Data Management for Analytics

Technology Week 2020

24 February 2020

SAFER, SMARTER, GREENER

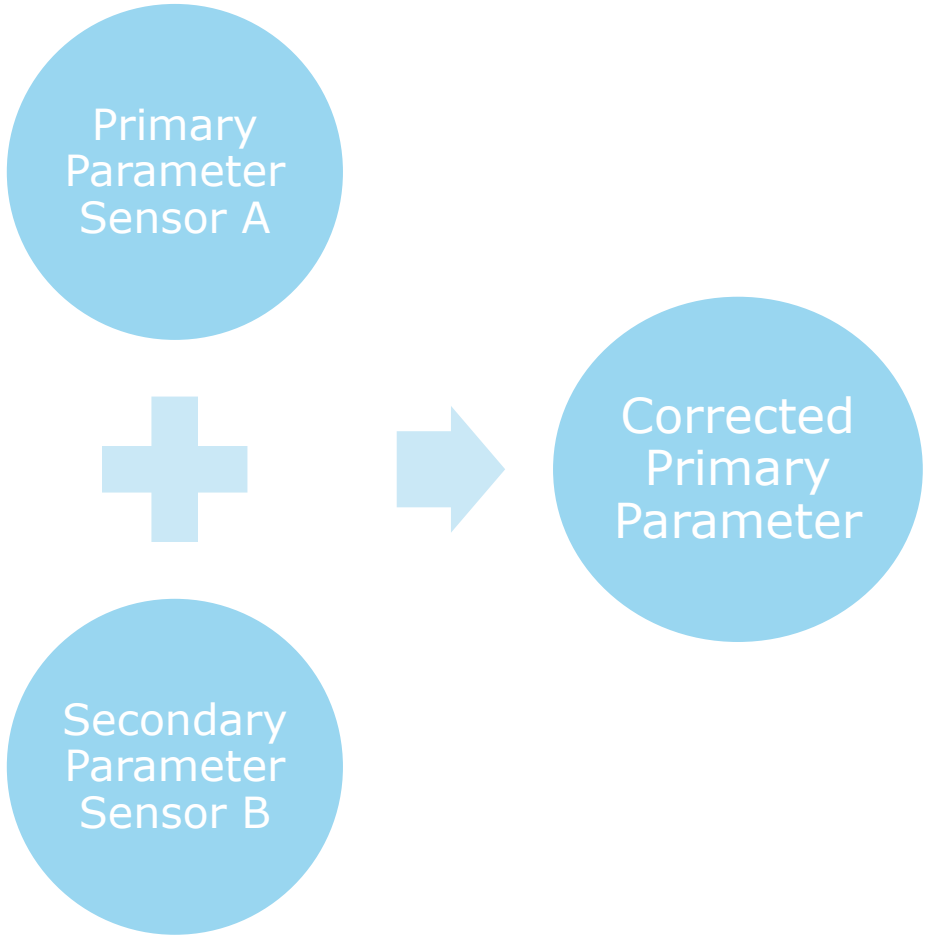
Outline

REAL WORLD EXAMPLES

DNVGL RP FOR DATA QUALITY ASSESSMENT

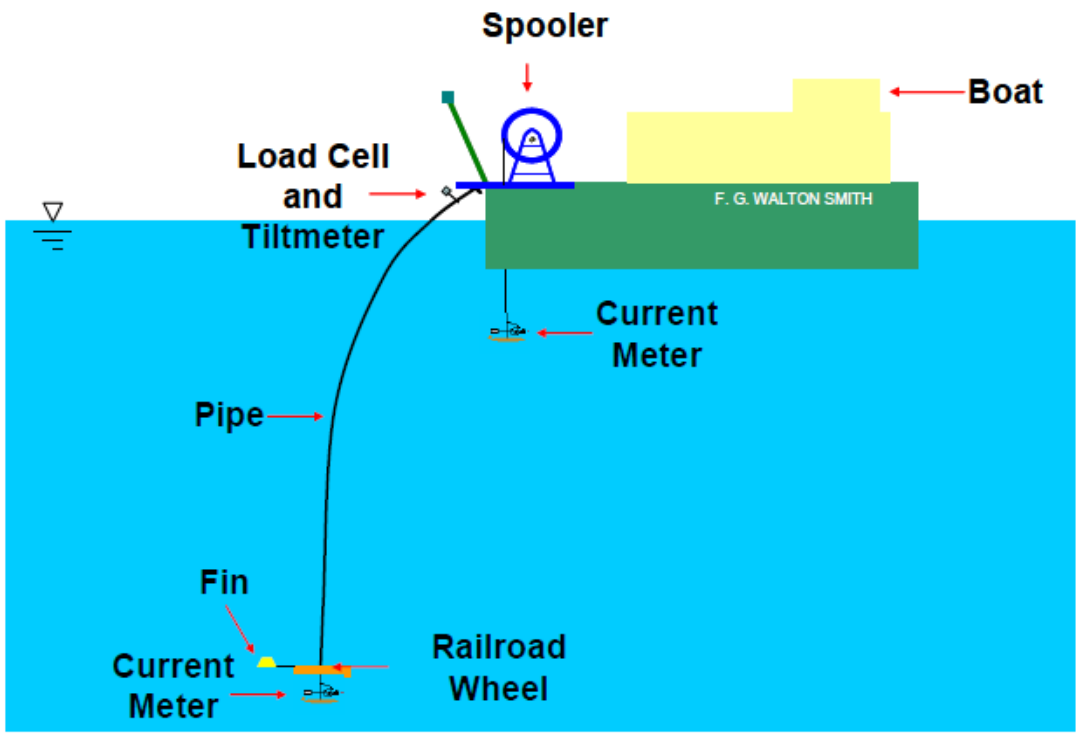
Example 1: Smart Mooring™ Monitoring System

- Objective: Develop mooring line tension and failure detection system using machine learning (ML) model and field measured data.
- Data from multiple sensors used in ML model.
- Large data set collected over 1-year period.
- Failure of sensor measuring a secondary parameter caused errors in primary parameter.
- Erroneous sensor readings identified by data validation process.

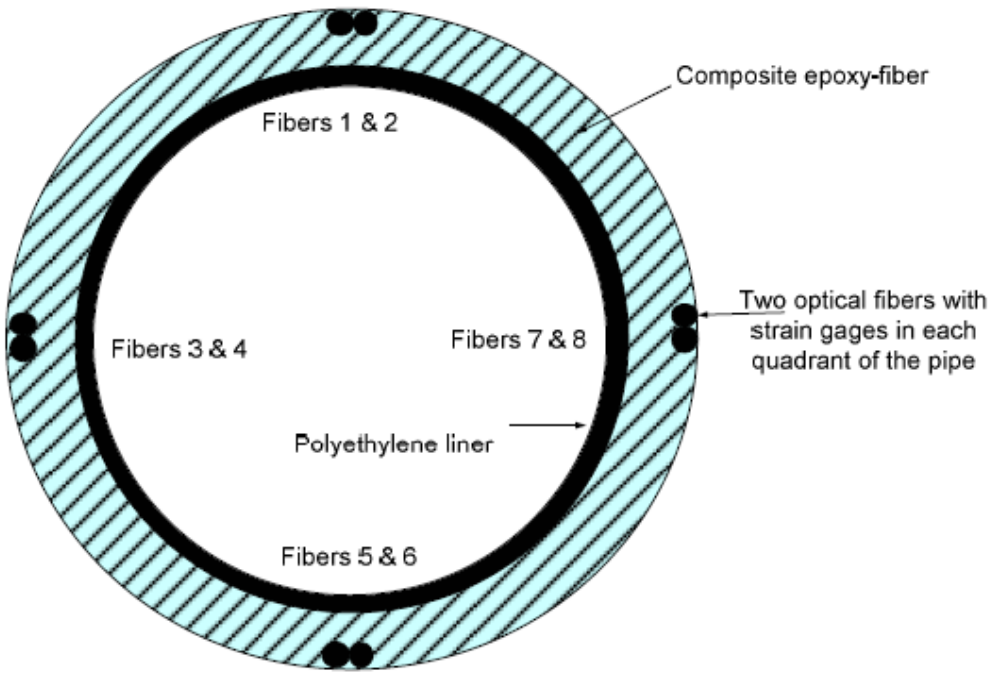


Example 2: Offshore Field Experiments

Sensors installed incorrectly and damaged during experiments



Experiment: Tow an instrumented pipe in ocean



Pipe cross-section

Images source: OMAE2007-29353, Jaiswal et al

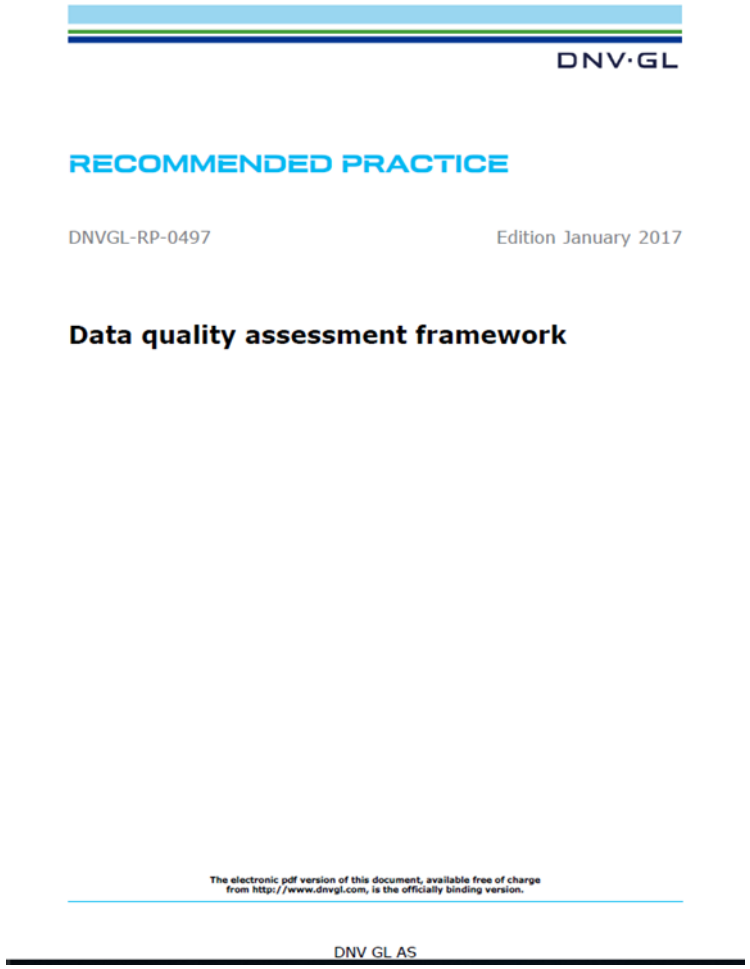
Example 3: Democratic Iowa Caucus 2020

Objective: New mobile app launched with the objective of speeding up the results tabulation and reporting

Outcome: Results were delayed; confusion

What went wrong: Mobile app had coding bugs resulting in inconsistent inputs and outputs; user interface was difficult to use; backup systems (phone lines) were overwhelmed due to understaffing.

DNVGL Recommended Practice - Data Quality Assessment Framework



DNV·GL

RECOMMENDED PRACTICE

DNVGL-RP-0497 Edition January 2017

Data quality assessment framework

The electronic pdf version of this document, available free of charge from http://www.dnvgl.com, is the officially binding version.

DNV GL AS

Data Quality Assessment

Organizational Maturity Assessment

Data Quality Risk Assessment

Data Quality Definition and Measurement Categories

Syntactic – does data conform to specified syntax?

Semantic – does the data represent reality?

Pragmatic – is the data quality appropriate for use case?

Organizational Maturity Assessment

Level 1 Initial

Management Perspective

Data is managed as a requirement for the implementation of projects

Pragmatic data quality

Data quality is adequate for local and dedicated usage or in data silos, but data reuse is difficult, the quality is not documented, and data analytics and automation are seldom very useful.

Level 5 Optimized

Management Perspective

Data is considered critical for survival in a dynamic and competitive market.

Pragmatic data quality

Data quality is used to brand the company as a preferred partner. Creation of new knowledge is based on analytics, with high trust in analytical results. Autonomous systems, with high trust and low risk, utilize the data.

Data Quality Process Maturity Assessment Framework

Image Source: DNVGL-RP-0497

Data Quality Risk Assessment

Risk management process

- define scope & scales for risk assessment, describing data usage and the datasets to be used
- perform risk identification
- rank risk according to criticality of consequences
- identify, evaluate, and prioritize responses and mitigating actions
- monitor risk scope, changes, incidents, and assessment process continuously
- review to improve the efficacy of the process and repeat based on needs or review criteria

Summary

Successful data analytics is underpinned by a robust data management strategy!

DNVGL-RP-0497 provides a framework for data quality assessment

Thank you!

Vivek Jaiswal
Vivek.Jaiswal@dnvgl.com
+1-281-396-1932

www.dnvgl.com

SAFER, SMARTER, GREENER

The trademarks DNV GL®, DNV®, the Horizon Graphic and Det Norske Veritas®
are the properties of companies in the Det Norske Veritas group. All rights reserved.