From Data to Intelligence

Steve Smith

9th January 2018

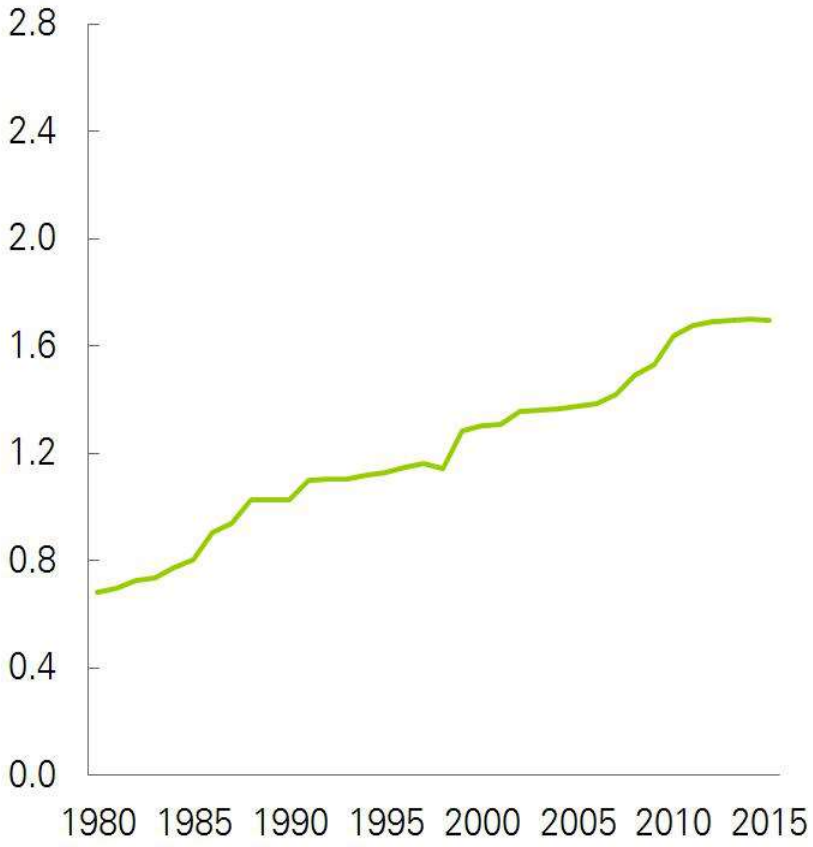
This document is

BP Energy Outlook
2017 edition

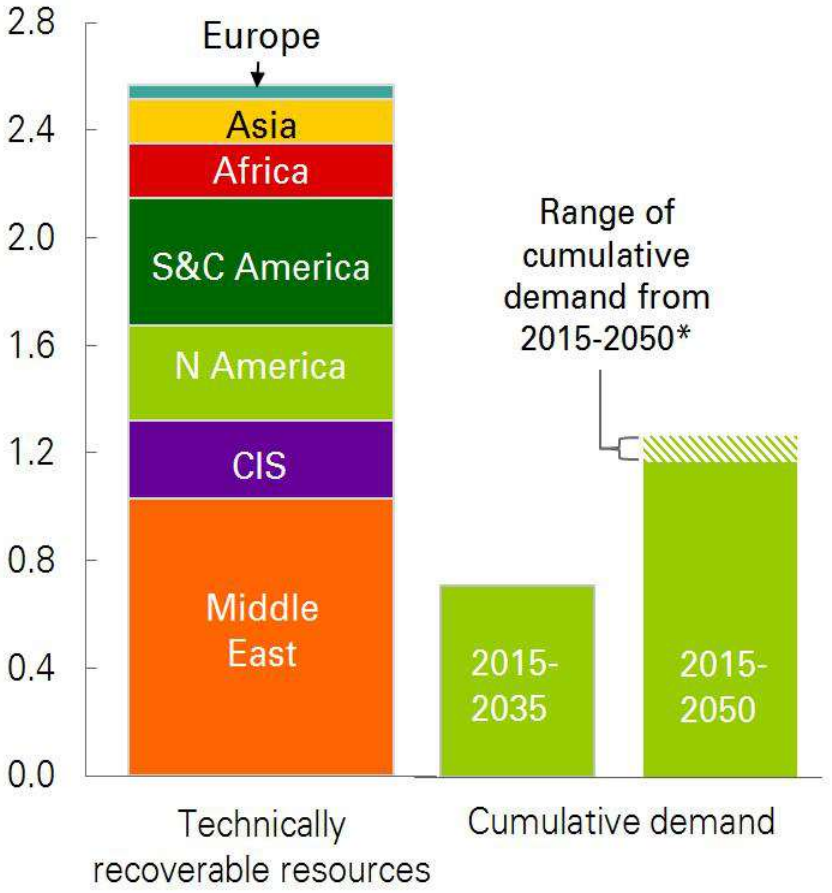
bp.com/energyoutlook
#BPstats

There is an abundance of oil resources...

Global proved oil reserves

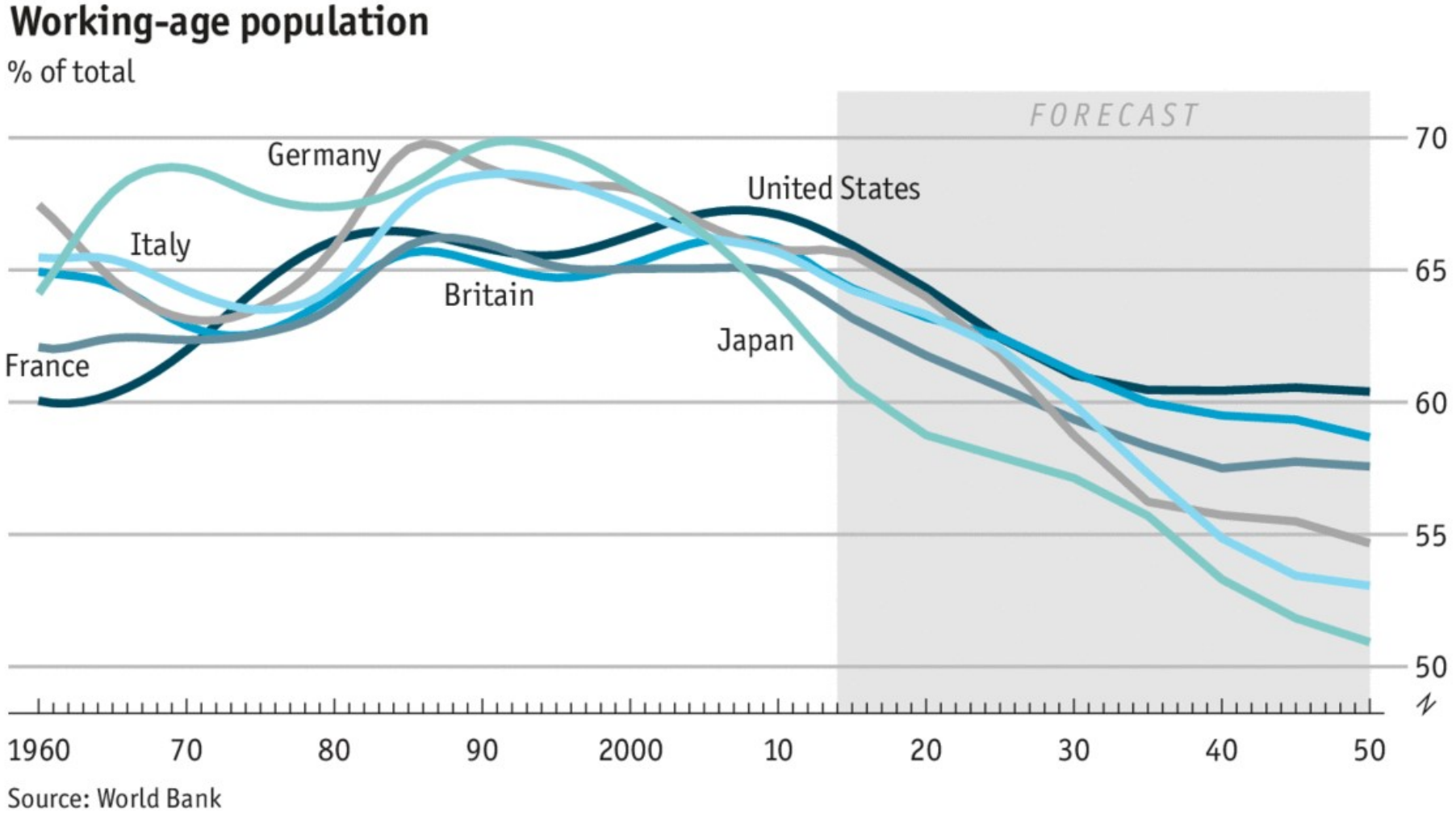


Estimates of technically recoverable resources and cumulative oil demand



*Based on range of outcomes shown on page 88

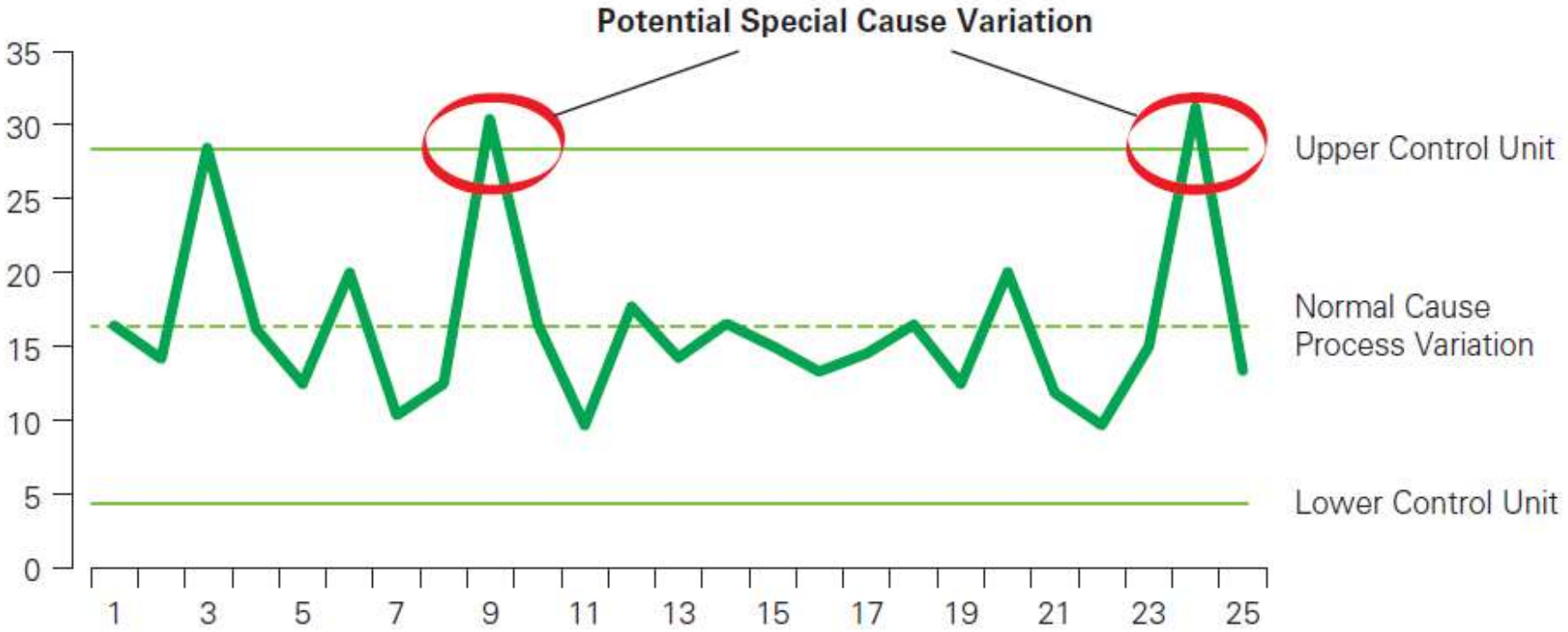
All this with less people!



Economist.com/graphicdetail

Our Objective:

To get better at spotting problems



Stakeholders

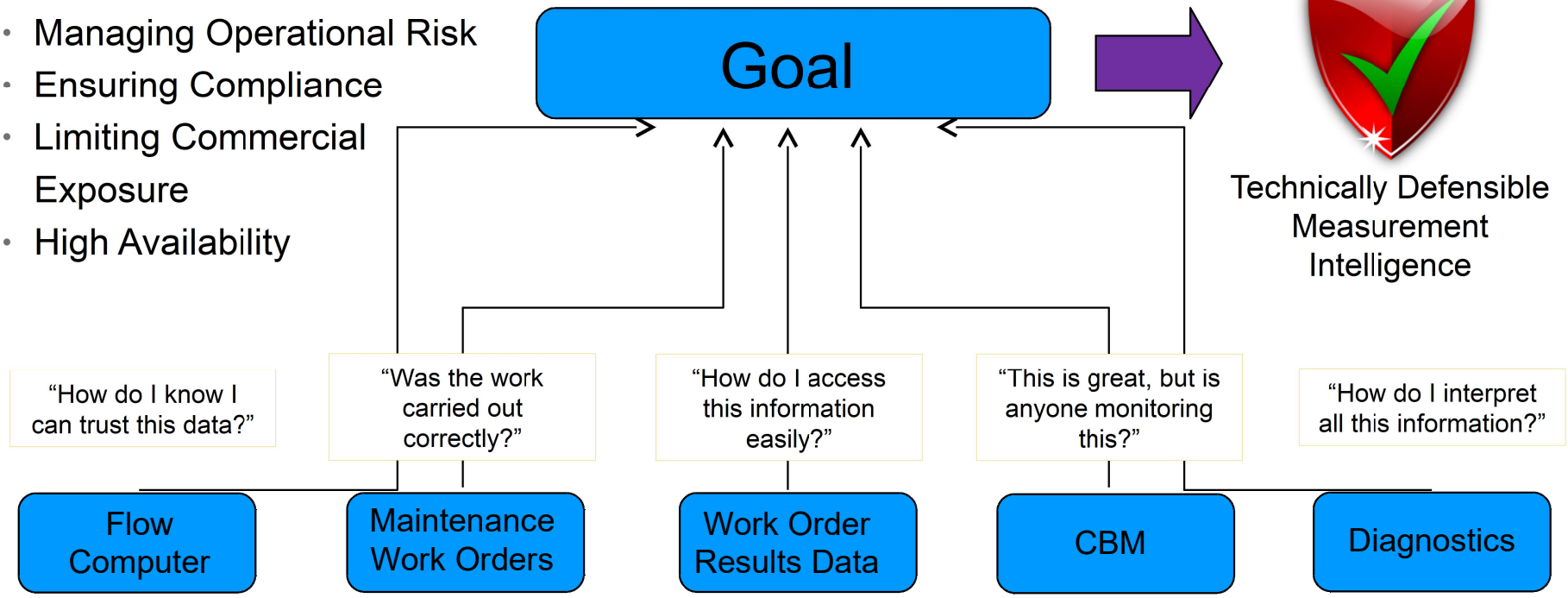
Safety

Key Metrics


Safety
Licence to Operate
Technically Defensible
End to End Mass Balance
Manage Uncertainty
Meet Production Targets
Risk Reduction
Control Reserve Inventory
Traceability
Compliance with Standards
Plant Efficiency
Contractual Obligations

Measurement Scope

- Managing Operational Risk
- Ensuring Compliance
- Limiting Commercial Exposure
- High Availability



- Daily Reports
- Monthly Reports
- System Status

- WO's Comp.
- WO's Outstanding
- Pass/Fail Status
- Pending WO's

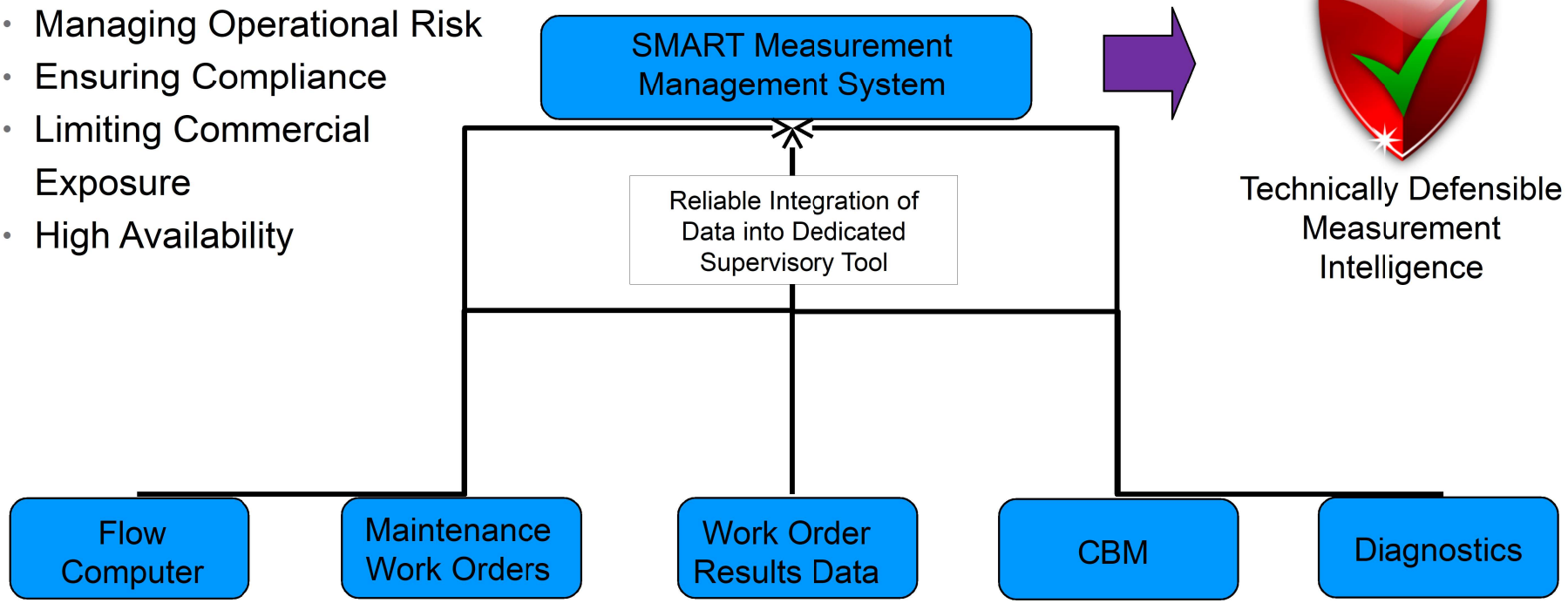
- Pass/Fail Criteria
- Test Results
- AF/AL Results
- Auditable Record System
- Results Trending
- Doc Control Log

- RAG Status
- Real time Degradation / Failure Alerts
- Troubleshooting tools

- System data validation
- Process stability
- Stream comparisons

SMART Measurement Ideal State:

- Managing Operational Risk
- Ensuring Compliance
- Limiting Commercial Exposure
- High Availability



- Daily Reports
- Monthly Reports
- System Status

- WO's Comp.
- WO's Outstanding
- Pass/Fail Status
- Pending WO's

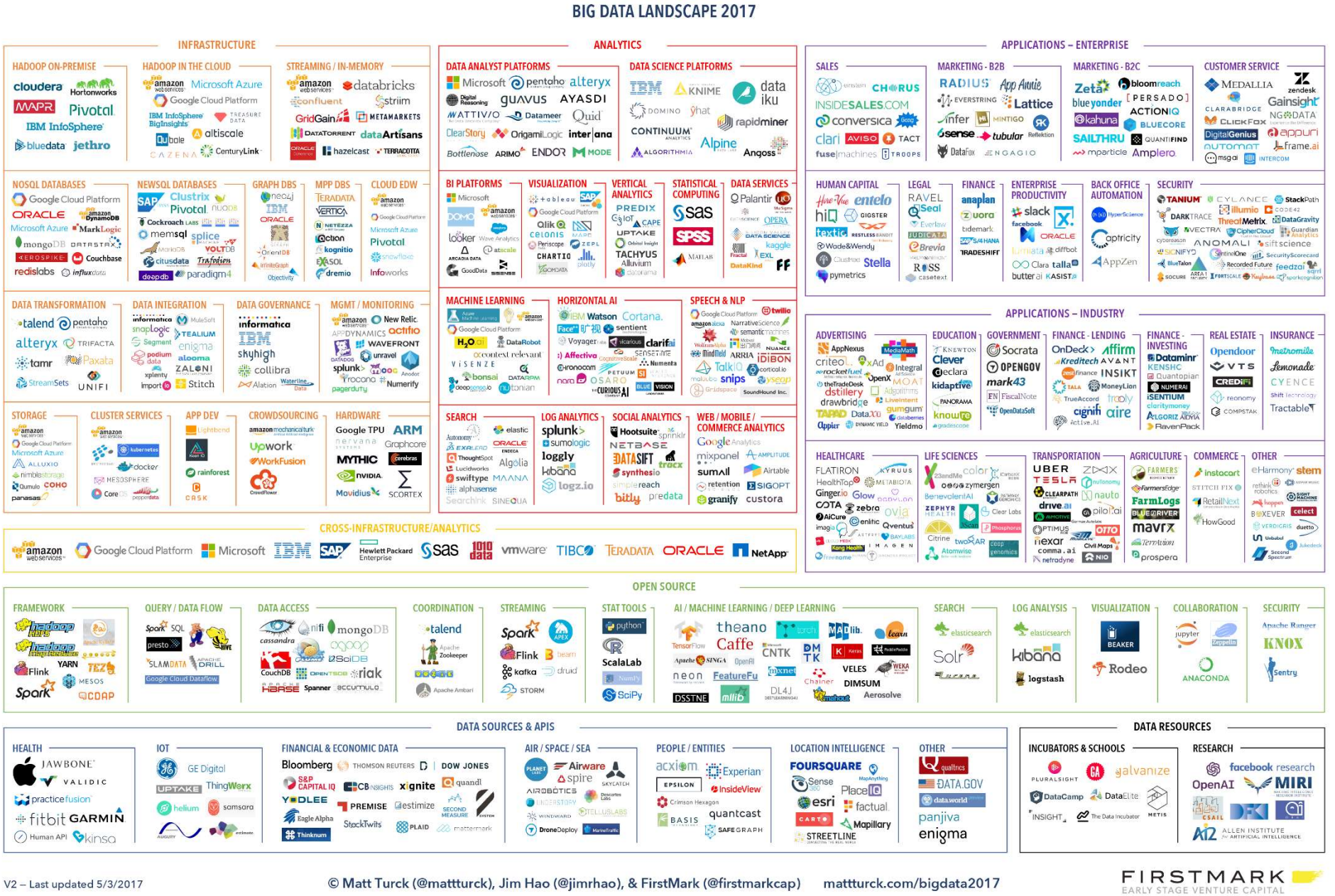
- Pass/Fail Criteria
- Test Results
- AF/AL Results
- Auditable Record System
- Results Trending
- Doc Control Log

- RAG Status
- Real time Degradation / Failure Alerts
- Troubleshooting tools

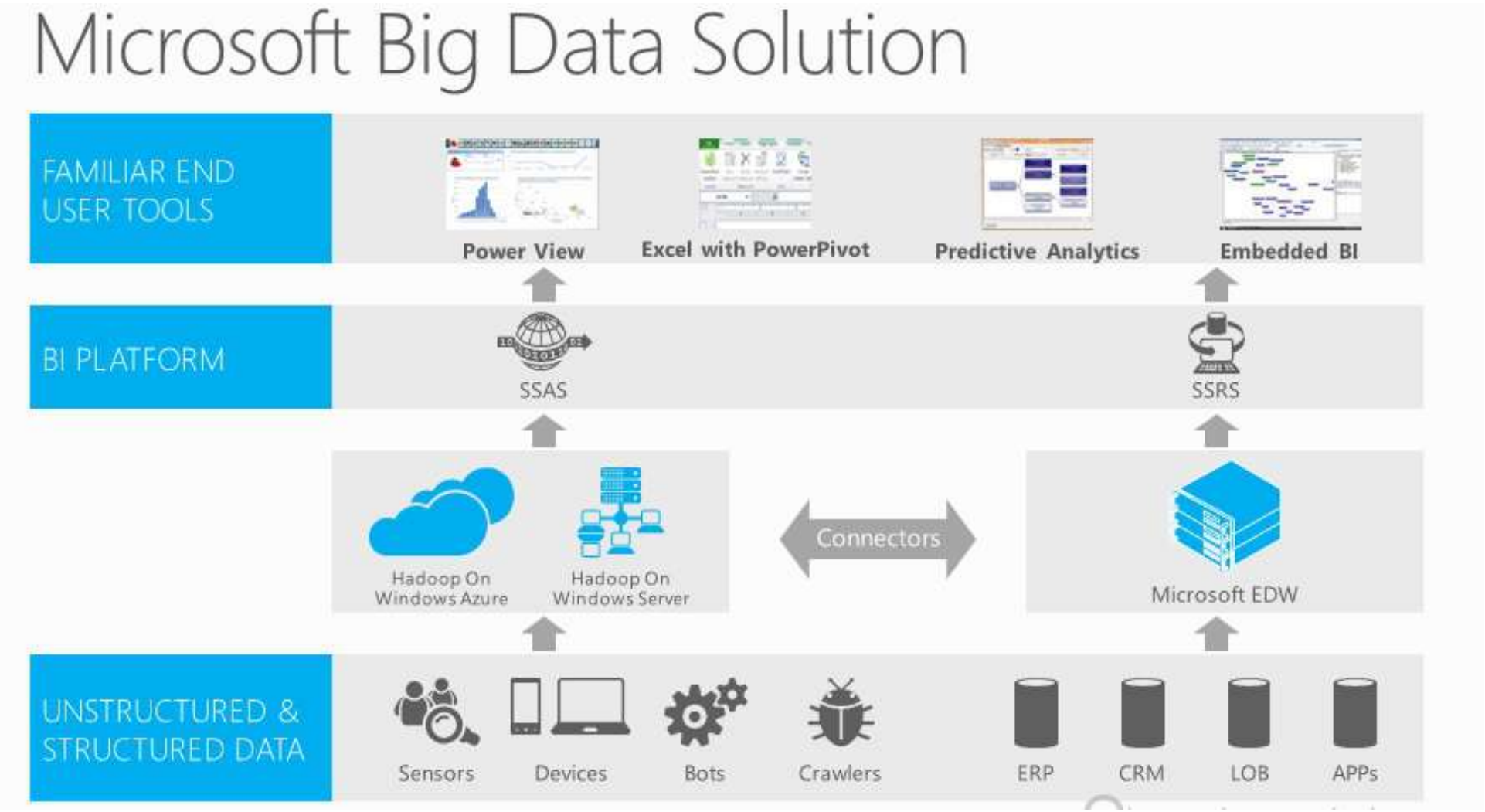
- System data validation
- Process stability
- Stream comparisons

So how do we achieve this???

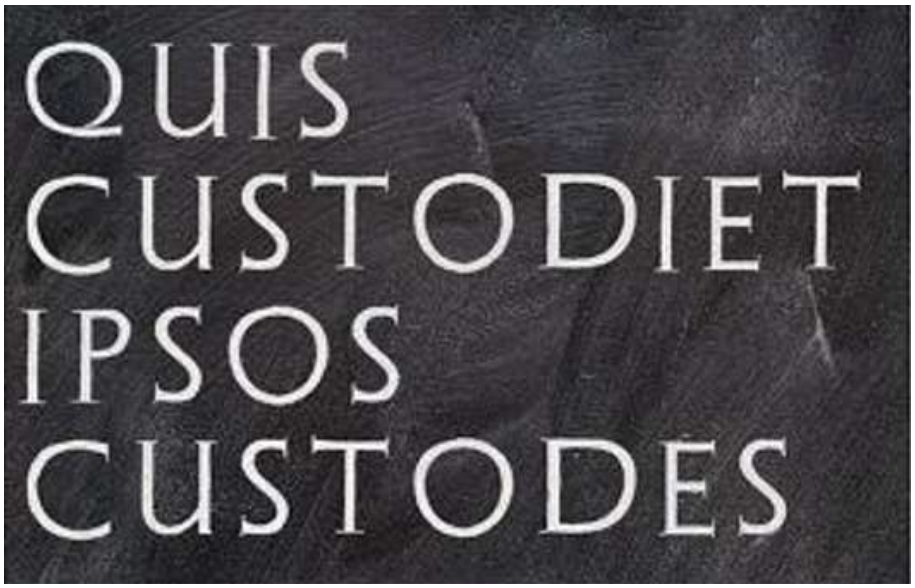
big
data
information
storage
search
sets
petabytes
large
management
analysis
technologies
research
science
development
software
size
billion
new
future
process
government
internet
volume
technology
exabytes
million
san
nas
databases
learning
mapreduce
business
processing
critiques
world
parallel
framework
analytics
applications
based
amount
approaches
initiative
challenges
social
implementation
increasing
factor
structure
sources
state
terms
nearly
second
processes
scientific
companies
shared
time
times
people
various
limits
high
private
simulations
statistics
distributed
flow
per
drive
online
hand
many
institute
paradigm
international
trends
algorithms
market
include
integration
tools
used
program
year
queries
traffic
determine
networks
terabytes
sensor
between
even
variety
effective
results
ability
privacy
higher
uses
making
countries
past
related
national
database
cost
most
handle
definition
worldwide
users
successful
work
end
rate
real
needed
funding
economic
day
systems
out
hypothesis
huge
every
across
critique
years
well
now
relevant
intelligence
records
visualization
complex
insight
massive
infrastructure
manage
center
set
one
replication
use
using
project
less
would


BIG DATA LANDSCAPE 2017
INFRASTRUCTURE
HADOOP ON-PREMISE
HADOOP IN THE CLOUD
STREAMING / IN-MEMORY
NOSQL DATABASES
NEWSQL DATABASES
GRAPH DBS
MPP DBS
CLOUD EDW
DATA TRANSFORMATION
DATA INTEGRATION
DATA GOVERNANCE
MGMT / MONITORING
STORAGE
CLUSTER SERVICES
APP DEV
CROWDSOURCING
HARDWARE
ANALYTICS
DATA ANALYST PLATFORMS
DATA SCIENCE PLATFORMS
BI PLATFORMS
VISUALIZATION
VERTICAL ANALYTICS
STATISTICAL COMPUTING
DATA SERVICES
MACHINE LEARNING
HORIZONTAL AI
SPEECH & NLP
SEARCH
LOG ANALYTICS
SOCIAL ANALYTICS
WEB / MOBILE / COMMERCE ANALYTICS
CROSS-INFRASTRUCTURE/ANALYTICS
APPLICATIONS – ENTERPRISE
SALES
MARKETING - B2B
MARKETING - B2C
CUSTOMER SERVICE
HUMAN CAPITAL
LEGAL
FINANCE
ENTERPRISE PRODUCTIVITY
BACK OFFICE AUTOMATION
SECURITY
APPLICATIONS – INDUSTRY
ADVERTISING
EDUCATION
GOVERNMENT
FINANCE - LENDING
FINANCE - INVESTING
REAL ESTATE
INSURANCE
HEALTHCARE
LIFE SCIENCES
TRANSPORTATION
AGRICULTURE
COMMERCE
OTHER
OPEN SOURCE
FRAMEWORK
QUERY / DATA FLOW
DATA ACCESS
COORDINATION
STREAMING
STAT TOOLS
AI / MACHINE LEARNING / DEEP LEARNING
SEARCH
LOG ANALYSIS
VISUALIZATION
COLLABORATION
SECURITY
DATA SOURCES & APIS
HEALTH
IOT
FINANCIAL & ECONOMIC DATA
AIR / SPACE / SEA
PEOPLE / ENTITIES
LOCATION INTELLIGENCE
OTHER
DATA RESOURCES
INCUBATORS & SCHOOLS
RESEARCH
V2 – Last updated 5/3/2017
© Matt Turck (@mattturck), Jim Hao (@jimrhao), & FirstMark (@firstmarkcap)
mattturck.com/bigdata2017
FIRSTMARK
EARLY STAGE VENTURE CAPITAL

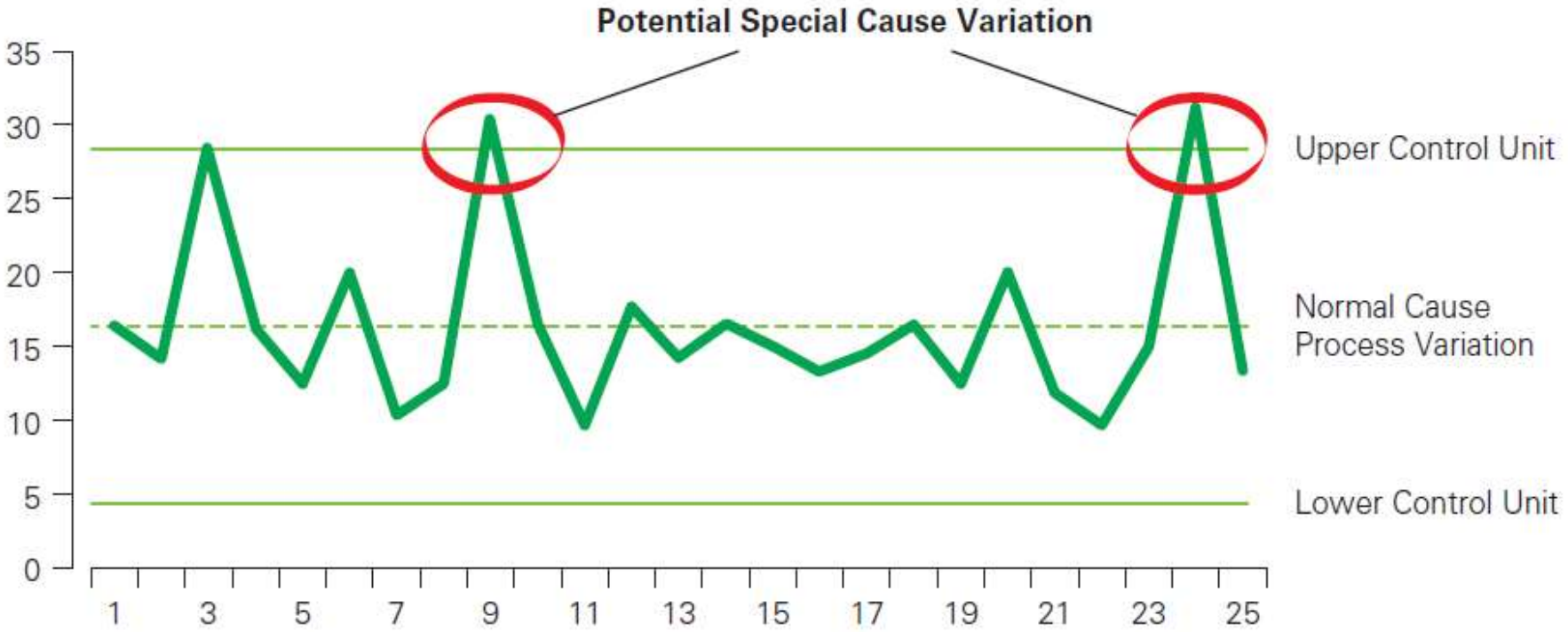
Big Data – Safeguarding our Future…?



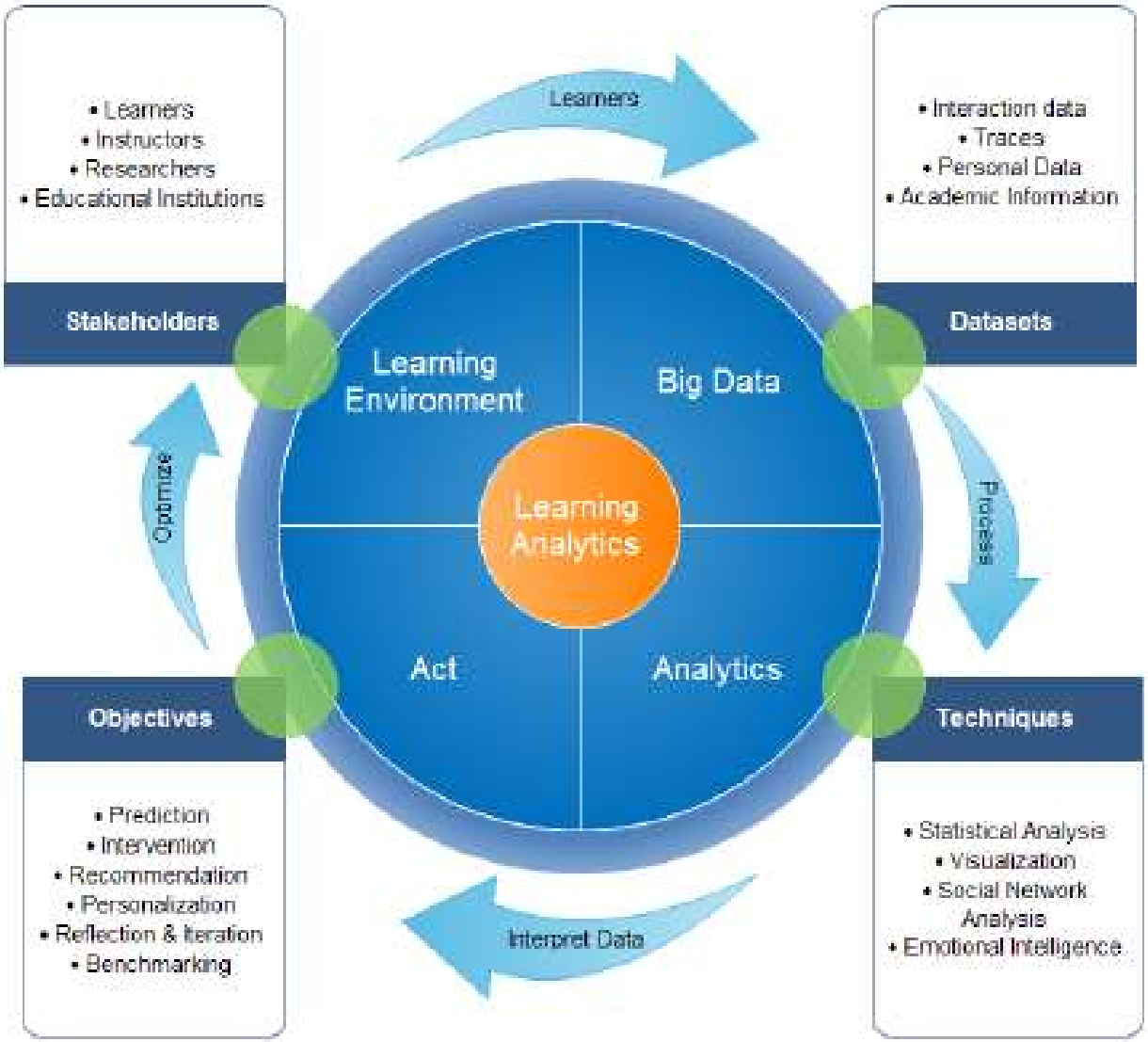
Big Data – Safeguarding our Future…?



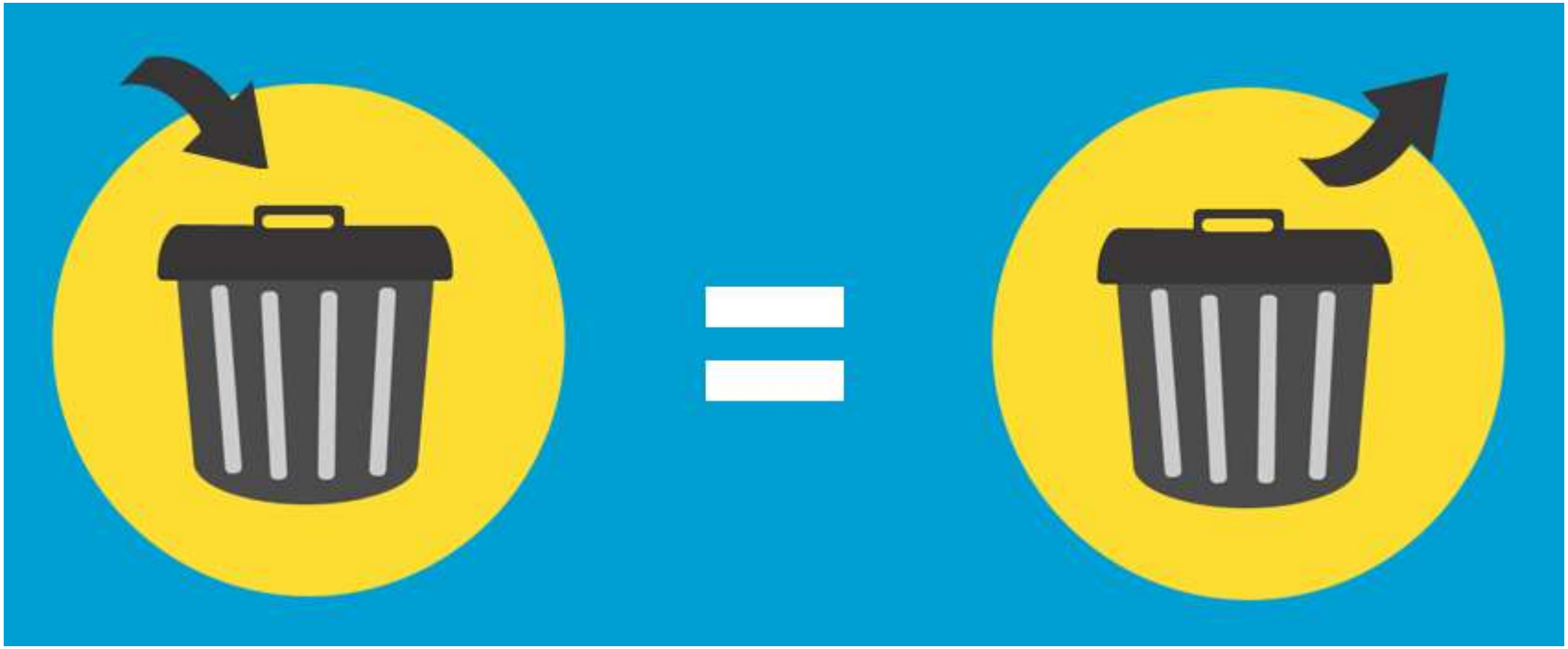
Identifying Emergent Risks



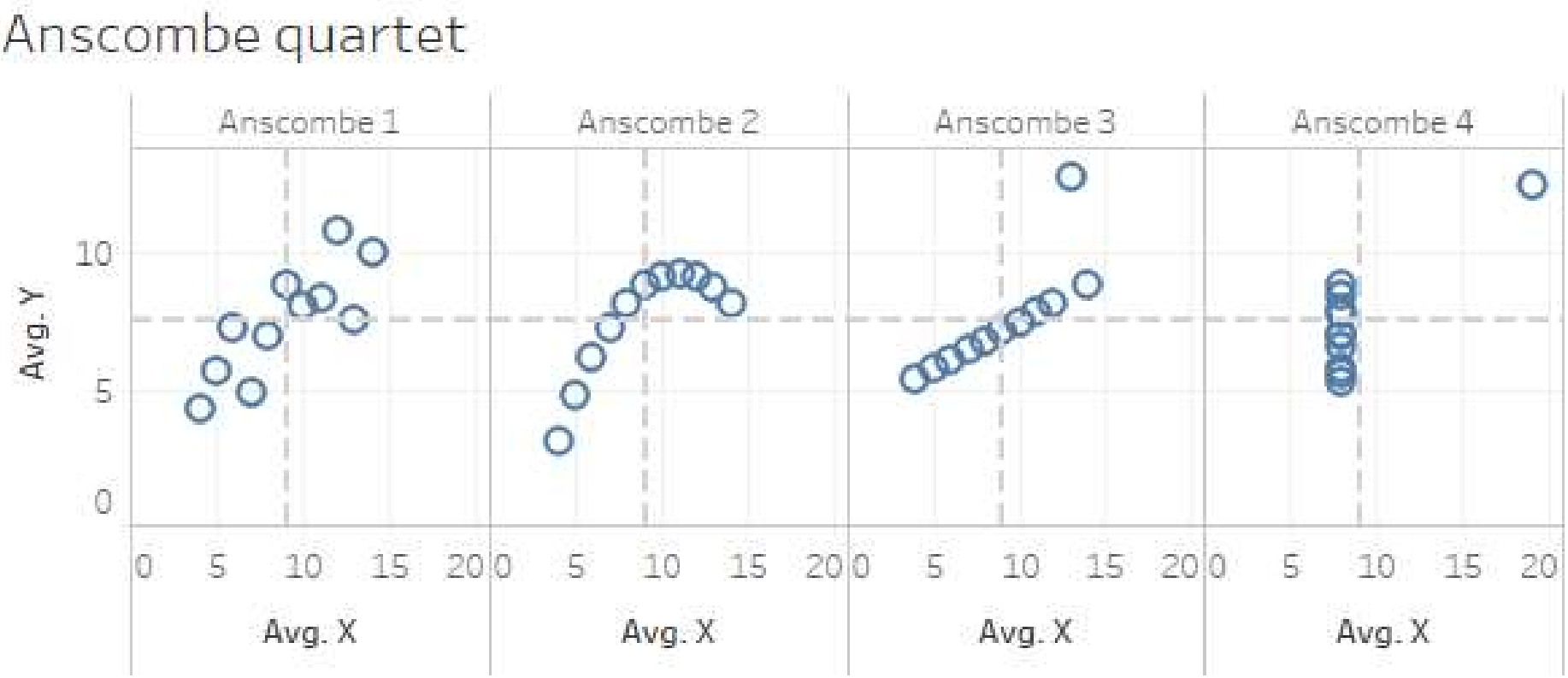
Solving this problem using big data



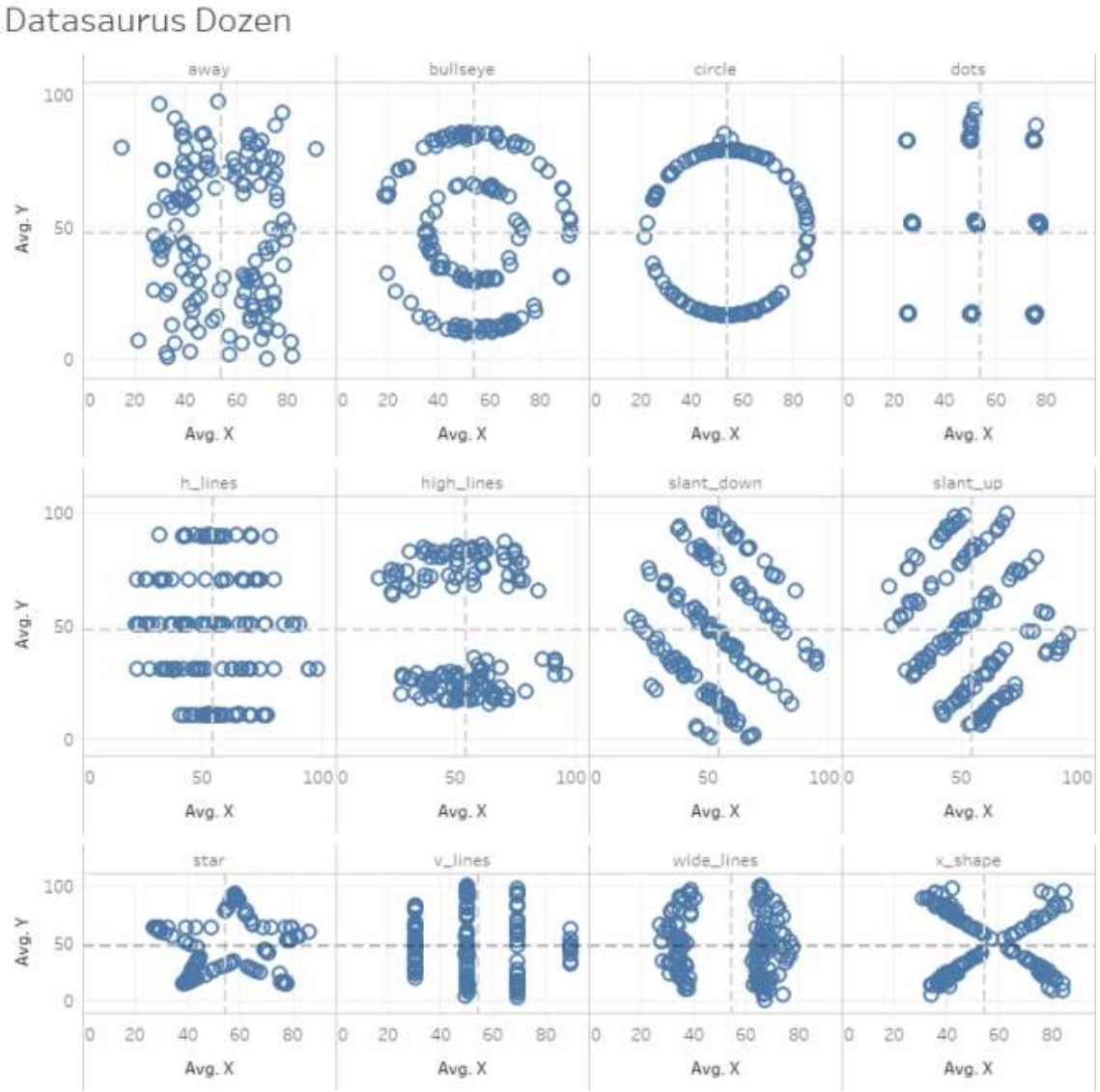
Solving this problem using big data



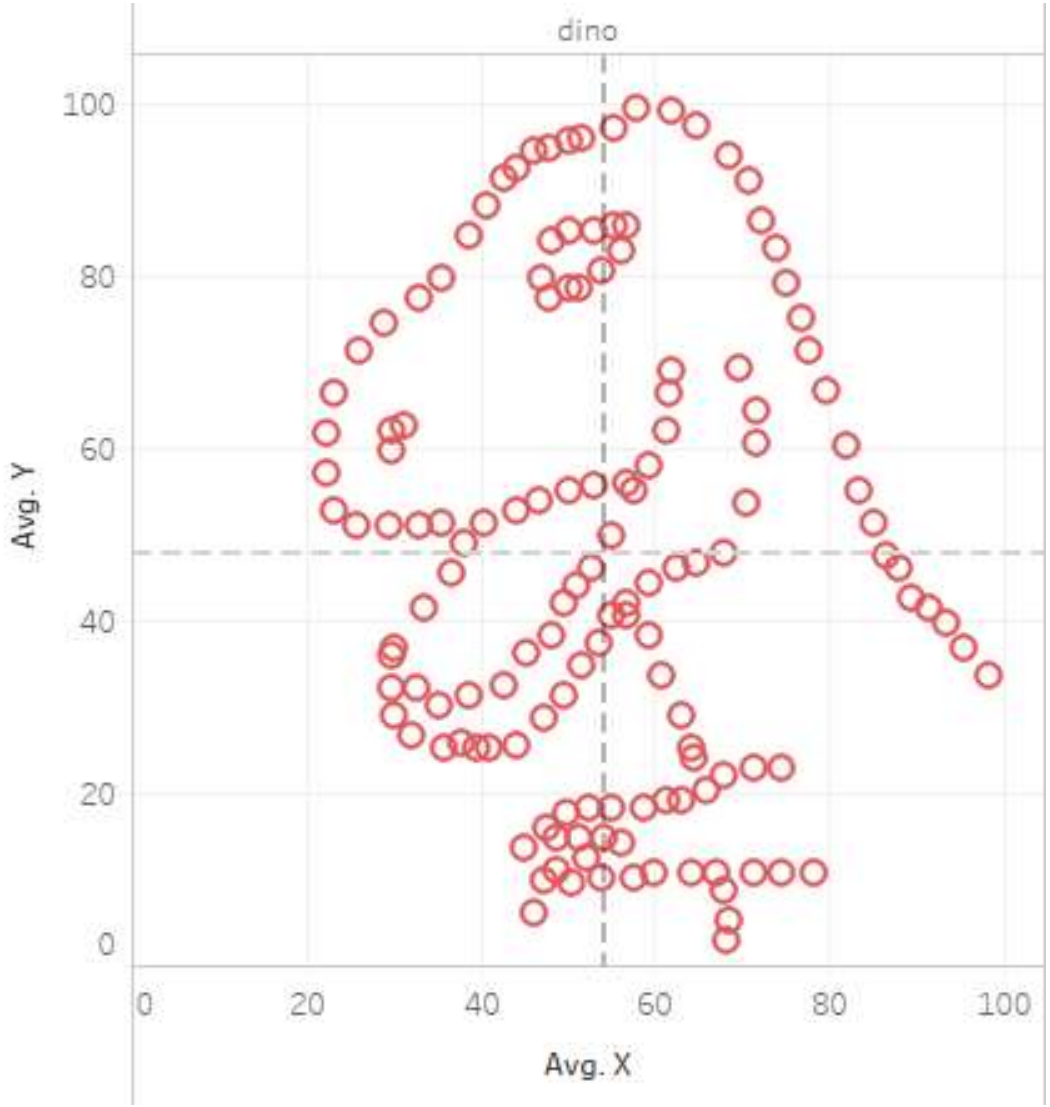
A hidden problem…



The Datasaurus dozen



The Datasaurus dozen



Where to Start?

Where to Start?

SAFETY NOTE:

Although many of these techniques are intended to improve engineer and technician efficiency, it is essential to ensure that safety systems shall not be compromised by changes made in the way these systems are maintained.

Where to Start?

Wh[illegible] Start?

Raw Data

- Process Control Network
- Data Historian
- Data Lakes
- Spreadsheets(!)

Quality Checks

- Risk
- Exsposure
- Uncertainty
- Status

- Process Data
- Maintenance Activities/Results
- CBM

Data Analytics

This document is

Measurement Management System Vision

Outputs to Business Big Data System

Field Inputs

- Flow Rate
- Pressure
- Temperature
- Gas CV
- HCDP
- Moisture
- Water Cut
- BS&W

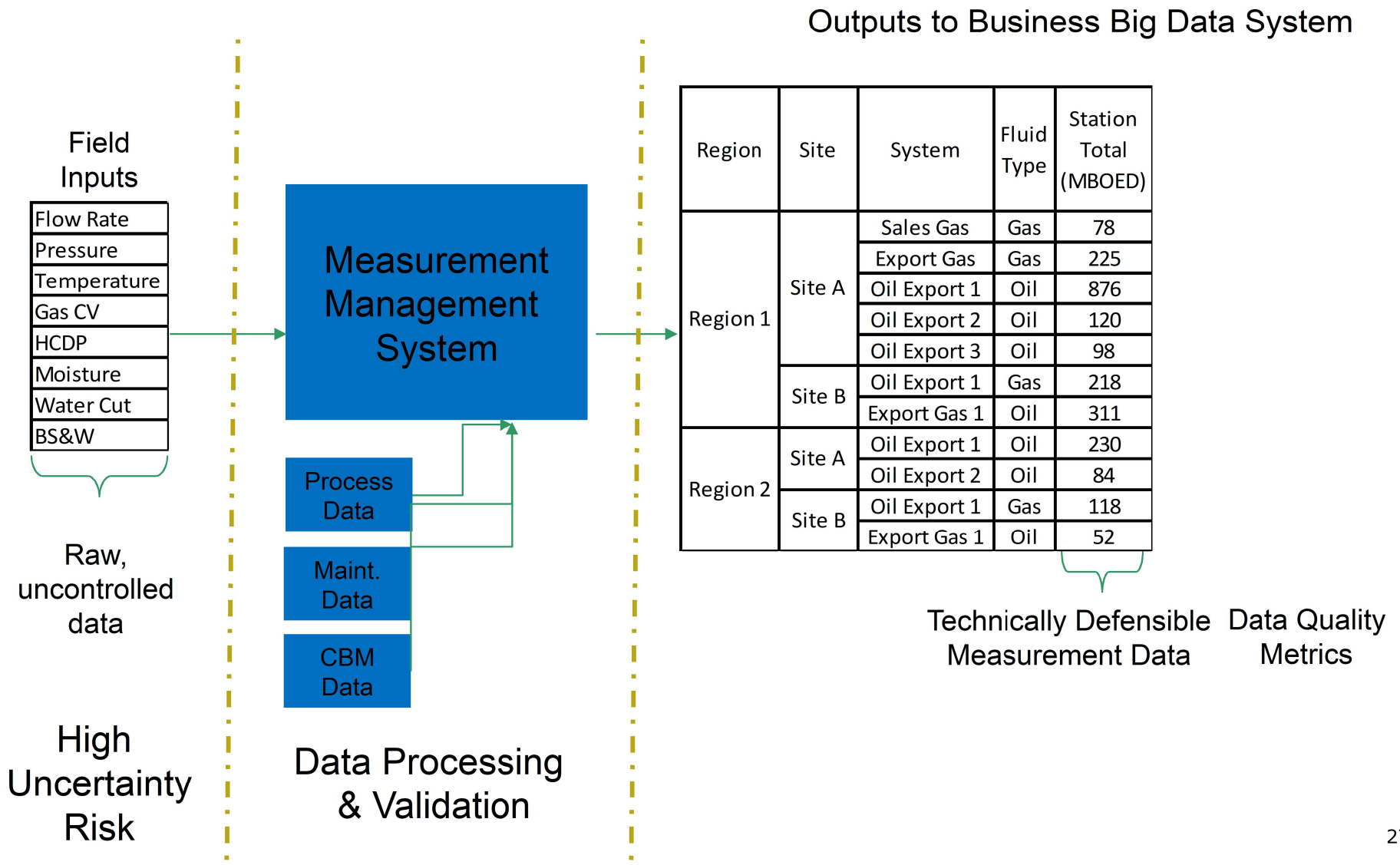
Raw, uncontrolled data

Region	Site	System	Fluid Type	Station Total (MBOED)
Region 1	Site A	Sales Gas	Gas	78
		Export Gas	Gas	225
		Oil Export 1	Oil	876
		Oil Export 2	Oil	120
		Oil Export 3	Oil	98
	Site B	Oil Export 1	Gas	218
		Export Gas 1	Oil	311
Region 2	Site A	Oil Export 1	Oil	230
		Oil Export 2	Oil	84
	Site B	Oil Export 1	Gas	118
		Export Gas 1	Oil	52

Technically Defensible Measurement Data?

High Uncertainty Risk

Measurement Management System Vision



Region	Site	System	Fluid Type	Station Total (MBOED)
Region 1	Site A	Sales Gas	Gas	78
		Export Gas	Gas	225
		Oil Export 1	Oil	876
		Oil Export 2	Oil	120
		Oil Export 3	Oil	98
	Site B	Oil Export 1	Gas	218
		Export Gas 1	Oil	311
Region 2	Site A	Oil Export 1	Oil	230
		Oil Export 2	Oil	84
	Site B	Oil Export 1	Gas	118
		Export Gas 1	Oil	52

Wh[illegible] Start?

Raw Data

Quality Checks

- Risk
- Exsposure
- Uncertainty
- Status

- Process Data
- Maintenance Activities/Results
- CBM

Data Analytics

How important is this Step to ME?

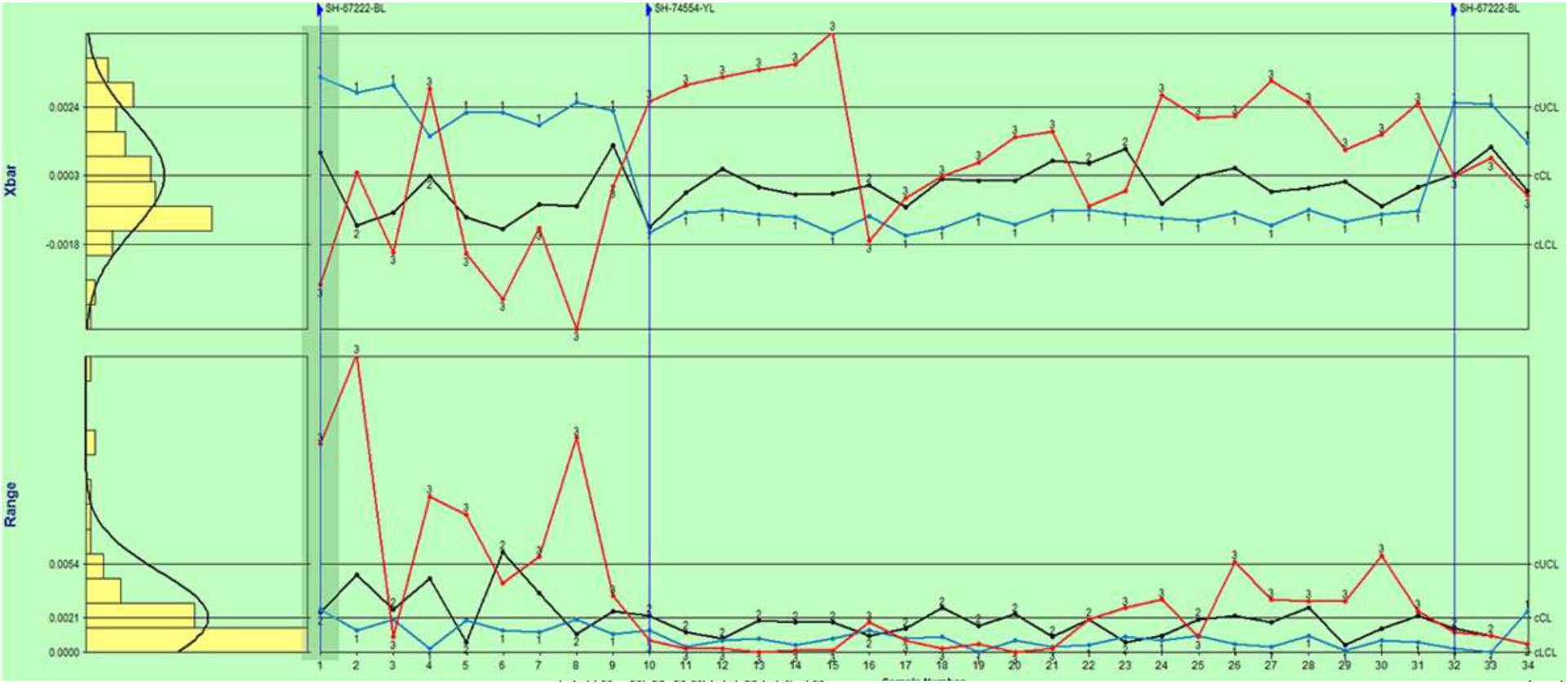
This document is

Process Data

Process Data

Process Data Validation Considerations:

- Operating within the intended design envelope of the system
- Cross checks between parallel measurement streams
- Range of measurement acceptable (volatility)

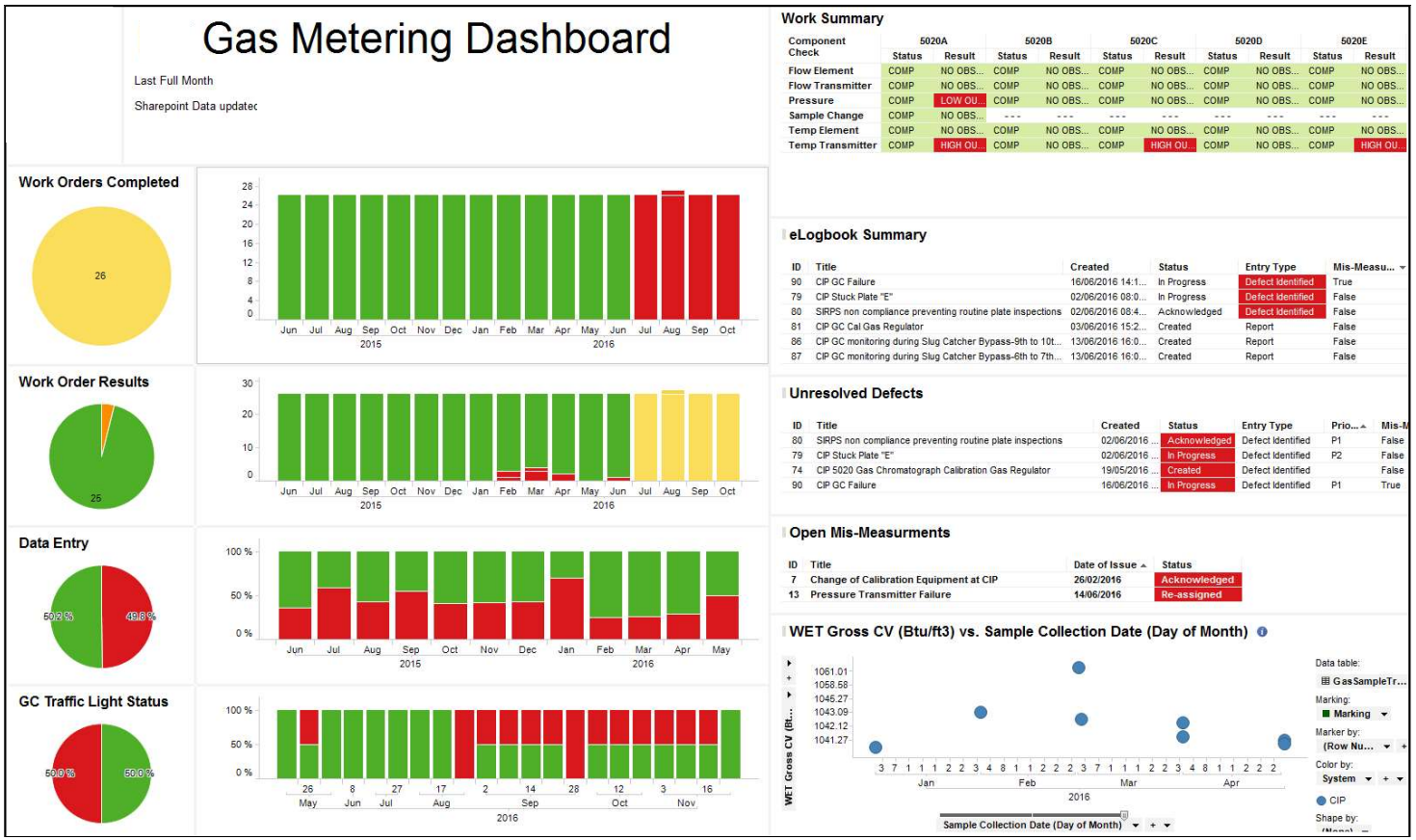


Maintenance Data

Maint. Data

Confidence in Maintenance Activities:

- Confirming that Maintenance *WAS* conducted
- Do results make sense compared to last time…?
- Early detection of failing components

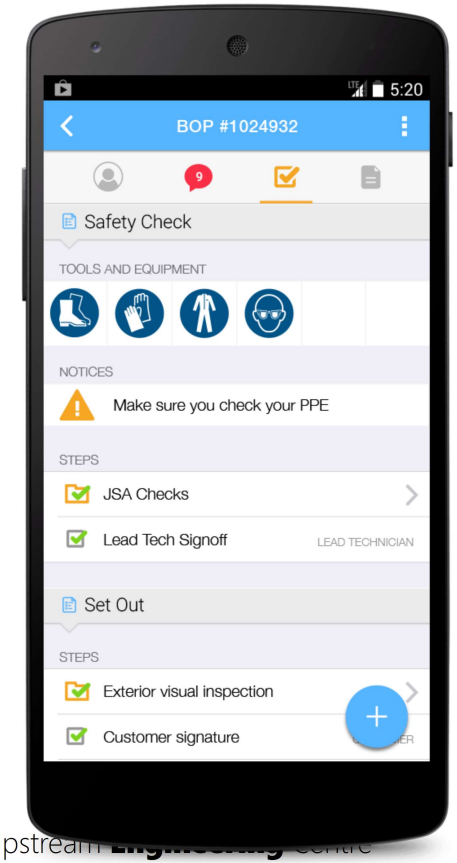


Maintenance Data

Maint. Data

Confidence in Maintenance Activities:

- Maintenance *WAS* conducted
- Results make sense compared to last time…
- Early detection of failing components



Flow Transmitter Data

Tag Number:

WO Number	725109
Result:	NO OBSERVED FAILURE MODE
Span:	200
Units:	InH2O
Tolerance %	0.25%
Absolute Tolerance	0.5

Test Equipment

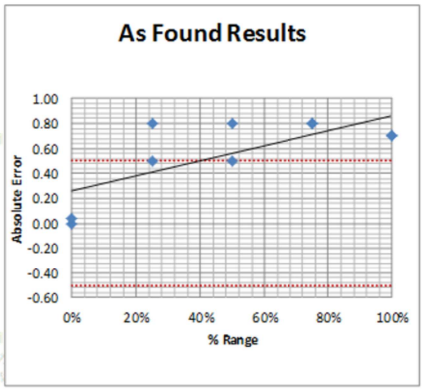
Name/Number:	**saddsf**
Serial Number:	
Calibration Date:	42012
Procedure Rev #	0
Procedure Rev Date	42086

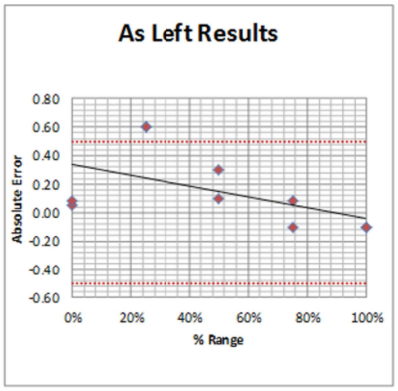
Automated Checks and Outlier Warnings

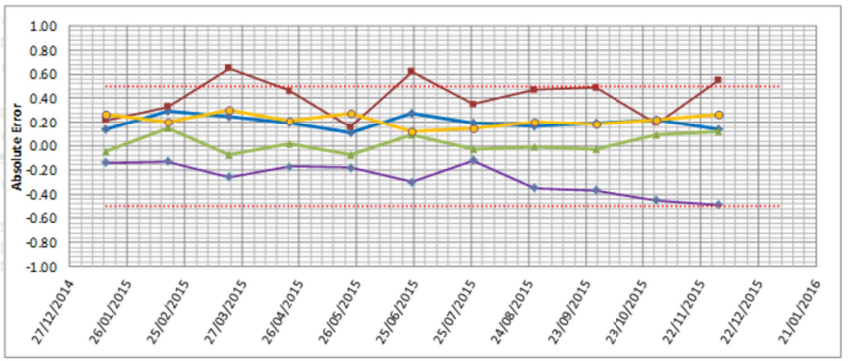
Equalization Valve Closed?	PASS
Impulse Lines Free from Plugging?	PASS
Test Loop Leak test:	PASS
AF HP Static Zero:	0.02
	PASS
AL HP Static Zero:	0.05
	PASS
AF Process Value:	645
AL Process Value:	74.5
Test Result AF/AL or AL?:	AL

Self Verification

Tolerance Limit	±0.5
Range variance	0.06
# Failed AF Calibrations in past 12 months	3.00







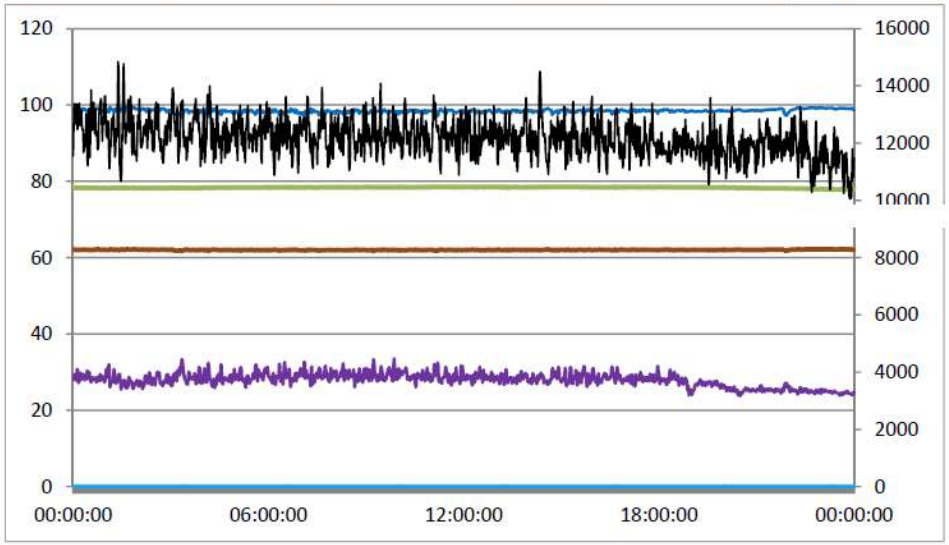
CBM Data

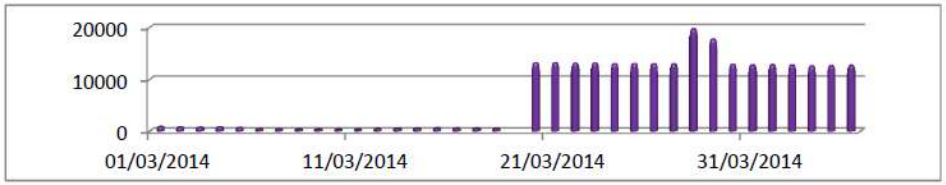
CBM Data

CBM/Diagnostic Data Validation Considerations:

- Operating within the intended design envelope of the system
- Cross checks between parallel measurement streams
- Aggregate of system diagnostics confirms that signal is good?

Stream		Today	Prev 24hrs	Units	% Change	Uptime	Process	Maint.	Cal Drift
	Total Oil	175	196	bbls	112%	OK	PROCESS	?	?
	Total Water	12,669	13,106	bbls	103%	OK	PROCESS	?	?
???	Total Liquid	0	0	0	#DIV/0!	OK	PROCESS	?	?





Coriolis Sensor Details	
Model	0
Serial Number	0
Stored Zero Value	0
Zero Cal. Temp.	80

Key	PI Tag Name	Description	Min	Avg	Max	Units
	.pv	Mixed Stream Flow Rate	10047.3	12172.9	14839.9	bbls/d
	\.pv	Mixed Stream Density	61.66	62.01	62.21	lb/ft3
	..pv	Drive Gain (%)	23.81	27.95	33.57	Var. 9.759
		Line Pressure	0	0	0	0%
	..pv	Line Temperature	77.81	78.34	78.55	0
	..pv	Sensor Failure	OK			OK/Fail
	\.pv	Mixed Stream Water Cut	96.46	98.42	99.57	%
		Red Eye Water Cut	0	0	0	0

Density Keypad Values		Last Sample Date
Dry Oil	0	?
Base Water	0	?

Process	Min	Max	Diagnostics	Amber	Red
Mixed Stream FR bbl	1500	30000	Drive Gain	20	30
Temperature F	60	90	DG VAR	4	8
Density lb/ft3	40	65	Water Cut %	70	85
Pressure psi	0	1000	Px VAR	5%	10%

Wh[illegible] Start?

Raw Data

Quality Checks

- Risk
- Exsposure
- Uncertainty
- Status

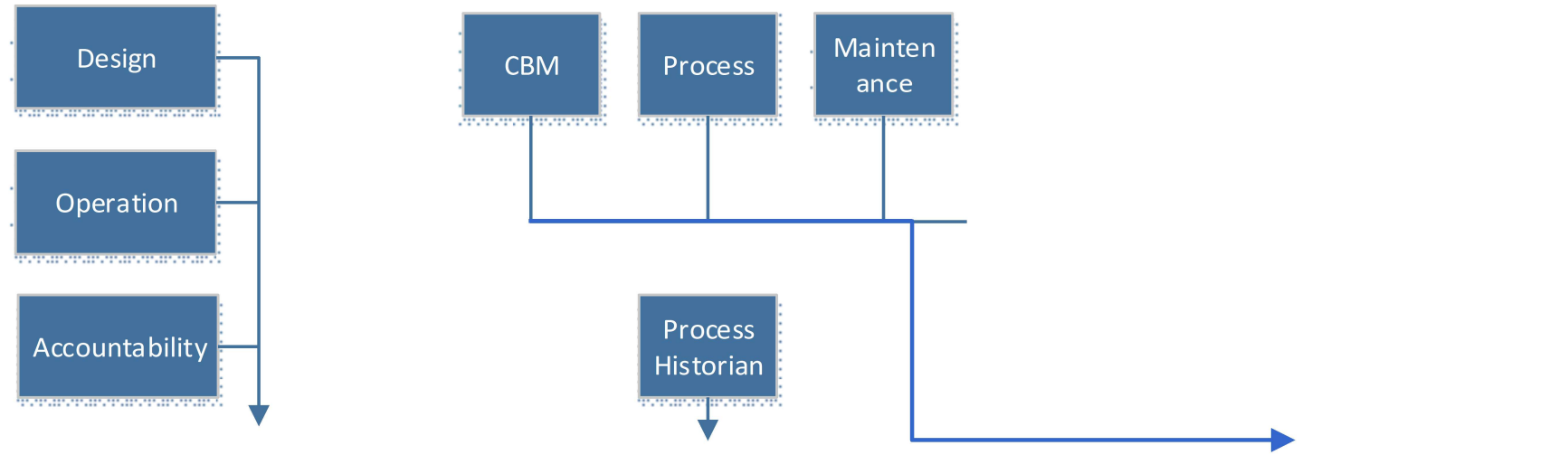
- Process Data
- Maintenance Activities/Results
- CBM

Data Analytics

How important is this Step to ME?

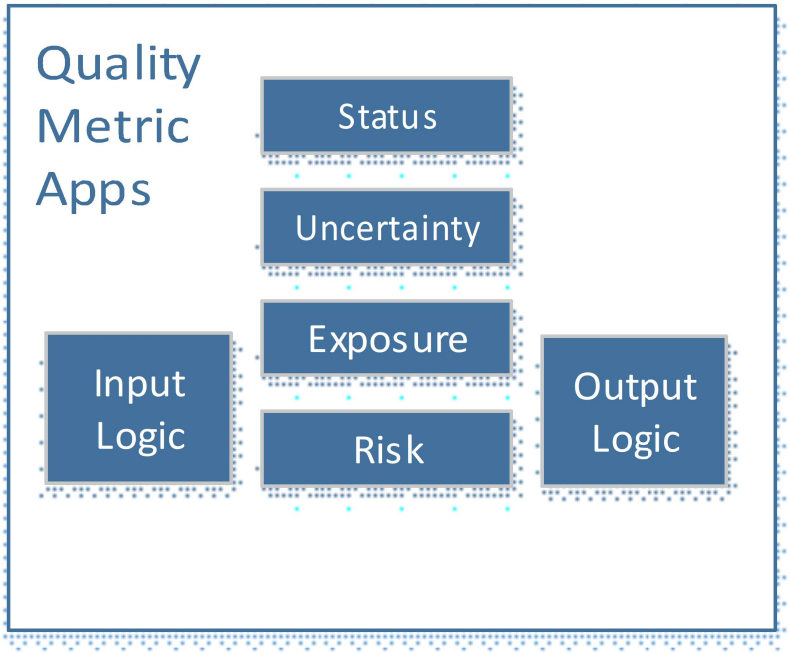
This document is

Pulling it Together

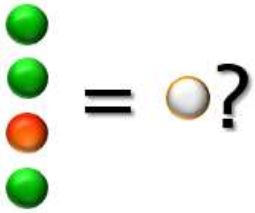


Region	Site	System/ Station	Fluid Type	Total Vol	Total Std Vol	Total Mass	Total BOE	Risk	Exposure	Uncert. %	Status
Region 1	Site 1	Sys 1	Oil	Total Vol	Total Std Vol	Total Mass	Total BOE	Risk	Exposure	Uncert. %	Status

Translating Quality Metrics



- **Status:**



- **Uncertainty**

$$CU = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

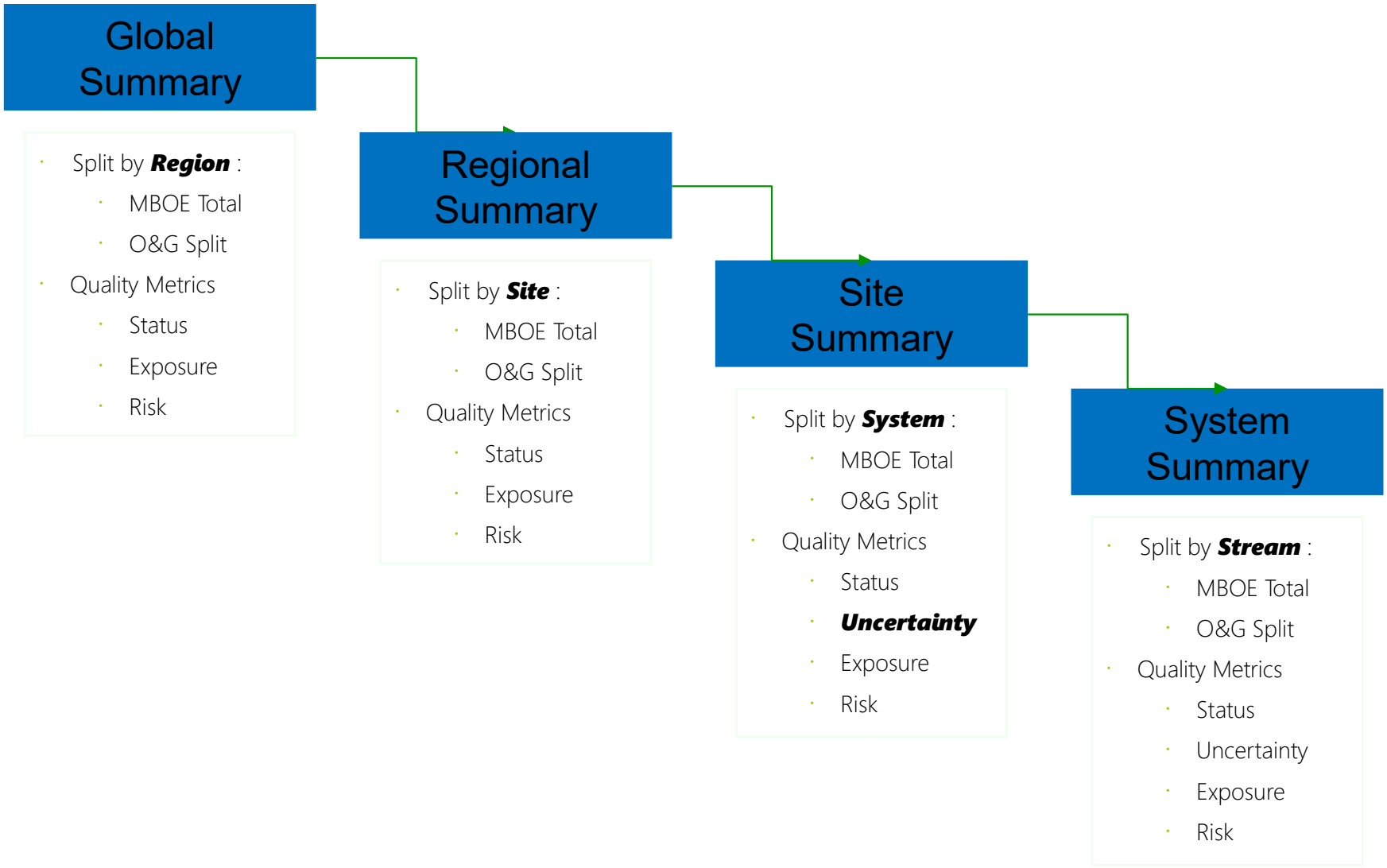
- **Ex**

Measured Value x Uncertainty

- **Risk**

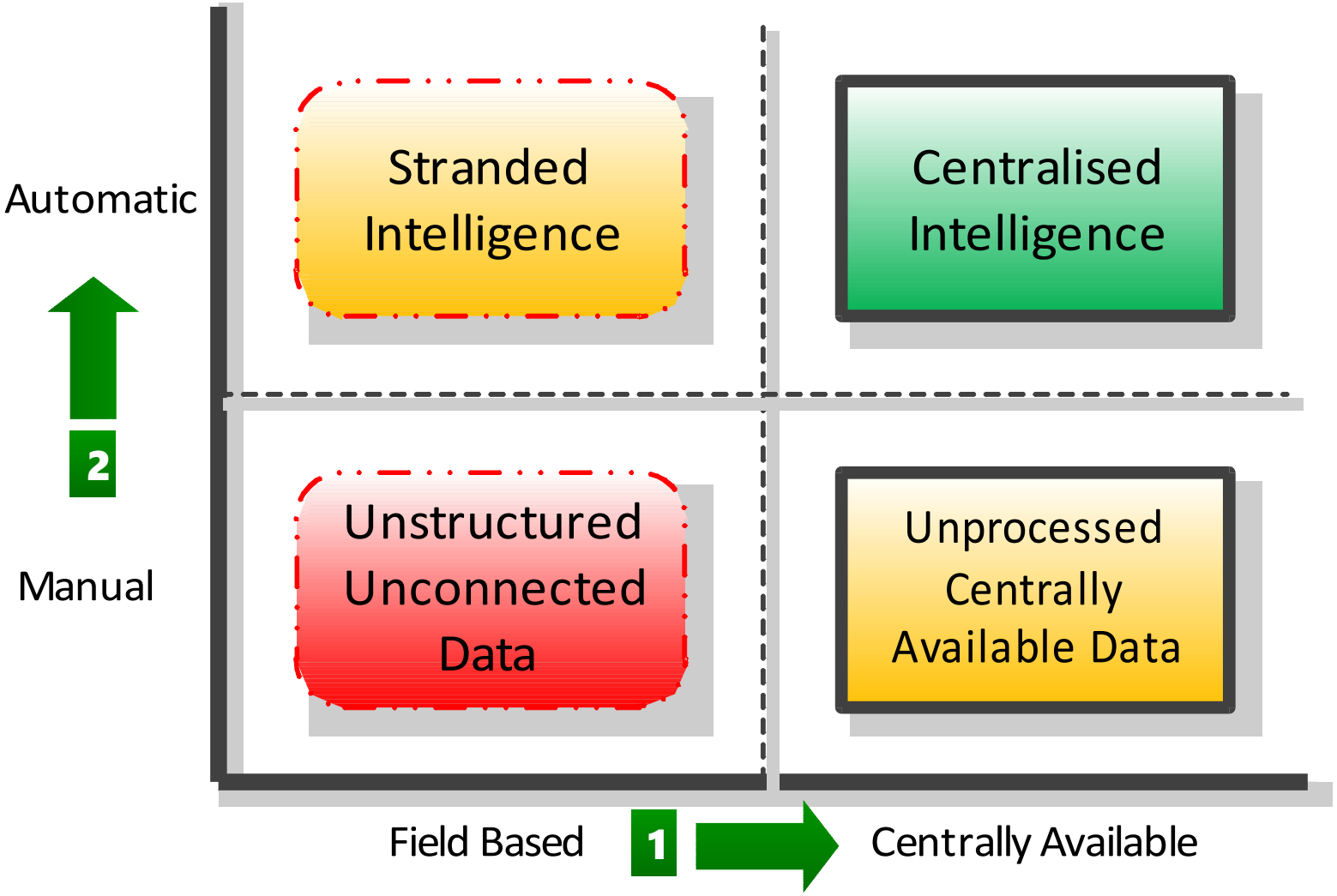
		A	B	C	D	E
		Negligible	Minor	Moderate	Significant	Severe
E	Very Likely	Low Med	Medium	Med Hi	High	High
D	Likely	Low	Low Med	Medium	Med Hi	High
C	Possible	Low	Low Med	Medium	Med Hi	Med Hi
B	Unlikely	Low	Low Med	Low Med	Medium	Med Hi
A	Very Unlikely	Low	Low	Low Med	Medium	Medium

Measurement Management Hierarchy

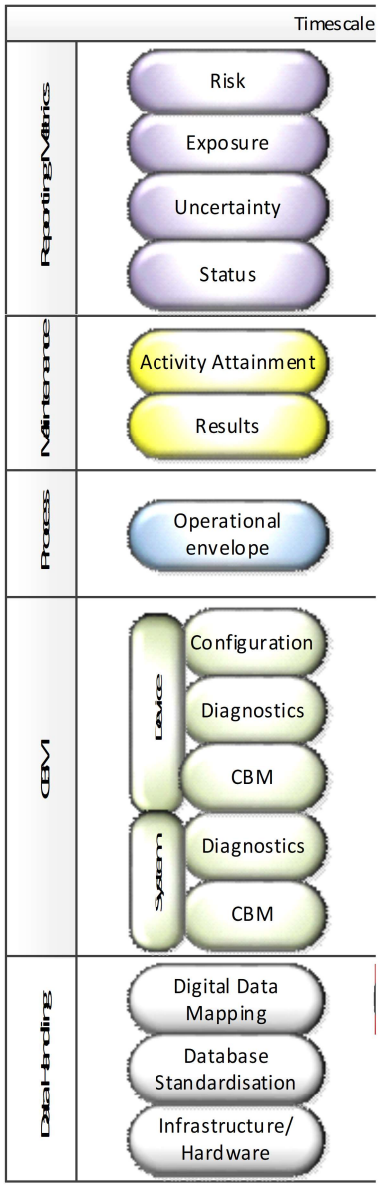


This document is

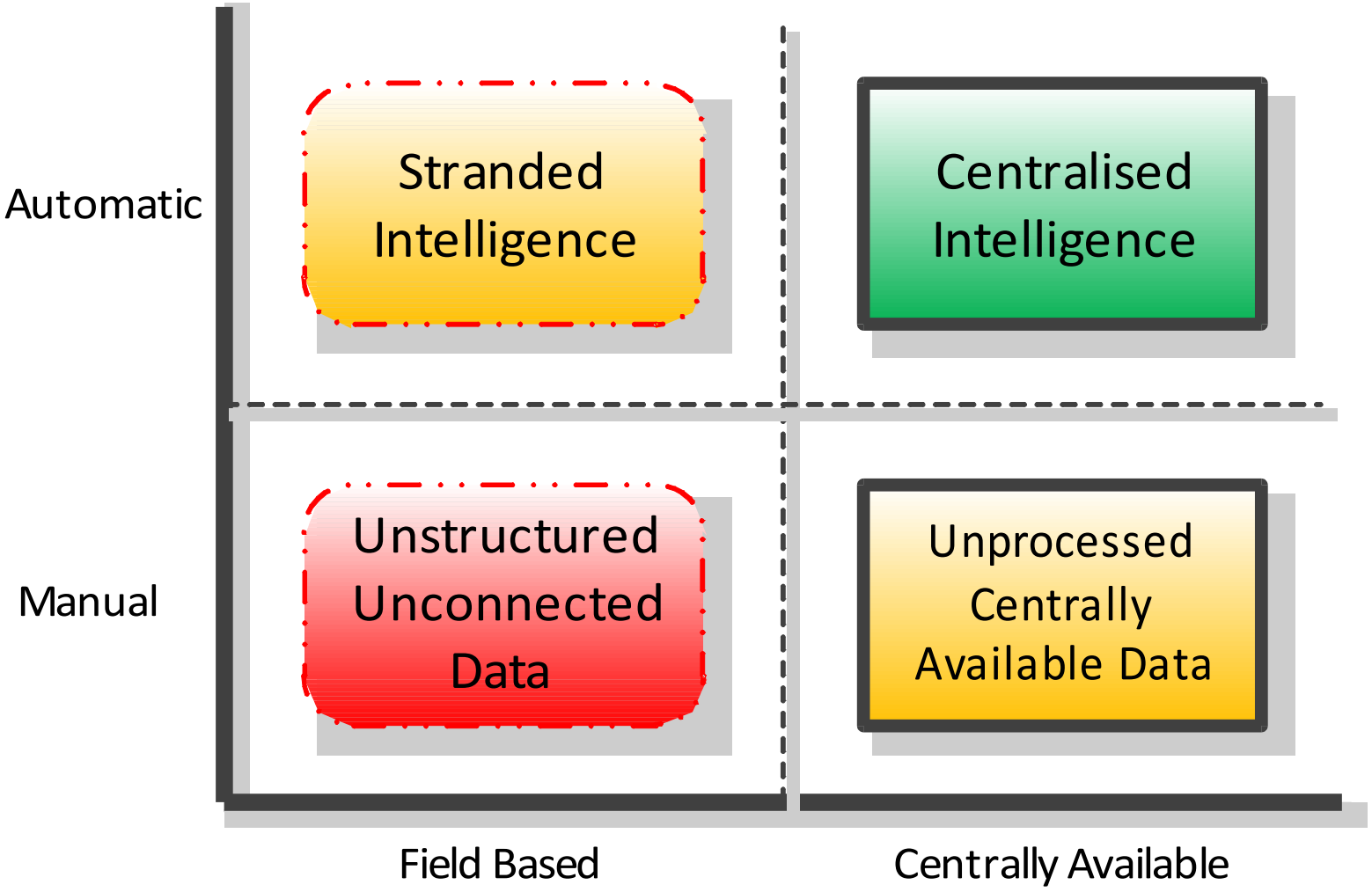
Where to Start?



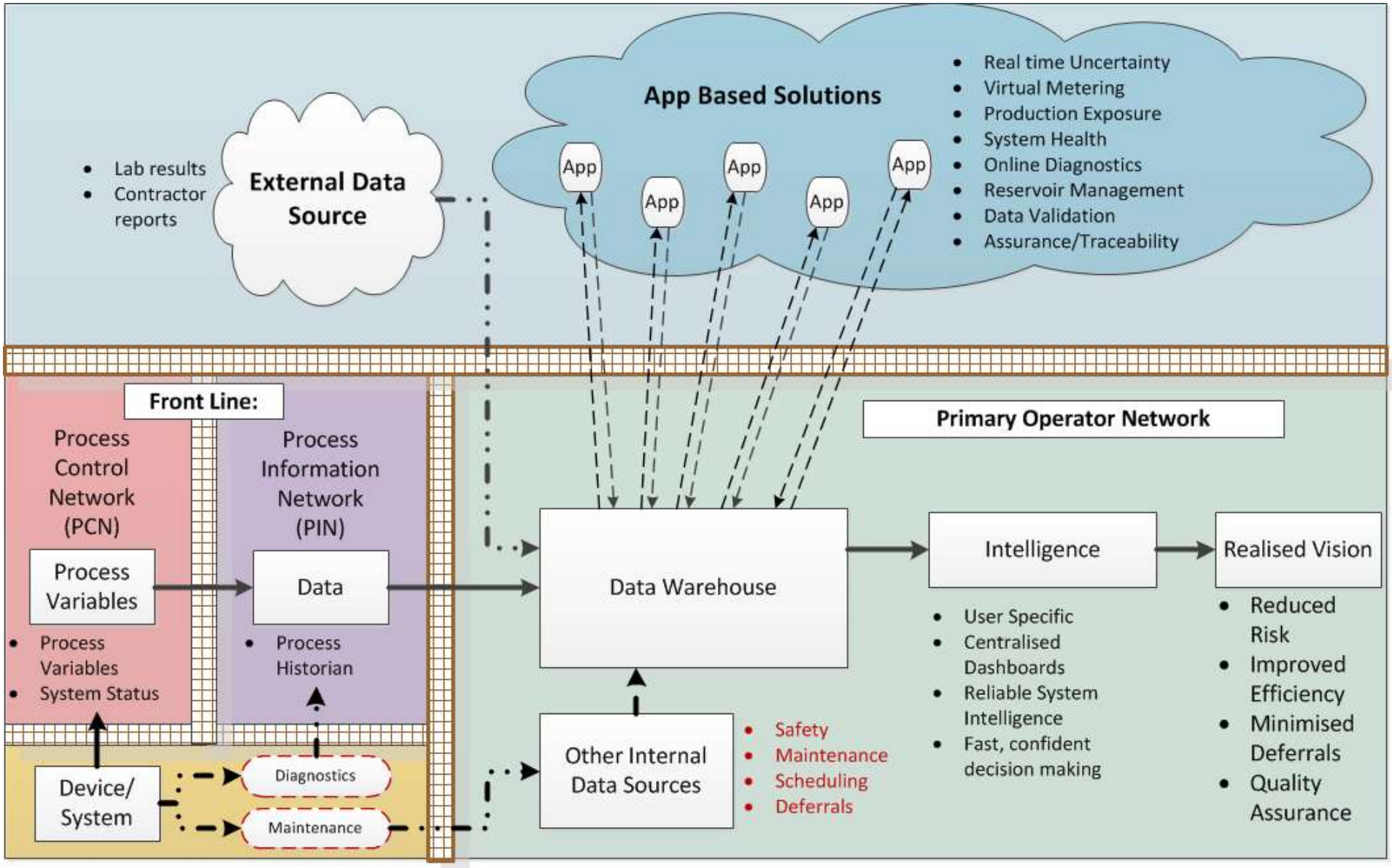
Considerations & Sequencing Options



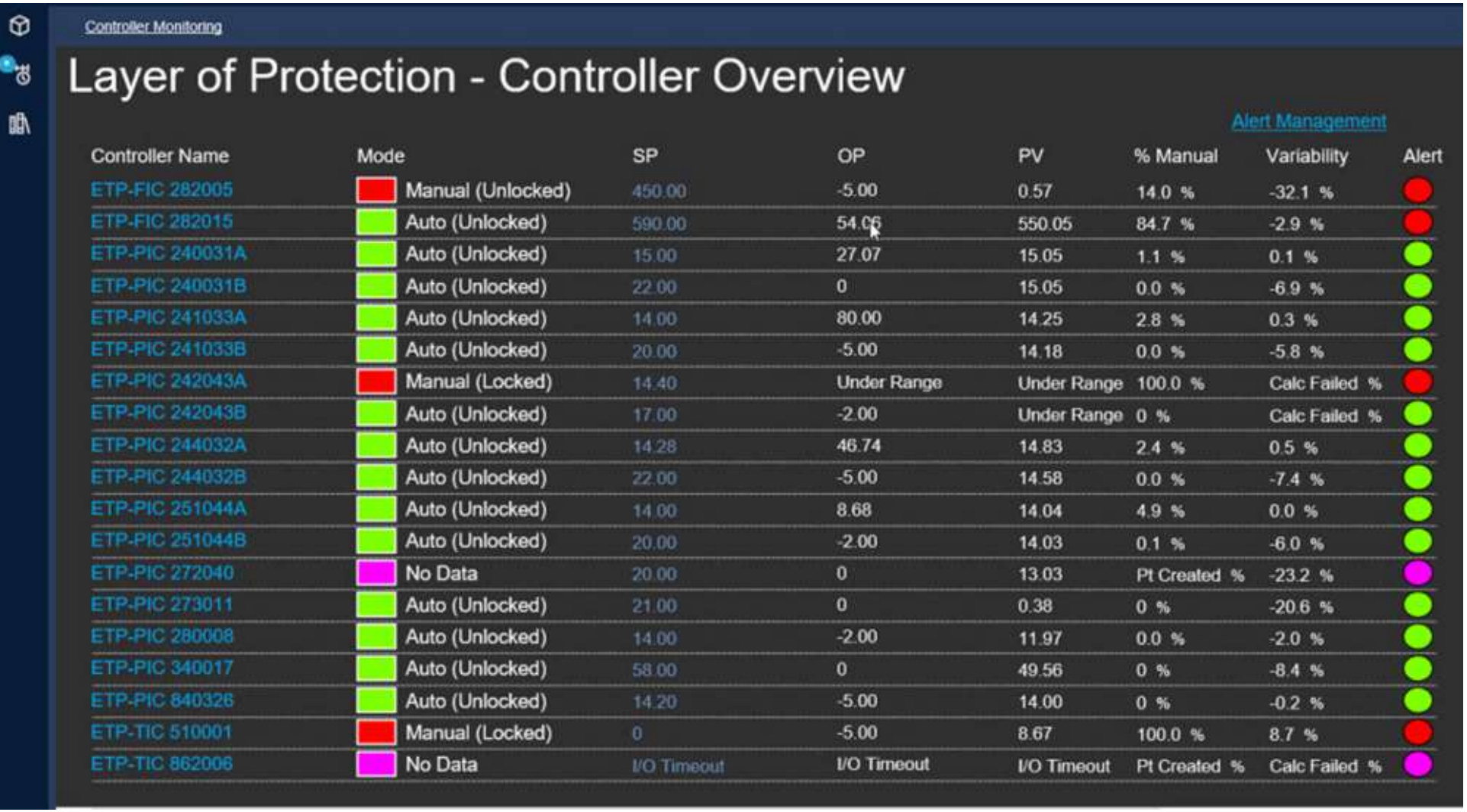
Automating for Efficiency



Infrastructure & Hardware Considerations

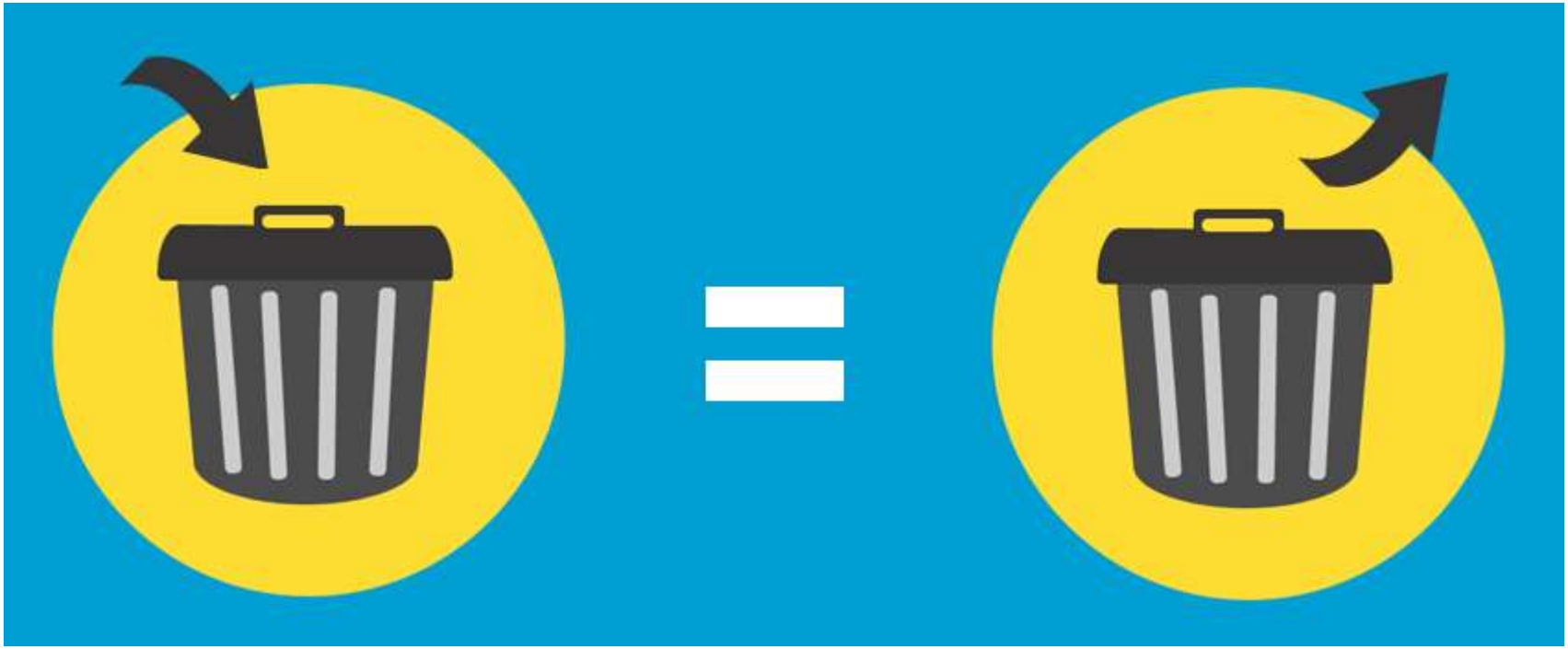


Reporting



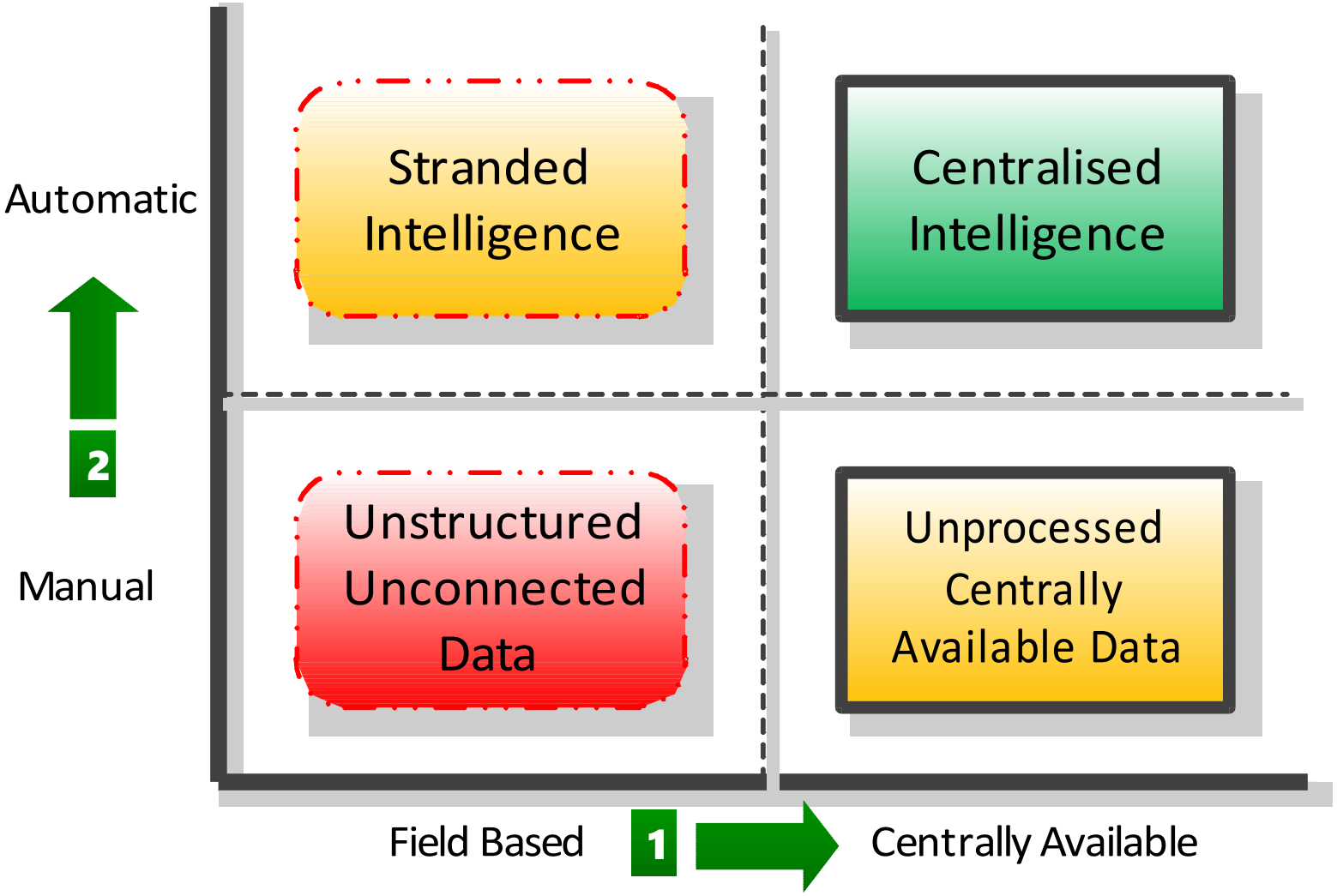
Controller Name	Mode	SP	OP	PV	% Manual	Variability	Alert
ETP-FIC 282005	Manual (Unlocked)	450.00	-5.00	0.57	14.0 %	-32.1 %	
ETP-FIC 282015	Auto (Unlocked)	590.00	54.06	550.05	84.7 %	-2.9 %	
ETP-PIC 240031A	Auto (Unlocked)	15.00	27.07	15.05	1.1 %	0.1 %	
ETP-PIC 240031B	Auto (Unlocked)	22.00	0	15.05	0.0 %	-6.9 %	
ETP-PIC 241033A	Auto (Unlocked)	14.00	80.00	14.25	2.8 %	0.3 %	
ETP-PIC 241033B	Auto (Unlocked)	20.00	-5.00	14.18	0.0 %	-5.8 %	
ETP-PIC 242043A	Manual (Locked)	14.40	Under Range	Under Range	100.0 %	Calc Failed %	
ETP-PIC 242043B	Auto (Unlocked)	17.00	-2.00	Under Range	0 %	Calc Failed %	
ETP-PIC 244032A	Auto (Unlocked)	14.28	46.74	14.83	2.4 %	0.5 %	
ETP-PIC 244032B	Auto (Unlocked)	22.00	-5.00	14.58	0.0 %	-7.4 %	
ETP-PIC 251044A	Auto (Unlocked)	14.00	8.68	14.04	4.9 %	0.0 %	
ETP-PIC 251044B	Auto (Unlocked)	20.00	-2.00	14.03	0.1 %	-6.0 %	
ETP-PIC 272040	No Data	20.00	0	13.03	Pt Created %	-23.2 %	
ETP-PIC 273011	Auto (Unlocked)	21.00	0	0.38	0 %	-20.6 %	
ETP-PIC 280008	Auto (Unlocked)	14.00	-2.00	11.97	0.0 %	-2.0 %	
ETP-PIC 340017	Auto (Unlocked)	58.00	0	49.56	0 %	-8.4 %	
ETP-PIC 840326	Auto (Unlocked)	14.20	-5.00	14.00	0 %	-0.2 %	
ETP-TIC 510001	Manual (Locked)	0	-5.00	8.67	100.0 %	8.7 %	
ETP-TIC 862006	No Data	I/O Timeout	I/O Timeout	I/O Timeout	Pt Created %	Calc Failed %	

Recap

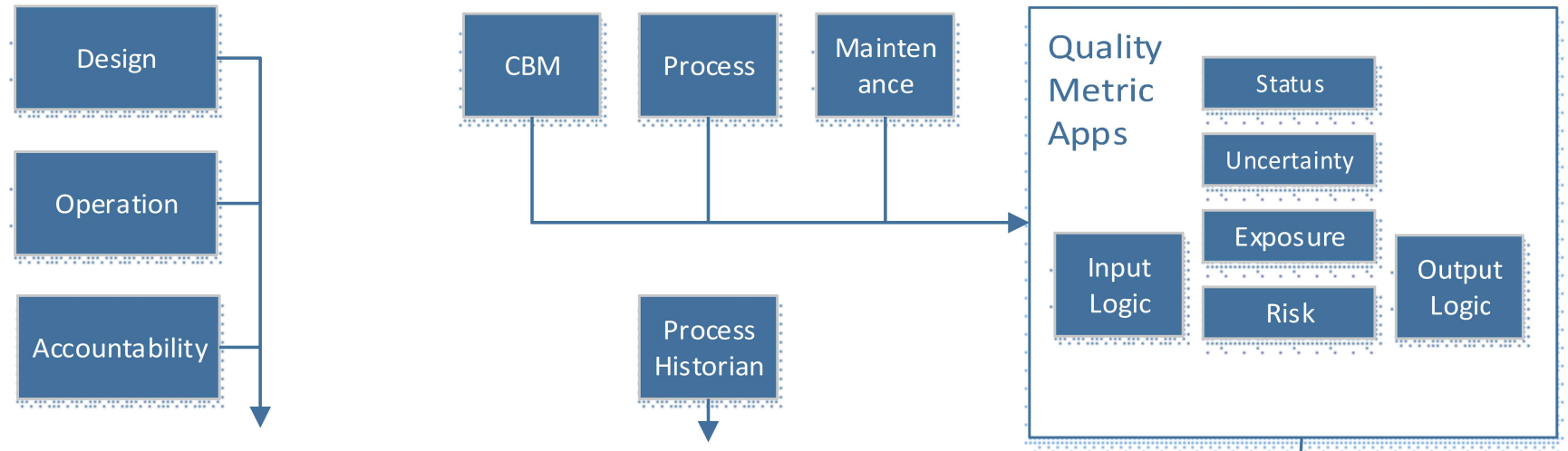


Recap

Recap

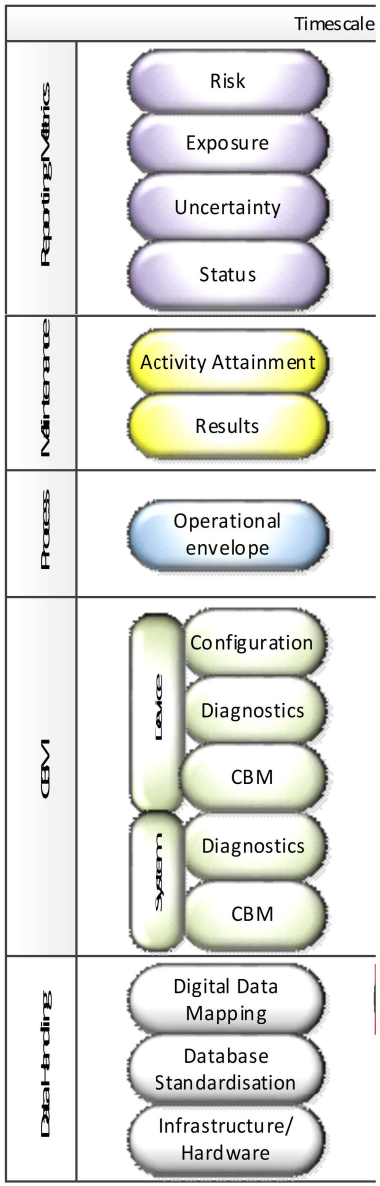

Automatic
2
Manual
Stranded Intelligence
Centralised Intelligence
Unstructured Unconnected Data
Unprocessed Centrally Available Data
Field Based
1
Centrally Available

Recap



Region	Site	System/ Station	Fluid Type	Total Vol	Total Std Vol	Total Mass	Total BOE	Risk	Exposure	Uncert. %	Status
Region 1	Site 1	Sys 1	Oil	Total Vol	Total Std Vol	Total Mass	Total BOE	Risk	Exposure	Uncert. %	Status

Recap


Timescale
Reporting/Metrics
Risk
Exposure
Uncertainty
Status
Maintenance
Activity Attainment
Results
Process
Operational envelope
CBM
Lead
Configuration
Diagnostics
CBM
System
Diagnostics
CBM
Data Handling
Digital Data Mapping
Database Standardisation
Infrastructure/ Hardware

Questions?

Technically Defensible
Measurement
Intelligence