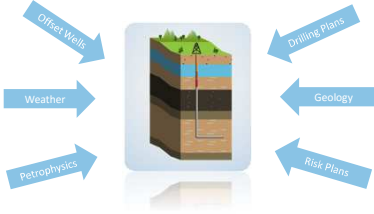



Analytics Can Help





Analytics Can Help



But...




Can you TRUST your data?



Can you TRUST your data?

- Confidence in decisions **starts** with confidence in data



Can you TRUST your data?

- Confidence in decisions **starts** with confidence in data



Poll: Your Concerns



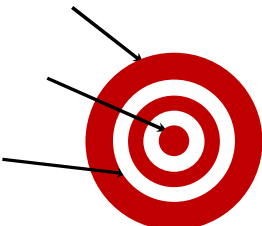
What is the biggest worry you have about Analytics and YOUR data?



Common Data Problems



- Accuracy
 - Subjective
 - Assessed in RT/pRT

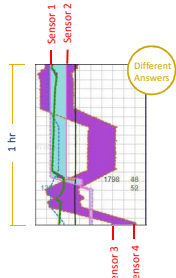



Accuracy – Sensor issues



4 types of invalid data

- Multiple values




Accuracy – Sensor issues

4 types of invalid data

1. Multiple values
2. Index issues
3. Amplitude shift

Accuracy – Sensor issues

4 types of invalid data

1. Multiple values
2. Index issues
3. Amplitude shift
4. Calibration

Impact on Analytics: MSE

$$MSE = \frac{WOB}{(P_{IPD})} + \frac{2\pi * RPM * Torq}{(P_{IPD}) * ROP}$$

What if

- A sensor off by 10%?

Warning – not intended to be a class on drilling best practices

Impact on Analytics: MSE



$$MSE = \frac{WOB}{(P_{PA})} + \frac{2\pi * RPM * Torque}{(P_{PD}) * ROP}$$

What if

- A sensor off by 10%?

	WOB	RPM	Torque	ROP	MSE	Impact %
Baseline	18	60	16	25	22342.88109	
WOB	19.8	60	16	25	22343.14776	0.0012%
RPM	18	66	16	25	24576.90253	9.9987%
Torque**	18	60	17.6	25	24576.90253	9.0899%
ROP	18	60	16	27.5	24576.90253	9.0899%



Warning - not intended to be a class on drilling best practices

Impact on Analytics: MSE



$$MSE = \frac{WOB}{(P_{PA})} + \frac{2\pi * RPM * Torque}{(P_{PD}) * ROP}$$

What if

- A sensor off by 10%
- ALL off by 10%

	WOB	RPM	Torque	ROP	MSE	Impact %
Baseline	18	60	16	25	22342.88109	
all of them	19.8	66	17.6	27.5	29737.75873	33.0972%



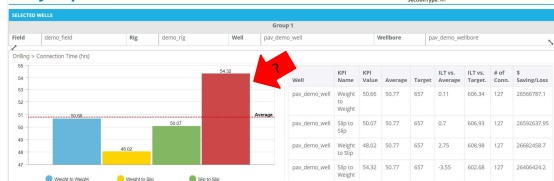
Warning - not intended to be a class on drilling best practices

Impact on Efficiency: NPT & \$\$



Daily Report

New: 10/14/16 12:00:00-10/14/16 12:00:00
Depth: 100 - 100
Location: 100 - 100



Impact on Predictive Drilling



- False Alarms WORSE than no alarm!



Common Data Problems



- Accuracy
 - Subjective
 - Assessed in RT/pRT
- Access
 - Formats
 - Units
 - Security



Integration and Interoperability



- Challenges:
 - Formats
 - Time-frame
 - Sources
 - Access
- Require a balance of Security / Openness
- The role for Standards:
 - Interoperability through common, consistent data structures
 - WITSML™, RESQML™, PRODML + ETP provide high-speed access in real-time or on-demand



Common Data Problems



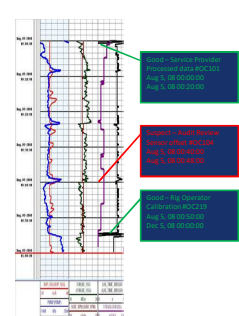
- Accuracy
 - Subjective
 - Assessed in RT/pRT
- Access
 - Formats
 - Units
 - Security
- Confidence
 - Data Lineage
 - Meta data
 - Processed/Raw




What REALLY can we do?




Ideas: Data Assurance Object



- Decorator objects
- Provide Lineage
- Process to assess data validity/conformity
- Applies to ALL objects
- Applies to ALL ML's




Ideas: Data Assessment Report



In Real-time:

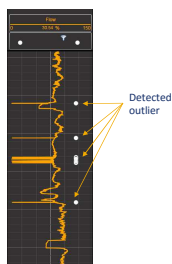
- Based on corporate objectives
- Review dimensions of "quality"
- Assess confidence
- Record and report on "fit for purpose"



Ideas: Soft Calibration



- Rule based methods
 - Knowledge + Heuristics
- Estimation methods
 - Correlation and prediction
- Time Series analysis
 - Model sensors and compare
- Learning-based methods
 - Model & train, use statistics
- Others...



Conclusions



- Standards industry is working on this
GET INVOLVED
- Analytics helps, check your data
- Quality is everyone's concerns



Questions?



David Johnson
david.Johnson@Petrolink.com

* Thanks to:

- Ali Moussadek, Sridharan Chandrasekaran, Mark Farnan, and HOST of other people
- Bill Chmela our MODERATOR
- SPE for hosting this
- YOU!