Data quality improvement

Speaker

- Simen Sandelien
- Data quality specialist
- DNV GL
- Simen.Sandelien@dnvgl.com

DNV·GL

Topics

- **DNV GL**

- **Digital Class and data historian infrastructures**

- **Data quality perspectives**

- **Sensor system data quality foundations**

- **Recommendations**

- **Questions**

Commercial in confidence

OUR PURPOSE

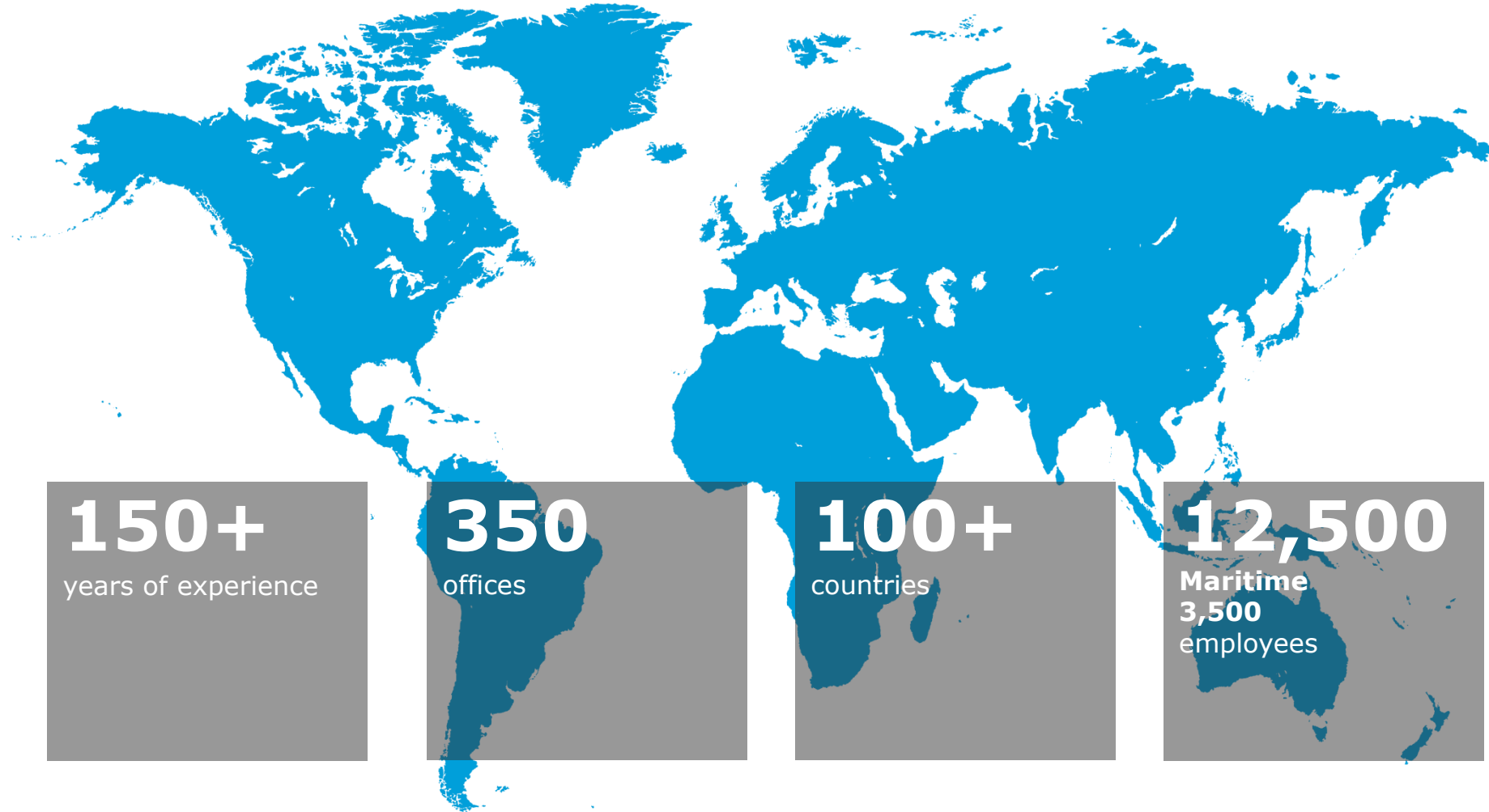# TO SAFEGUARD LIFE, PROPERTY AND THE ENVIRONMENT

Commercial in confidence

DNV·GL

Highly skilled people across the world

150+
years of experience

350
offices

100+
countries

12,500
Maritime
3,500
employees

DNV·GL

Our vision: global impact for a safe and sustainable future

MARITIME

OIL & GAS

POWER & RENEWABLES
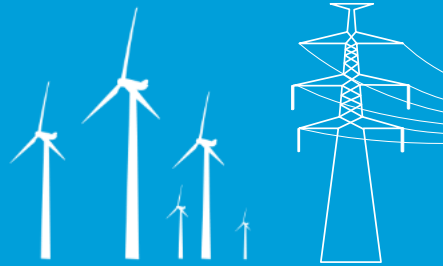
BUSINESS ASSURANCE

DIGITAL SOLUTIONS

Technology & Research

Global Shared Services

Commercial in confidence

Digital solutions for managing risk, improving safety and performance

We support business-critical activities across industries, including maritime, oil and gas, energy and healthcare

| Industry software solutions | Data management and analytics | Consulting and advisory services | Industry data platform | Cyber security |

Commercial in confidence

DNV·GL

Digital Class drivers
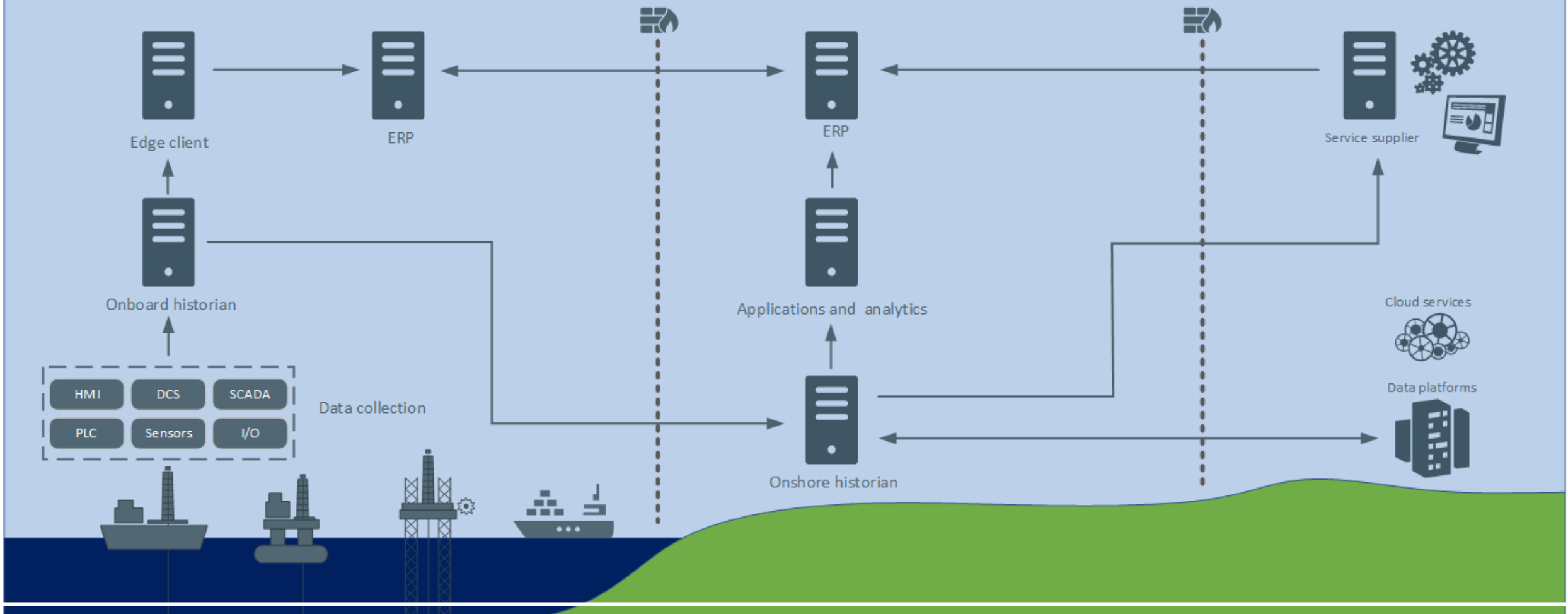
DNV·GL

Complicated data chains, corporate interdependencies and technical options
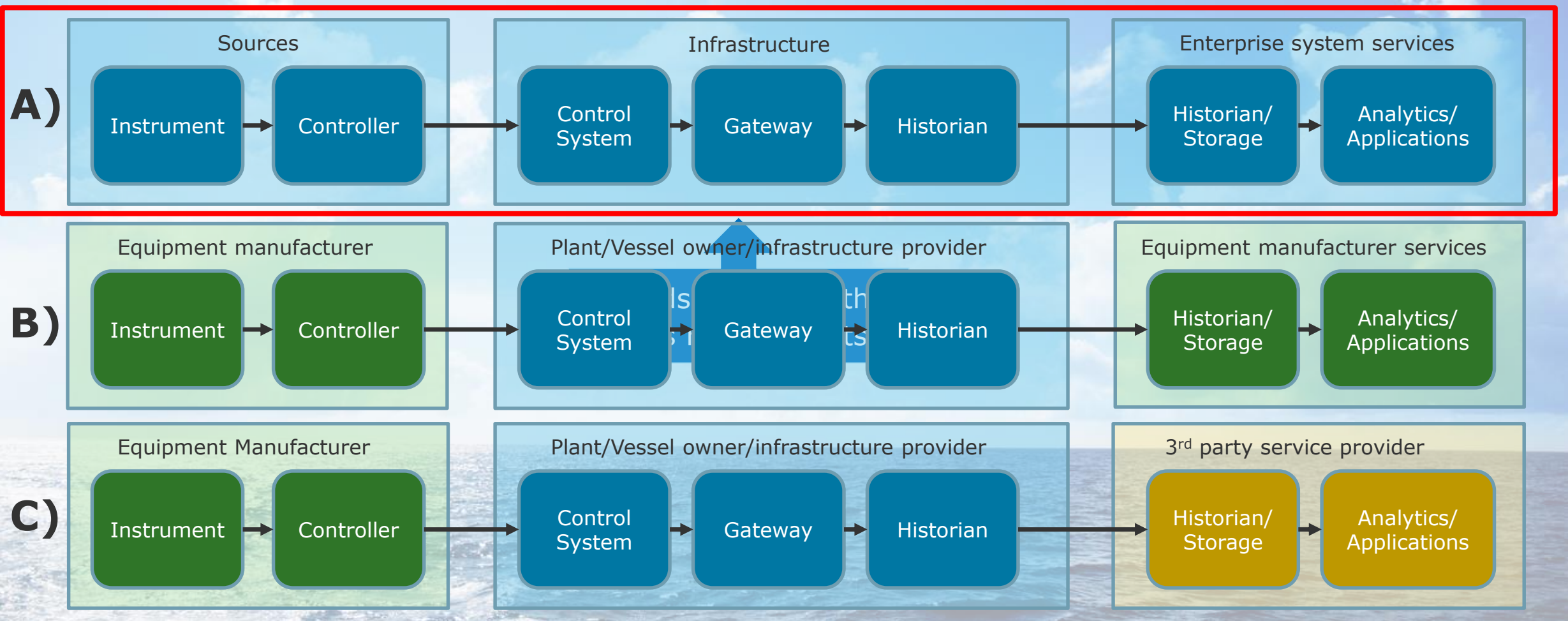
Commercial in confidence

DNV·GL

DNV GL customer benefits of qualifying products like the OSIsoft PI System

Collaboration between stakeholders in the industry to tackle common issues

Reducing complexity, centralize and increase standardization

Improving cost efficiency and **lower the OPEX**

Optimization by utilizing new technology and **capture the value in data**

Commercial in confidence

DNV·GL

Digital trust

Managing risks and opportunities of increasingly data-driven operations and finding new ways of building trust

Data quality

Data management maturity

Data flow and sensor systems

ISO 8000-8

DNV·GL

Data quality perspectives

Macro:
- Organizational processes and adherence to quality systems and business goals

Meso level:
- Architecture/integration
- Planning and tools

Micro:
- Connectivity issues
- Dead/stale points

Commercial in confidence

DNV·GL

Data quality - Day to day issues

Recurring data quality issues never fully resolved

Degradation of metadata over time

Poor data quality not sufficiently understood by data consumers

Commercial in confidence

DNV·GL

Data quality - Big incidents

Data collection stopped working for several hours at the worst possible time

Permanent loss of years of plant data during a software upgrade

Equipment monitoring package rendered useless due to poor event quality

Predictive maintenance and analytics projects failed due to data quality issues

Commercial in confidence

DNV·GL

Data quality - Data based projects challenges

Importance of contribution

Data relevant for the task

Unidentified potential

Known unavailable

Known insufficient

Known good

a b c d e f g h i j k l m n o p q r s u v w x y z
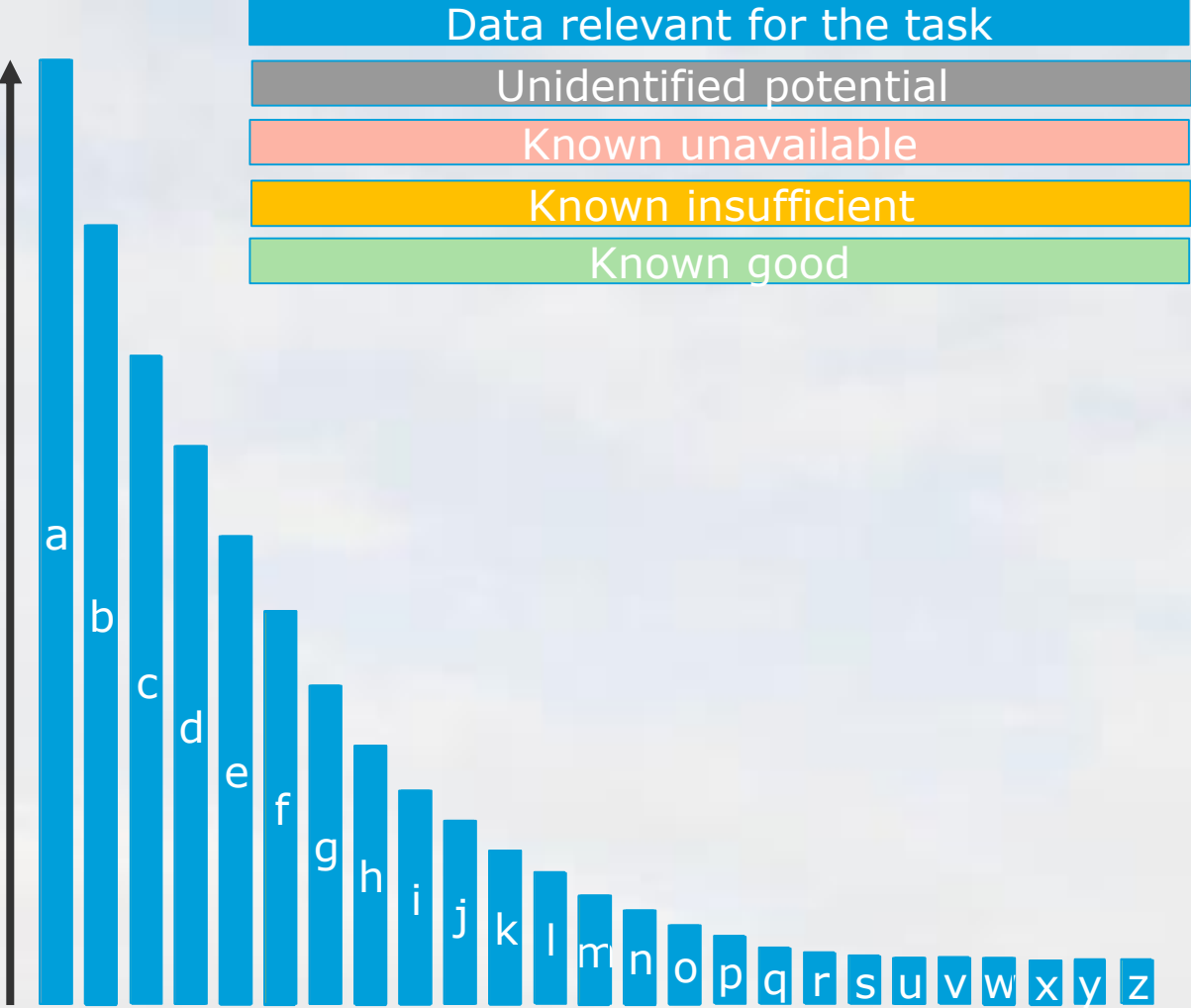
$$Y(t) = at+bt+ct+ dt+et+ft+ gt+ht+ it+ jt+kt+ lt+ mt+nt+ot+pt+qt+ rt+st+ut+vt+wt+ xt+yt +zt$$

DNV·GL

Raise your hand:

Who has started working on data quality improvement due to such issues?

Sensor system data quality dimensions

Governance	Organization/people	Processes	Efficiency	Requirement definition	Metrics	Tools and technologies		Data standards
Management focus	End user training	End user work processes	Historization pipeline extent	Legacy system debt	End user quality feedback	Calibration	Sufficient instrumentation	Plant model synchronization and scaling
	System owner training	System owner work processes	Aggregation or sampling freq. loss	Maintenance and service contracts	Heartbeat signals	Structured I/O federation (DCS)	Sufficient granularity of system flags	Meta data synchronization across layers
	Infrastr. and MES training	Infrastr. & MES system work processes	Latency prevention	Upgrade and refitting practices	Asset management systems	I/O Control module load	Sufficient HW/server capacity	Inconsistent data transfer standards and schemas
	ICSS training	ICSS work processes	Design redundancy	Initial planning and development practices	Alarm monitoring and statistics	I/O Control Loop Quality	Bandwidth	Single point of truth
	Instrument training	Instrument work processes	Dynamic and flexible design		Infrastructure service monitoring	Intrinsic instrumentation qualities		
			Graceful buffering and recovery			Time synch-ronization		

Legend:
- Semantic
- Syntactic
- Pragmatic

Completeness

Commercial in confidence

DNV·GL

Data quality - Day to day issues

Recurring data quality issues never fully resolved

Degradation of metadata over time

Poor data quality not sufficiently understood by data consumers

DNV·GL

Recurring data quality issues never fully resolved

Governance	Organization/people	Processes	Efficiency	Requirement definition	Metrics	Tools and technologies		Data standards
Management focus	End user training	End user work processes	Historization pipeline extent	Legacy system debt	End user quality feedback	Calibration	Sufficient instrumentation	Plant model synchronization and scaling
	System owner training	System owner work processes	Aggregation or sampling freq. loss	Maintenance and service contracts	Heartbeat signals	Structured I/O federation (DCS)	Sufficient granularity of system flags	Meta data synchronization across layers
	Infrastr. and MES training	Infrastr. & MES system work processes	Latency prevention	Upgrade and refitting practices		I/O Control	Sufficient HW/server capacity	Inconsistent data transfer standards and schemas
	ICSS	ICSS processes		Initial planning	Alarm and sta...		...andwidth	Single point of truth
			...flex...		Infrastructu... service monitoring		qualities	
			Gr... buffering and recovery				Time synch-ronization	

Legend:
- Semantic
- Syntactic
- Pragmatic

Infrastr. and MES system training

Infrastr. & MES system work processes

Maintenance and service contracts

Commercial in confidence

DNV·GL

Data quality - Day to day issues
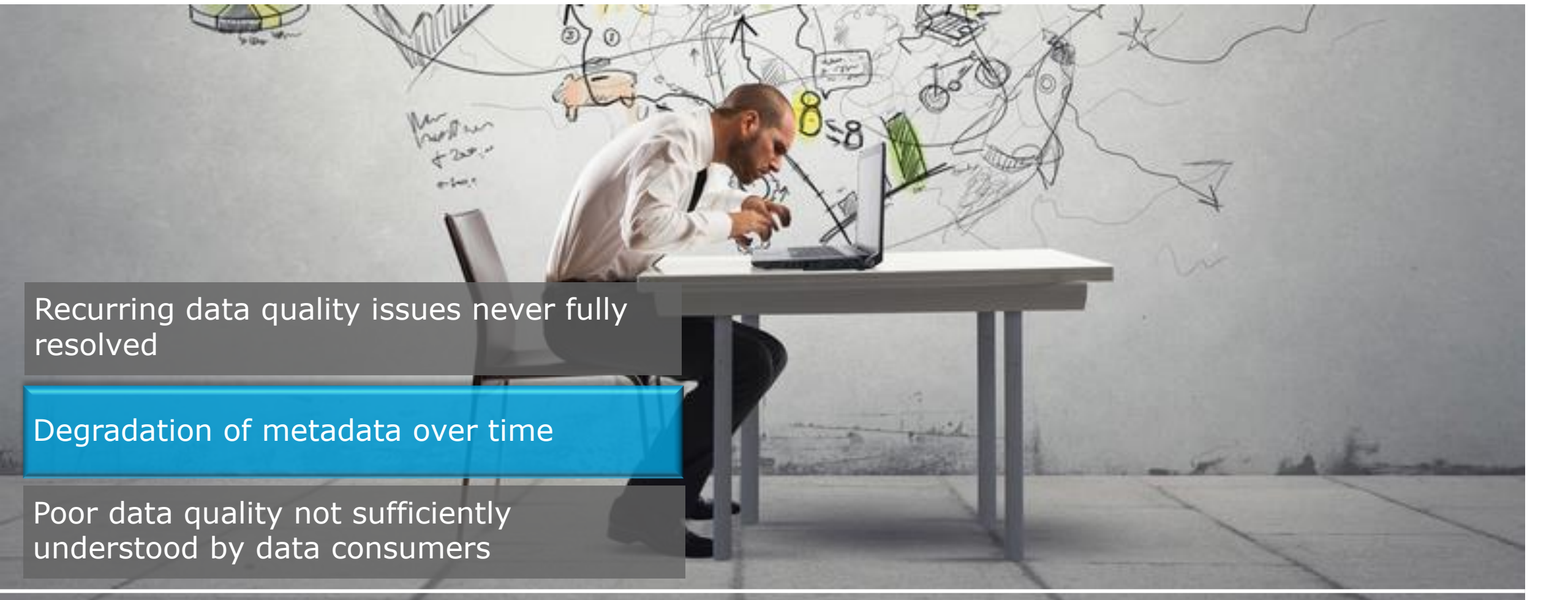

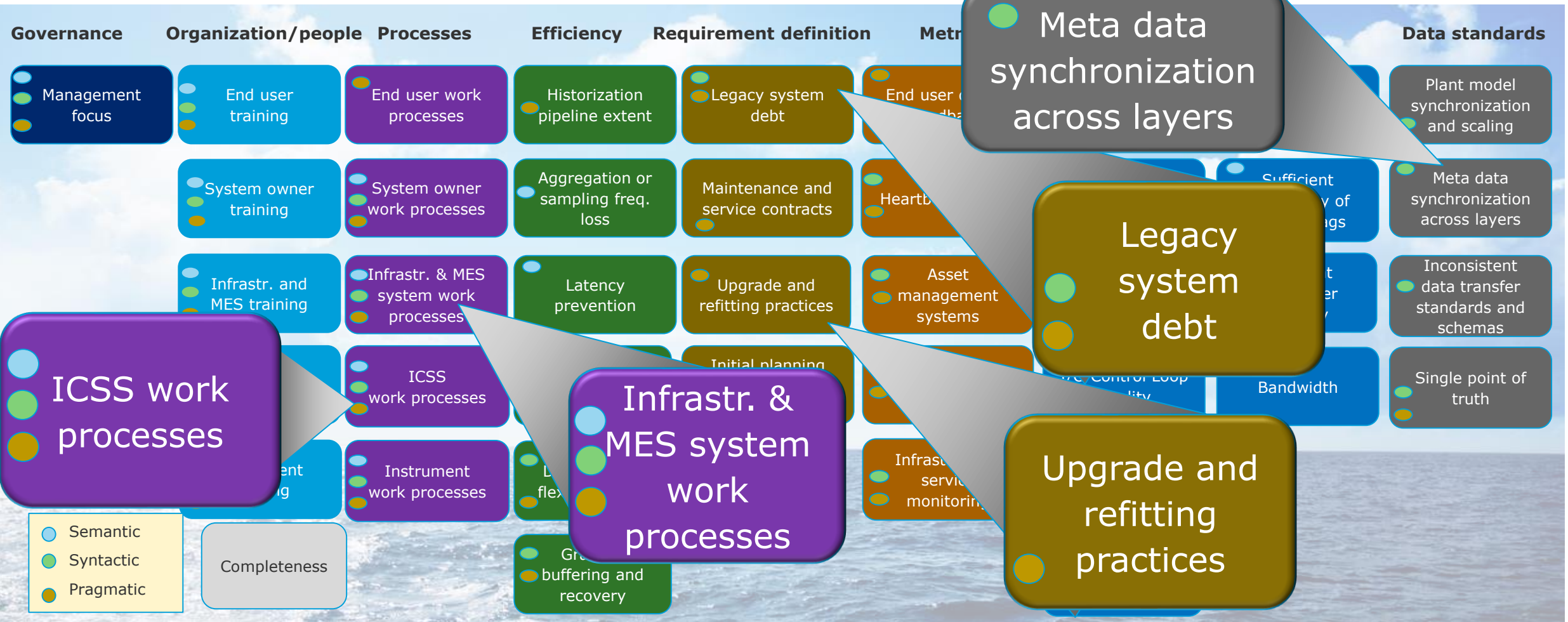Recurring data quality issues never fully resolved

Degradation of metadata over time

Poor data quality not sufficiently understood by data consumers

Commercial in confidence

DNV·GL

Degradation of metadata over time

Governance	Organization/people	Processes	Efficiency	Requirement definition	Metr...	Data standards
Management focus	End user training	End user work processes	Historization pipeline extent	Legacy system debt	End user ...dba...	Plant model synchronization and scaling
	System owner training	System owner work processes	Aggregation or sampling freq. loss	Maintenance and service contracts	Heartb...	Meta data synchronization across layers
Infrastr. and MES training	Infrastr. & MES system work processes	Latency prevention	Upgrade and refitting practices	Asset management systems	...t ...er ...y	Inconsistent data transfer standards and schemas
	ICSS work processes		Initial planning	I/C Control Loop ...ity	Bandwidth	Single point of truth
...ent ...g	Instrument work processes	...D... flex...		Infras... servic... monitorin...		
		Gr... buffering and recovery				

Legend:
- ● Semantic
- ● Syntactic
- ● Pragmatic

Completeness

Callout boxes:
- ● ICSS work processes
- ● Infrastr. & MES system work processes
- ● Meta data synchronization across layers
- ● Legacy system debt
- ● Upgrade and refitting practices

Commercial in confidence

DNV·GL

Data quality - Day to day issues

Recurring data quality issues never fully resolved

Degradation of metadata over time

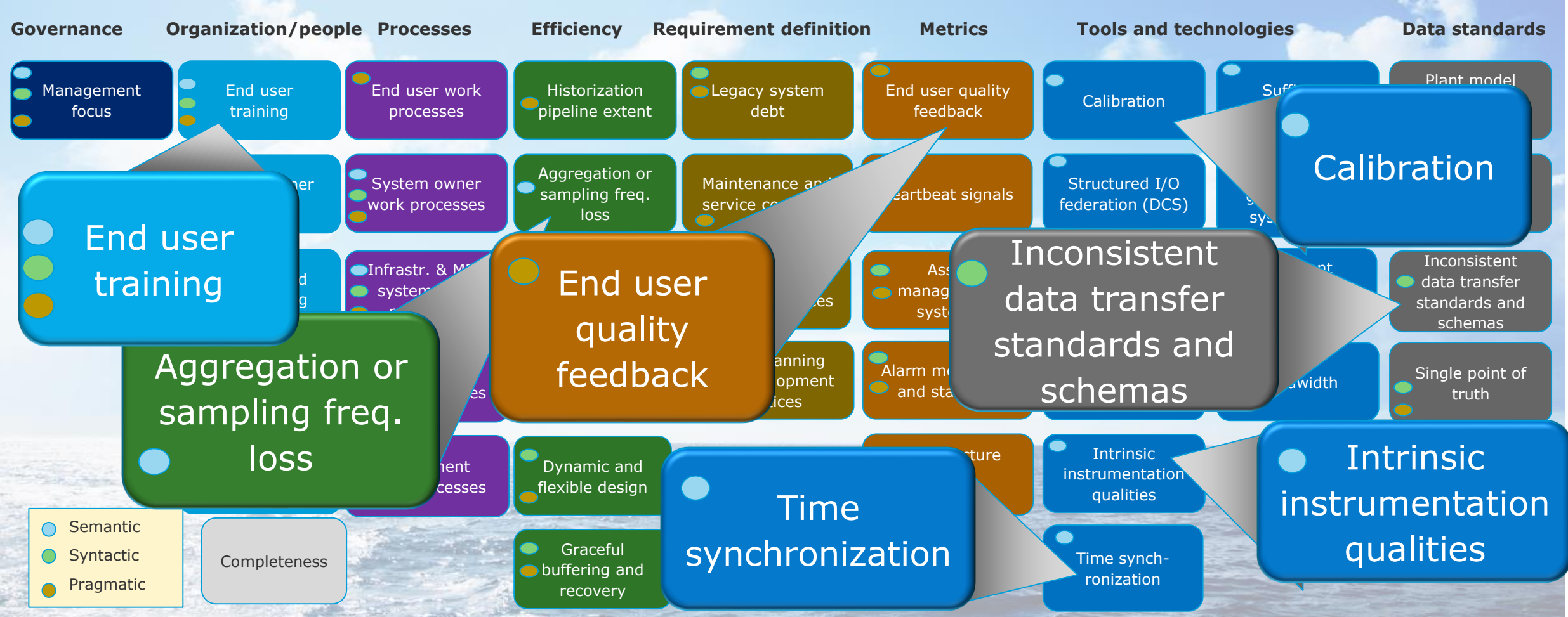Poor data quality not sufficiently understood by data consumers

DNV·GL

Poor data quality not sufficiently understood by data consumers

Governance	Organization/people	Processes	Efficiency	Requirement definition	Metrics	Tools and technologies		Data standards
Management focus	End user training	End user work processes	Historization pipeline extent	Legacy system debt	End user quality feedback	Calibration	Suffi...	Plant model
...ner	System owner work processes	Aggregation or sampling freq. loss	Maintenance and service co...	...eartbeat signals	Structured I/O federation (DCS)	...g sys...		
...d ...g	Infrastr. & M... system...			Ass... manag... syst...			Inconsistent data transfer standards and schemas
	...ces	...anning ...opment ...ices	Alarm m... and sta...	...width		Single point of truth
	...ent ...cesses	Dynamic and flexible design	...cture	Intrinsic instrumentation qualities		
		Graceful buffering and recovery		Time synch-ronization		

End user training

Aggregation or sampling freq. loss

End user quality feedback

Inconsistent data transfer standards and schemas

Calibration

Intrinsic instrumentation qualities

Time synchronization

- ○ Semantic
- ○ Syntactic
- ○ Pragmatic

Completeness

DNV·GL

Data quality - Big incidents

Data collection stopped working for several hours at the worst possible time.
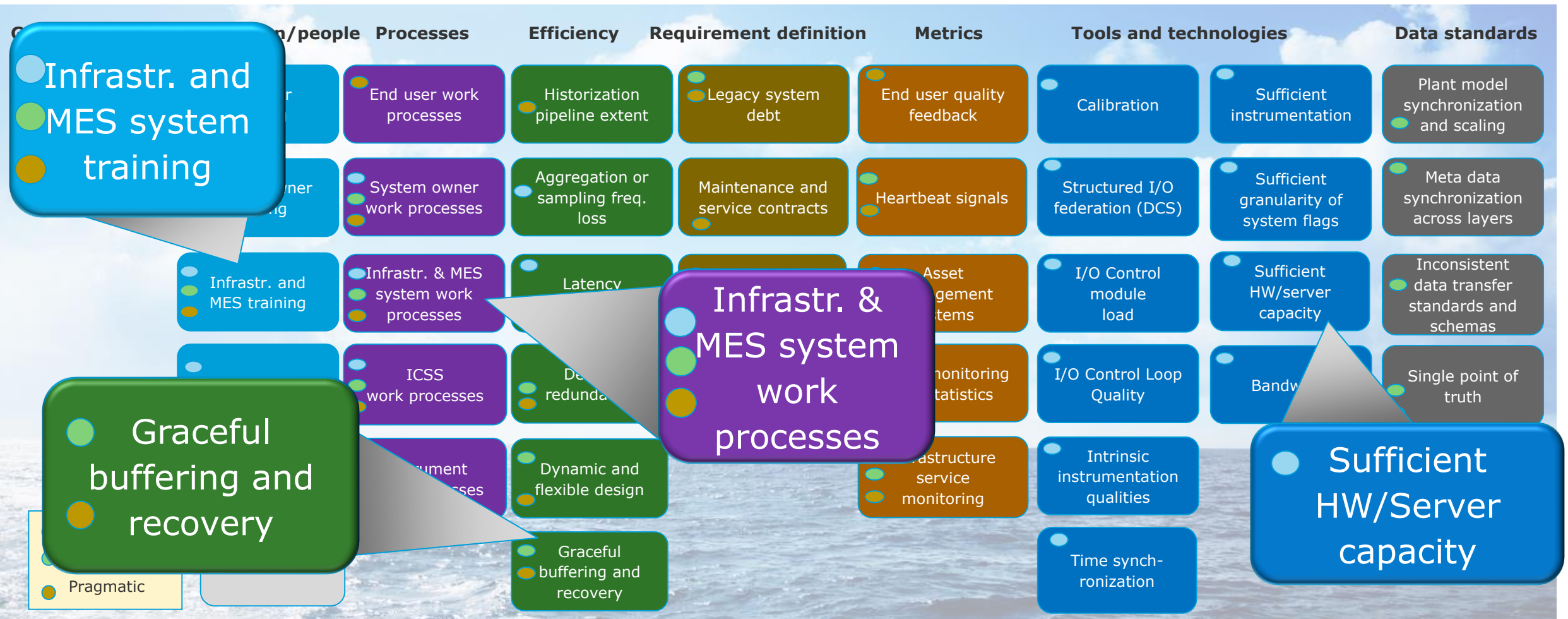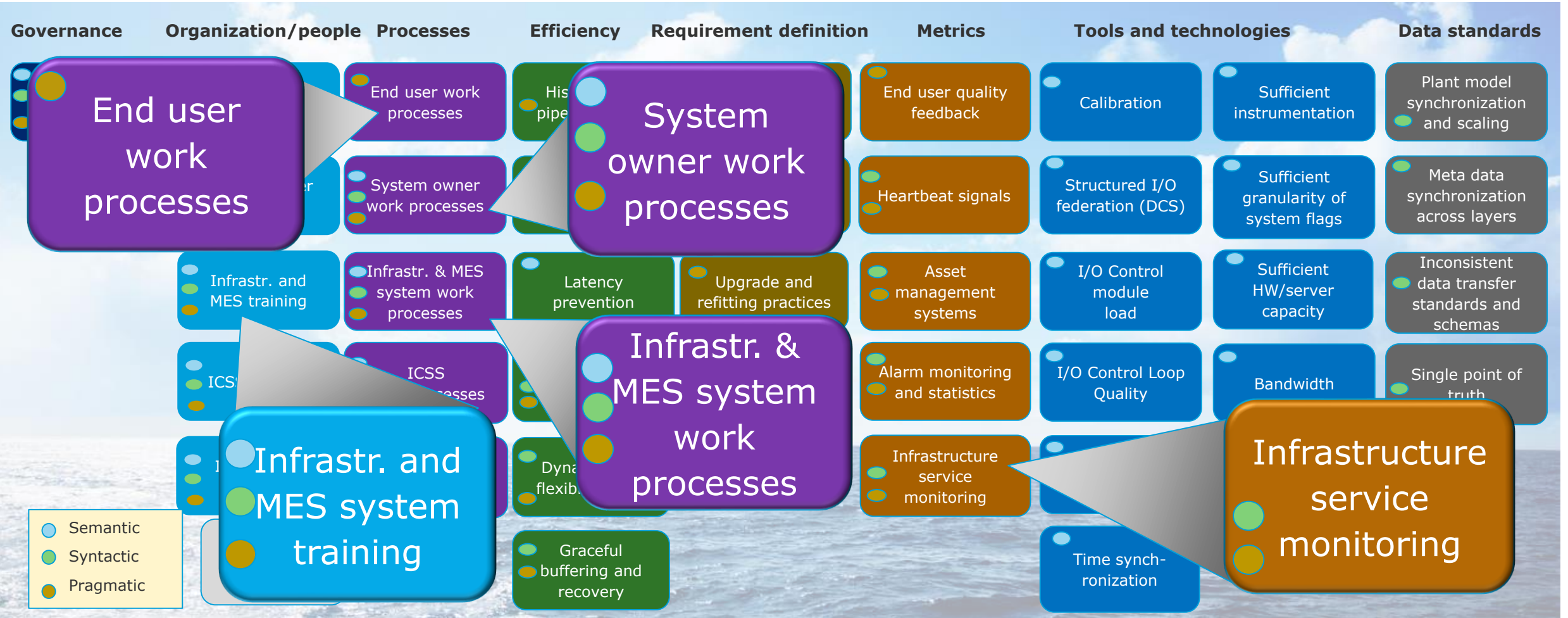
Permanent loss of years of plant data during a software upgrade

Equipment monitoring package rendered useless due to poor event quality

Predictive maintenance and analytics projects failed due to data quality issues

Commercial in confidence

DNV·GL

Data collection stopped working for several hours at a critical time

Commercial in confidence

DNV·GL

Data quality - Big incidents

Data collection stopped working for several hours at the worst possible time
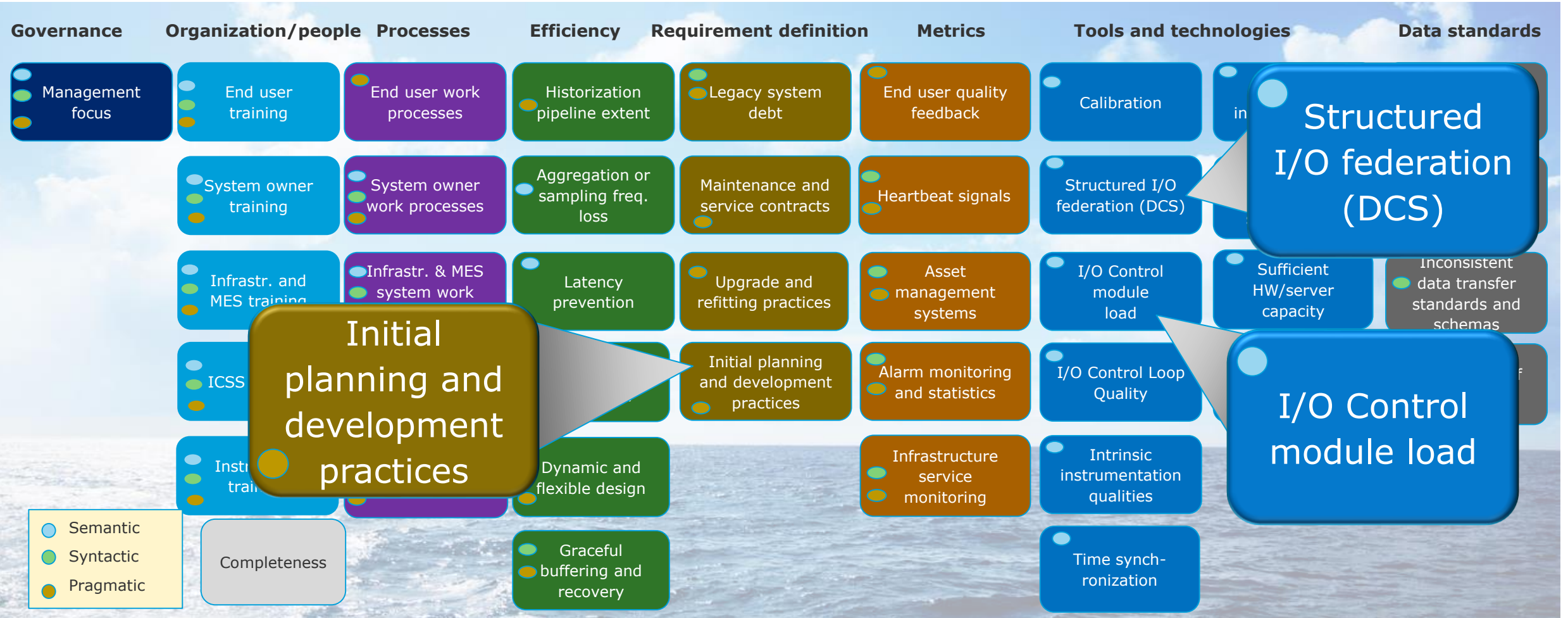
Permanent loss of years of plant data during a software upgrade

Equipment monitoring package rendered useless due to poor event quality

Predictive maintenance and analytics projects failed due to data quality issues

Commercial in confidence

DNV·GL

Permanent loss of plant data during software upgrade

Governance	Organization/people	Processes	Efficiency	Requirement definition	Metrics	Tools and technologies		Data standards
End user work processes	End user work processes	His pipe		**System owner work processes**	End user quality feedback	Calibration	Sufficient instrumentation	Plant model synchronization and scaling
	System owner work processes				Heartbeat signals	Structured I/O federation (DCS)	Sufficient granularity of system flags	Meta data synchronization across layers
	Infrastr. and MES training	Infrastr. & MES system work processes	Latency prevention	Upgrade and refitting practices	Asset management systems	I/O Control module load	Sufficient HW/server capacity	Inconsistent data transfer standards and schemas
	ICSS	ICSS processes	**Infrastr. & MES system work processes**		Alarm monitoring and statistics	I/O Control Loop Quality	Bandwidth	Single point of truth
	Infrastr. and MES system training		Dyna flexib		Infrastructure service monitoring			**Infrastructure service monitoring**
			Graceful buffering and recovery			Time synch-ronization		

Legend:
- ● Semantic
- ● Syntactic
- ● Pragmatic

DNV·GL

Data quality - Big incidents

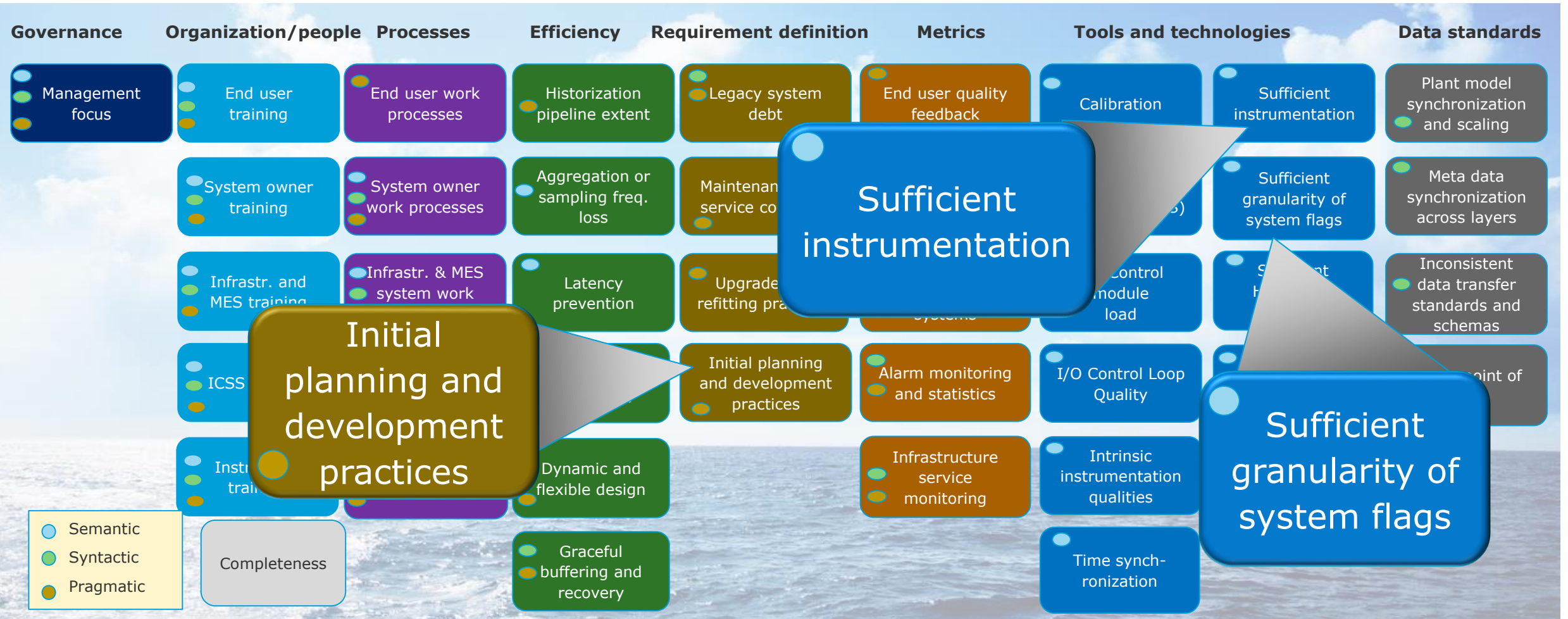Data collection stopped working for several hours at the worst possible time

Permanent loss of years of plant data during a software upgrade

Equipment monitoring package rendered useless due to poor event quality

Predictive maintenance and analytics projects failed due to data quality issues

DNV·GL

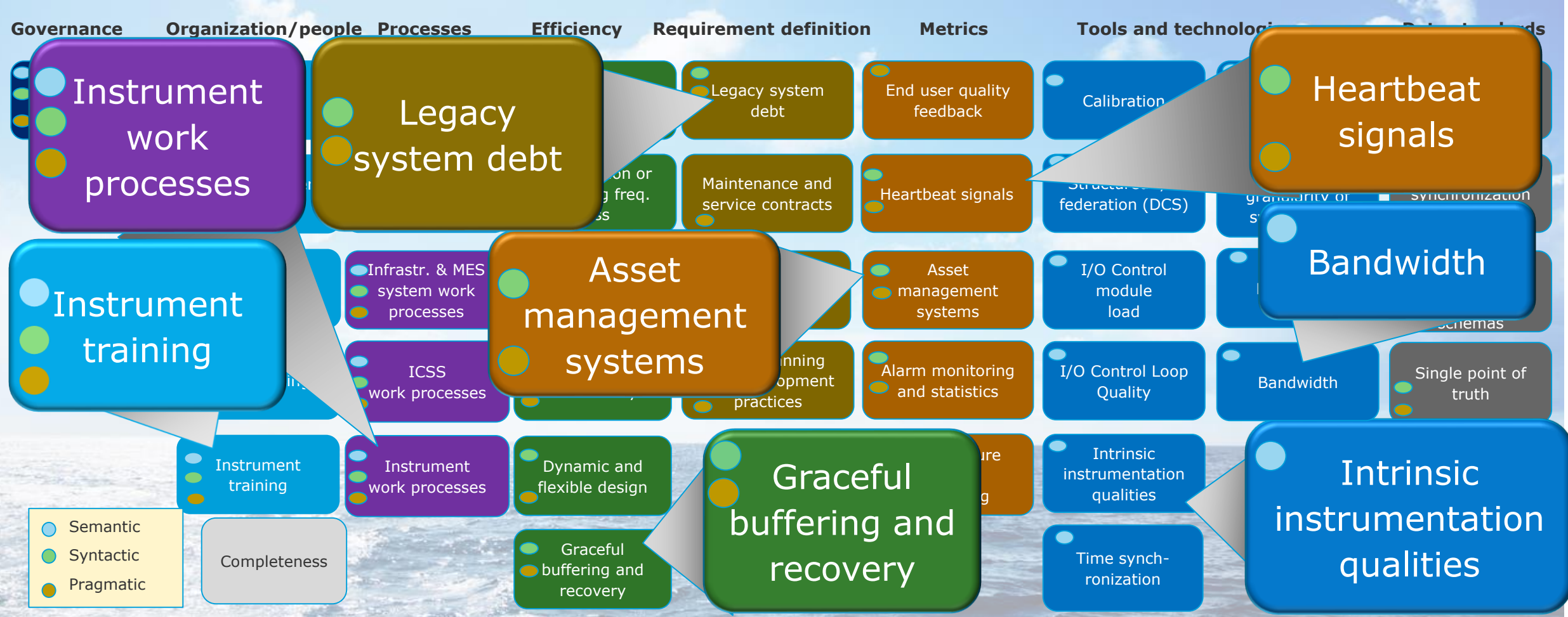# Equipment monitoring package rendered useless

Governance	Organization/people	Processes	Efficiency	Requirement definition	Metrics	Tools and technologies	Data standards
Management focus	End user training	End user work processes	Historization pipeline extent	Legacy system debt	End user quality feedback	Calibration	in...
	System owner training	System owner work processes	Aggregation or sampling freq. loss	Maintenance and service contracts	Heartbeat signals	Structured I/O federation (DCS)	
	Infrastr. and MES training	Infrastr. & MES system work	Latency prevention	Upgrade and refitting practices	Asset management systems	I/O Control module load	Sufficient HW/server capacity / Inconsistent data transfer standards and schemas
	ICSS		Initial planning and development practices		Alarm monitoring and statistics	I/O Control Loop Quality	
	Instr... trai...		Dynamic and flexible design		Infrastructure service monitoring	Intrinsic instrumentation qualities	
		Completeness	Graceful buffering and recovery			Time synch-ronization	

Legend:
- Semantic
- Syntactic
- Pragmatic

Callouts:
- Structured I/O federation (DCS)
- I/O Control module load
- Initial planning and development practices

Commercial in confidence

DNV·GL

Data quality - Big incidents

Data collection stopped working for several hours at the worst possible time

Permanent loss of years of plant data during a software upgrade

Equipment monitoring package rendered useless due to poor event quality

Predictive maintenance and analytics projects failed due to data quality issues - **unavailable**

DNV·GL

Predictive analytics projects fail due to poor data quality – unavailable data

Governance	Organization/people	Processes	Efficiency	Requirement definition	Metrics	Tools and technologies	Data standards
Management focus	End user training	End user work processes	Historization pipeline extent	Legacy system debt	End user quality feedback	Calibration	Plant model synchronization and scaling
	System owner training	System owner work processes	Aggregation or sampling freq. loss	Maintenance service co...		Sufficient instrumentation	Meta data synchronization across layers
	Infrastr. and MES training	Infrastr. & MES system work	Latency prevention	Upgrade refitting pra...		Control module load	Sufficient granularity of system flags
	ICSS		Initial planning and development practices		Alarm monitoring and statistics	I/O Control Loop Quality	Inconsistent data transfer standards and schemas
	Instr... train...		Dynamic and flexible design		Infrastructure service monitoring	Intrinsic instrumentation qualities	...int of...
			Graceful buffering and recovery			Time synch-ronization	

Sufficient instrumentation

Initial planning and development practices

Sufficient granularity of system flags

Legend:
- Semantic
- Syntactic
- Pragmatic

Completeness

DNV·GL

Data quality - Big incidents

Data collection stopped working for several hours at the worst possible time

Permanent loss of years of plant data during a software upgrade

Equipment monitoring package rendered useless due to poor event quality

Predictive maintenance and analytics projects failed due to data quality issues - **insufficient**

DNV·GL

Predictive analytics projects fail due to poor data quality – insufficient data

Instrument work processes

Legacy system debt

Heartbeat signals

Instrument training

Asset management systems

Bandwidth

Graceful buffering and recovery

Intrinsic instrumentation qualities

Legacy system debt

End user quality feedback

Calibration

...on or ...g freq.

Maintenance and service contracts

Heartbeat signals

Structures / federation (DCS)

granularity of s...

synchronization

Infrastr. & MES system work processes

Asset management systems

I/O Control module load

ICSS work processes

...nning ...opment practices

Alarm monitoring and statistics

I/O Control Loop Quality

Bandwidth

Single point of truth

Instrument training

Instrument work processes

Dynamic and flexible design

Intrinsic instrumentation qualities

Graceful buffering and recovery

Time synch-ronization

- Semantic
- Syntactic
- Pragmatic

Completeness

Generic recommendations

- Establish data requirements early in new builds

- Include data quality KPIs in service contracts

- Facilitate cross discipline communication

- Have good data quality feedback systems

- Data is an asset. Think condition monitoring

Commercial in confidence

DNV·GL

Specific recommendations – based on advisory services

Define scope

- Intended use
- Context

- Datasets & criticality
- Data origin and path

Data exploration and profiling
Get access to data, prepare, explore

Continuous improvement cycle

Data quality assessment
- Identify and define req.
- Prepare and configure
- Perform assessment

Organizational maturity assessment
- Organizational scope
- Interview & review
- Analyse and score

Sensor system assessment
- Clarify scope
- Choose methodology
- Analyse and score

Algorithm assessment
- Clarify scope
- Choose methodology
- Analyse and score

Extended scope

Data usage risk assessment
Fit for use within acceptable risk
- Define scope
- Consequence dimensions
- Perform risk assessment

- Root causes
- Measures
- Prioritize

Risk based data quality improvement
- Data
- IT Systems

- Processes
- Organization

DNV·GL

Organizational Maturity Assessment

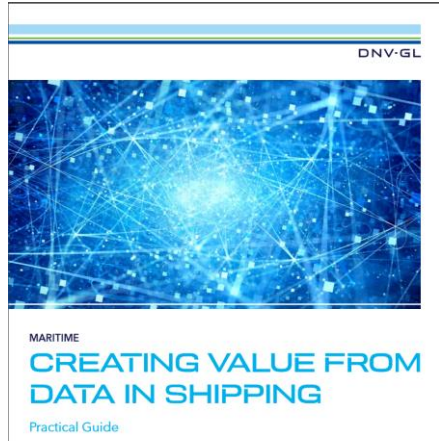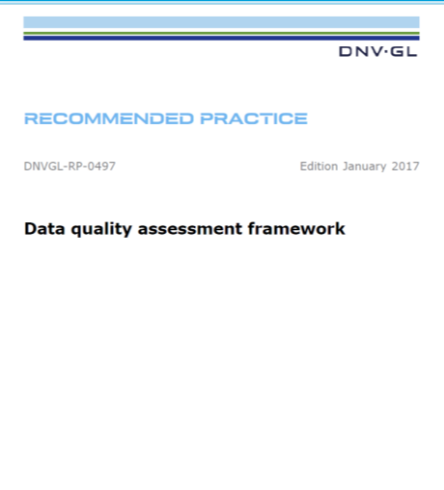| Maturity level	Governance	Organization and people	Processes	Process Efficiency	Requirement definition	Metrics and dimensions	Architecture, tools and technologies	Data standards
LEVEL 5 - Optimized	Data management policies governs and drives improvements	Data management board oversees improvement activities	Processes for continuous improvement in place	Processes provides feed-back and feed-forward to support continuous improvement	Baseline established and improvements measured according to requirements	Metrics defines baseline to support continuous improvement	Tools support policy driven continuous improvement cycle	Standard compliance and domain models are subject to continuous improvement
LEVEL 4 – Managed	Policies defined in relation to business objectives	Skillset extended to include risk analysis of quality issues aligned with business objectives	Processes for impact analysis and risk mgmt. in place	Monitoring is performed across enterprise and published as KPI's and trends	Requirements are linked to business impacts	Metrics are linked to business impacts and risk analysis	Tools are driven by business objectives and include support for root cause analysis and risk mgmt.	Standards are used actively to reduce risk for critical business operations
LEVEL 3 – Defined	Policies defined at enterprise level	Roles and required skills defined at enterprise level	Processes are defined and implemented consistently across enterprise	Defined metrics are monitored in advance of business impact	Requirements defined and communicated at enterprise level	Framework for metrics and dimensions defined at enterprise level	Architecture in place at enterprise level supporting full stack data management	Standards, domain models and semantics used at enterprise level
LEVEL 2 – Repeatable	Local initiatives address the requirement for policies	Locally defined roles and some basic skills	Best practices in place but not used consistently	Generic metrics are monitored at point of impact	Local initiatives define requirements	Metrics are reused locally in projects	Tools and technologies used consistently in selected projects	Industry standards and domain models used selectively across projects
LEVEL 1 - Initial	Only ad-hoc or temporal policies in place	No formally defined roles or skillset	Ad-hoc or reactive responses to quality issues	No baseline and no monitoring of quality issues	Re-engineering used to derive requirements	Project specific metrics	Tools are used ad-hoc per project	Ad-hoc and inconsistent use of standards
Objectives,	Policy, Culture, Awareness, Risks, Capabilities to handle DQ issues	Organization, roles, responsibilities, authority, skillsets	Structured and vetted ways of handling and preventing DQ issues	Measure, monitor and use metrics to mitigate DQ issues	DQ Requirements defined, communicated and acted upon	DQ metrics defined, setup, measured and monitored	DQ Tools for processing, analysing and correcting DQ issues with data assets	Use available standards, models, ontologies and taxonomies – a corporate «DQ language»

Commercial in confiden

DNV·GL

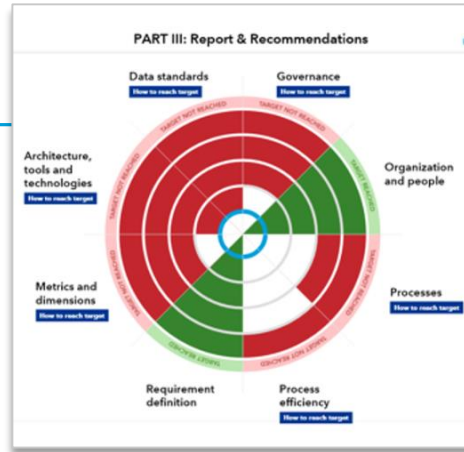# Risk based continuous improvement, prevent and fix

Threats

Vulnerability

Consequence

Data quality related incident

Preventive Measures (barriers)

Reactive Measures (barriers)
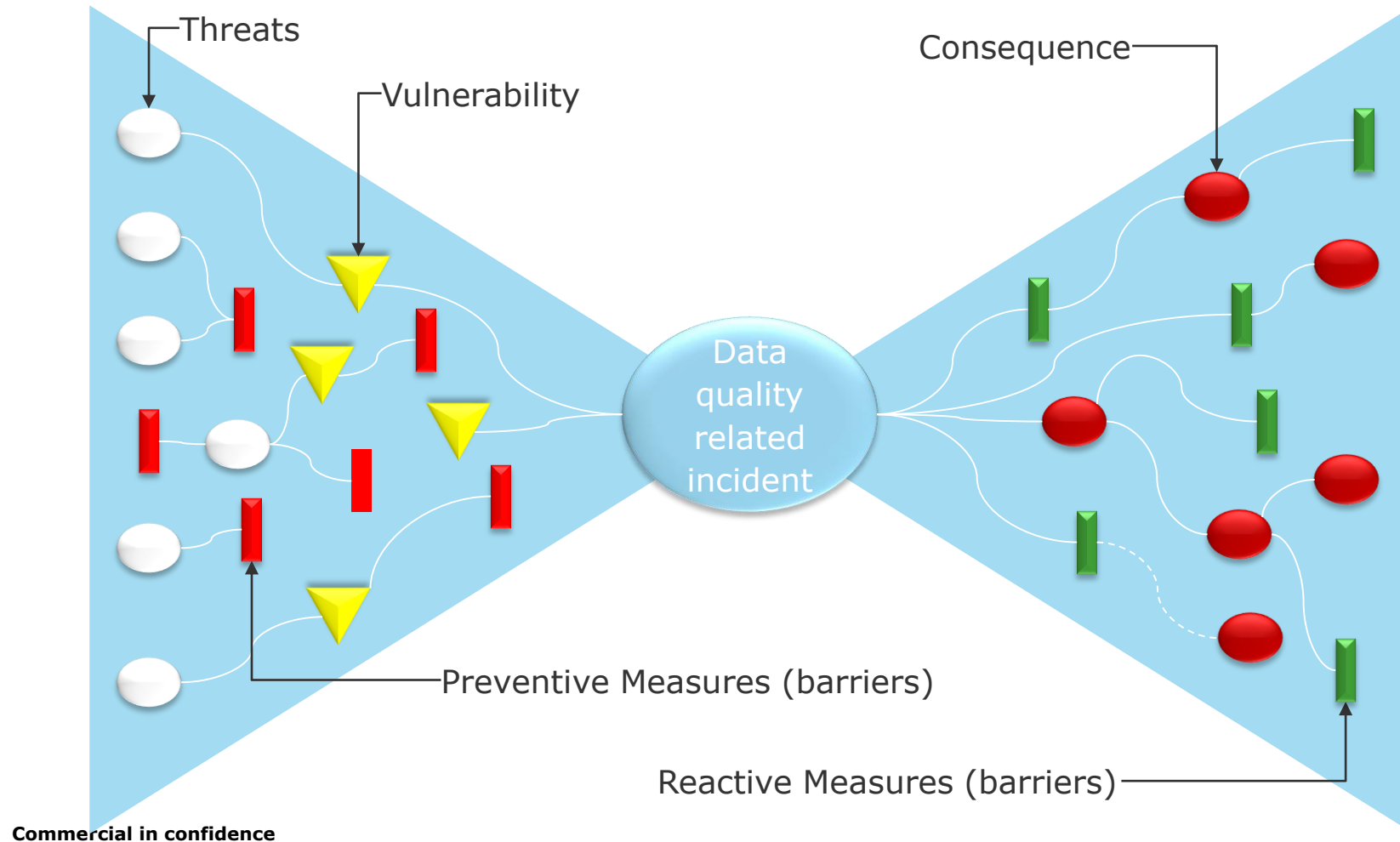
DNV·GL

Supporting documents

www.dnvgl.com

Commercial in confidence
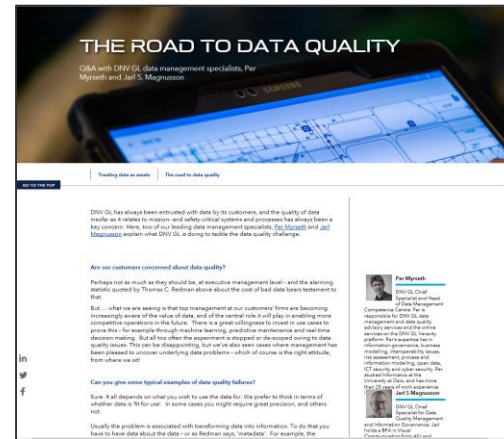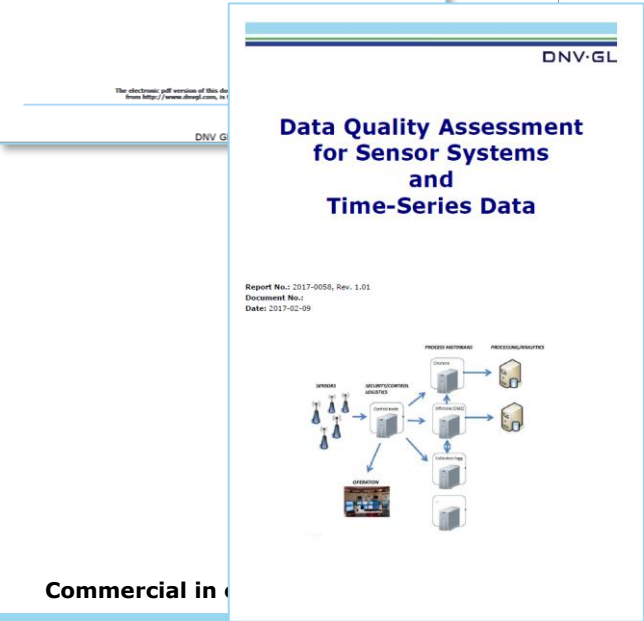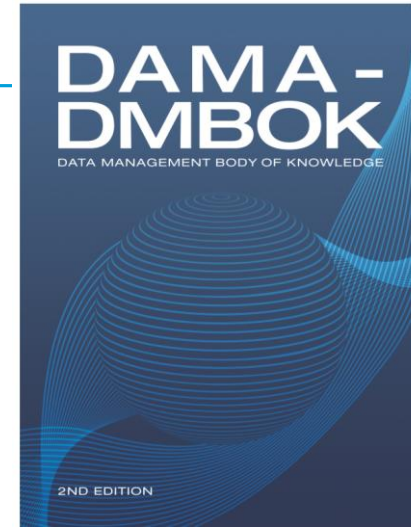
Data quality improvement

CHALLENGES

- Poor data quality negatively impacts business

- Efforts to achieve structured data quality improvement become either too ad-hoc or too generic

SOLUTION

- Relate data quality issues to familiar concepts and categories

- Target improvement efforts in dedicated areas

BENEFITS

- Increased value from data

- Alignment with digital class requirements

> " Investing in data quality is a prerequisite for successful data science. "
>
> Per Myrseth, Manager of Data management and analytics, DNV GL

Speaker

- Simen Sandelien
- Data quality specialist
- DNV GL
- Simen.Sandelien@dnvgl.com

DNV·GL

Questions?

Please wait for the **microphone**

State your **name & company**

Please remember to…

Complete Survey!
Navigate to this session in mobile agenda for survey

TO DOWNLOAD APP, SEARCH OSISOFT PI WORLD

Download on the App Store

GET IT ON Google Play

OSIsoft PI World