Guidance Notes On

Data Integrity for Marine and Offshore Operations

ABS CyberSafety™ Volume 3

September 2016

GUIDANCE NOTES ON

DATA INTEGRITY FOR MARINE AND OFFSHORE OPERATIONS SEPTEMBER 2016

ABS CYBERSAFETY™ VOLUME 3

American Bureau of Shipping
Incorporated by Act of Legislature of
the State of New York 1862

ABS Plaza
1701 City Plaza Drive
Spring, TX 77389 USA

Foreword

The marine and offshore industries are integrating connected sensors, communications, storage and processing capabilities into vessels, offshore units and facilities as networking and computational power penetrates all aspects of industry operations. The "Big Data" phenomenon has emerged as a direct result of this growth, enabling development of tremendous new sources of data and information. But challenges have also emerged. Sensors and data must be trustworthy in order to support the new analytic and decision methods available for maritime industry use.

These Guidance Notes are intended to clarify the basic principles and concepts of Data Integrity for marine and offshore assets. The document is intended to help the industry realize the new benefits from data sources and data analytics systems via implementation of Data Integrity concepts. It also supports owners who are increasingly required to provide data reporting to regulatory agencies. The intended users for these Guidance Notes are cybersecurity specialists, data specialists, owners, shipyards, operators, designers, suppliers, review engineers and Surveyors.

These Guidance Notes are Volume 3 of the ABS CyberSafety™ series, and are intended to be used in conjunction with other volumes.

These Guidance Notes become effective on the first day of the month of publication.

Users are advised to check periodically on the ABS website www.eagle.org to verify that this version of these Guidance Notes is the most current.

We welcome your feedback. Comments or suggestions can be sent electronically by email to rsd@eagle.org.

Terms of Use

The information presented herein is intended solely to assist the reader in the methodologies and/or techniques discussed. These Guidance Notes do not and cannot replace the analysis and/or advice of a qualified professional. It is the responsibility of the reader to perform their own assessment and obtain professional advice. Information contained herein is considered to be pertinent at the time of publication, but may be invalidated as a result of subsequent legislations, regulations, standards, methods, and/or more updated information and the reader assumes full responsibility for compliance. This publication may not be copied or redistributed in part or in whole without prior written consent from ABS.

GUIDANCE NOTES ON

DATA INTEGRITY FOR MARINE AND OFFSHORE OPERATIONS

CONTENTS

SECTION 1
General

1 Purpose and Scope

The maritime industry is beginning to use data as an asset. Historically, marine and offshore owners' and operators' major concerns were safety, asset integrity, and environmental protection. Increased presence and use of cyber-enabled data systems introduces data as both an enabler to address safety risks and as another area for concern. Vessels and their sensors generate large amounts of data from multiple sources.

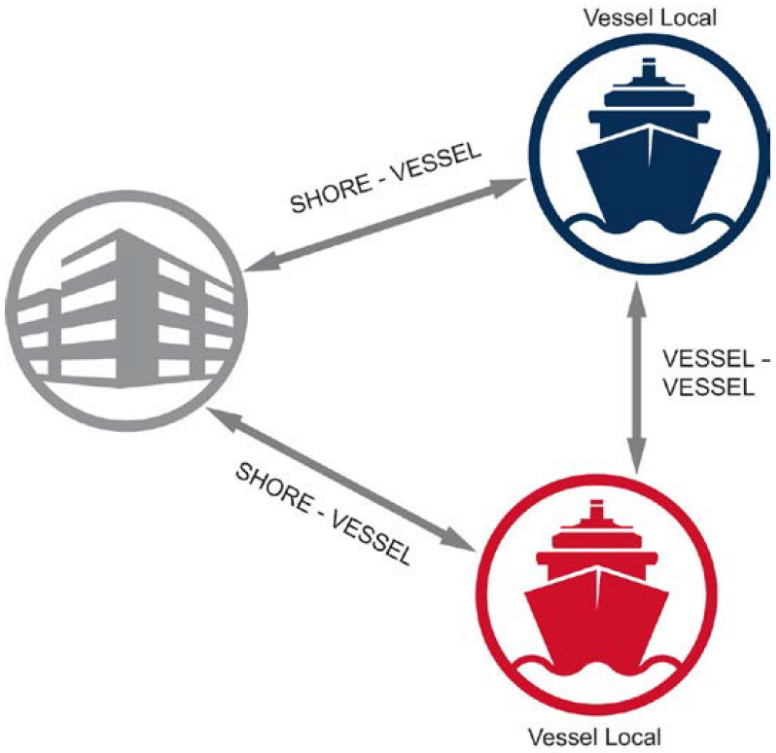
Note:

The general term "vessel" used throughout these Guidance Notes denotes a ship, a barge, an offshore unit or facility, or any other floating or fixed structure.

Section 1, Figure 1 illustrates nominal data sources and flows in marine and offshore operations. Three typical data sources and paths include:

- *Data Generated and Communicated Locally*. With the development of modern electronics and control technologies, extensive data can be generated and captured. The data covers a wide range of on-onboard systems and instruments, both from permanently-installed equipment and from cargo or portable equipment.
- *Data Communicated between Vessels*. The data communicated between vessels could be vessel status such as position, speed, direction, etc., but it may also include performance data, cargo carriage data, rig operational status data, or other raw or composite data sources.
- *Data Communicated between Shore and Vessel*. Operational, performance, and commercial data may come from shore-based systems as well as onboard systems. Data transferred from vessel to shore for data processing such as fleet management and benchmarking and maintenance management.

FIGURE 1
Data Source and Flow



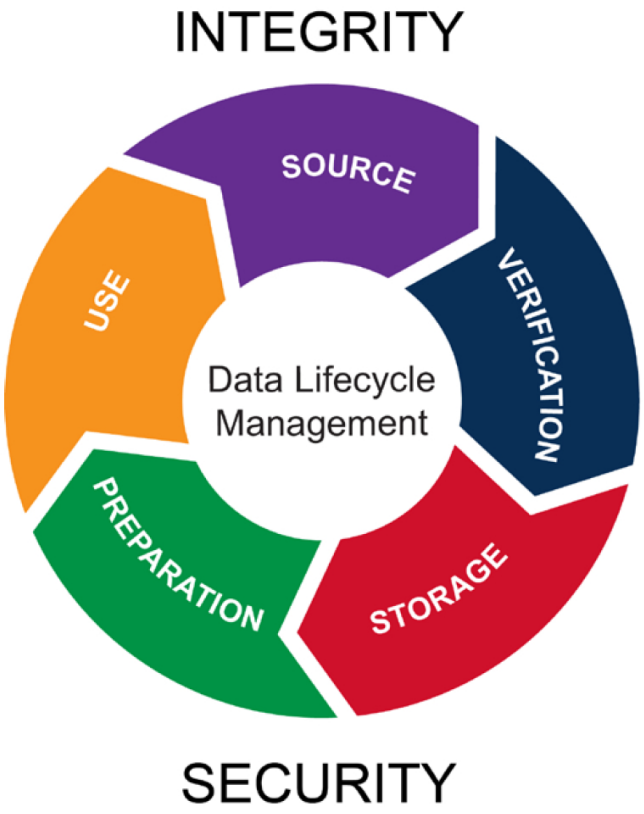
Data may be used for many functions beyond the traditional domains – for health and performance monitoring, operation, accomplishment prediction, business decision support, and others. New technical sources of data imply greater concern for data integrity and data security. Trustworthiness of data is vital to decision processes and to decision support.

These Guidance Notes are intended to clarify the concept and principles of Data Integrity for Marine and Offshore operations. The document will address data integrity as it relates to asset safety including human safety, safety of the vessel and/or threat to the environment. The objective of these Guidance Notes is to help the review engineers, surveyors, suppliers, shipyards, owners and operators to understand the application of Data Integrity in marine and offshore operations. The practices described in these Guidance Notes will improve the integrity of data.

3 Data Lifecycle Management

Data Lifecycle Management is illustrated in Section 1, Figure 2. The focus of this document is the security and integrity of the data management lifecycle. These Guidance Notes do not encompass all phases of the data lifecycle, but rather are directed at data source, data use, and data verification from the perspective of data security and data integrity.

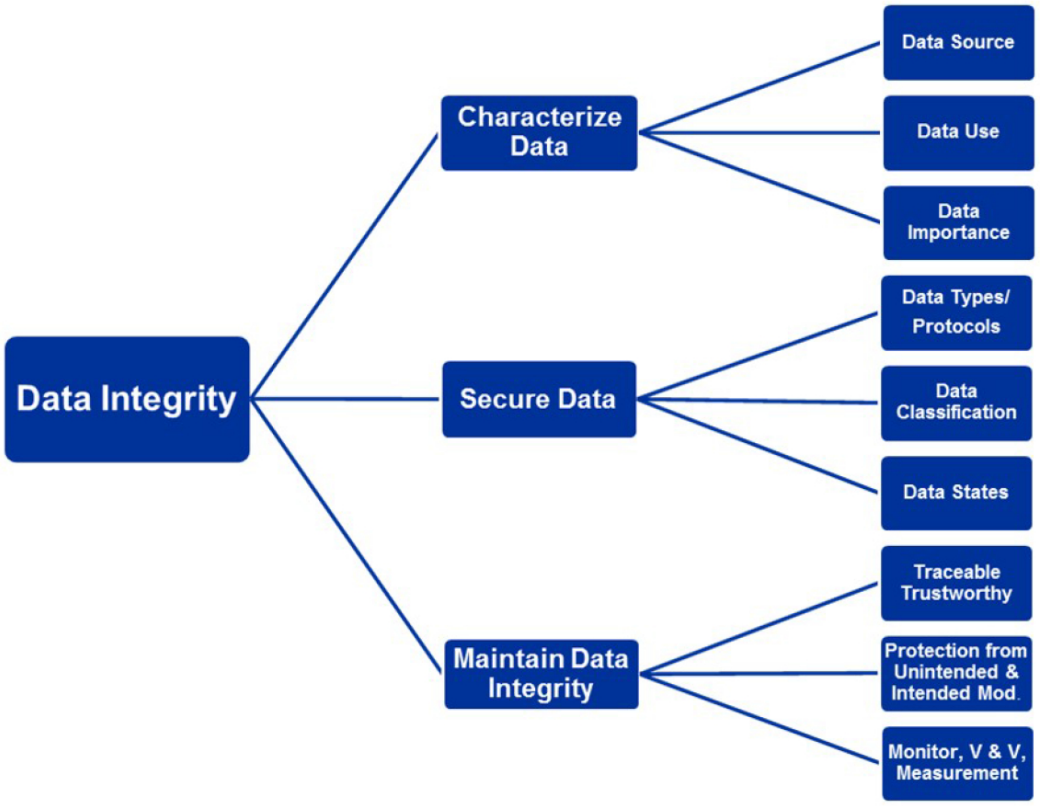
FIGURE 2
Data Lifecycle Management



5 Outline

The outline of these Guidance Notes is illustrated in Section 1, Figure 3. The outline indicates the development plan for these Guidance Notes. It shows how to characterize data, how to secure data, and how to maintain data integrity.

FIGURE 3
Guidance Notes Outline



Note:

V & V: Verification and Validation

7 Definitions

The following definitions are applied to the terms used in these Guidance Notes.

Access Control: Means to ensure that access to assets is authorized and restricted based on business and security requirements. [ISO/IEC 27000: Information technology – Security techniques – Information security management systems – Overview and vocabulary]

Bus Data: Data transmitted among multi-point networks.

CANbus: The data link layer of CAN; open transmitting of data between equipment, system controllers and data analyzers.

Chain of Custody (CoC): Chronological documentation, showing the seizure, custody, control, transfer, analysis, and disposition of physical or electronic evidence.

Data At-Rest (DAR): Data that resides in storage (a device or backup medium in any form) but excludes any data frequently transferred in the network or residing in temporary memory.

Data In-Motion (DIM): Data in transit, traveling across a network or contained in a computer's RAM ready to be read, updated, or processed.

Data Integrity: Accuracy, consistency (validity), and completeness of data over its lifecycle.

Data In-Use (DIU): Data that is being processed by one or more applications.

Data Schema: Skeleton structure that represents the logical view of the entire database.

Data Security: Protecting data from the unwanted actions of unauthorized users.

Database: An organized collection of data.

Integrity Level: A number assigned by an Owner and/or Driller or Crew Organization (DCO) to a computer-based function based upon the severity of the consequence of a failure of the function. Where 0 has little consequence to 3 where the consequence of a function failure is of significant concern. For control systems refer to ABS *Guide for Integrated Software Quality Management (ISQM Guide)*.

MODBUS: A common serial communications protocol for connecting industrial electronic devices such as sensors and programmable logic controllers (PLCs).

Penetration Test: Testing that places an attack on a computer system looking for security weaknesses and potentially gaining access to the computer's features and data.

PROFIBUS: PROFIBUS (Process Field Bus) is a standard for fieldbus communication in automation technology.

Sensor: An electronic device that produces electrical, optical, or digital data derived from a physical condition or event. [*IEEE 1451: IEEE Standard for a Smart Transducer Interface for Sensors and Actuators*]

Serial Data: Data transmitted in serial communication.

Software Management of Change (SMOC): The process of how to manage software changes or software evolution.

Streaming Data: Sequence of message-oriented data in-sequence used to transmit or receive information in a real-time application among the networks.

System of Systems: Vessels with multiple existing standalone and networked systems.

9 Abbreviations and Acronyms

The following abbreviations and acronyms are applied to the terms used in these Guidance Notes.

AAA: Authentication, Authorization, and Accounting

ABS: American Bureau of Shipping

ACLs: Access Control Lists

ARQ: Automatic Repeat Request

BSEE: Bureau of Safety and Environmental Enforcement

CRCs: Cyclic Redundancy Checks

CPU: Central Processing Unit

DAR: Data At-Rest

DCO: Driller or Crew Organization

DCS: Distributed Control System

DIM: Data In-Motion (equivalent to Data In-Transit)

DIU: Data In-Use

DP and DPS: Dynamic Positioning (Systems)

ECC: Error-Correcting Code

ERP: Enterprise Resource Planning

FDD: Functional Description Document

FEC: Forward Error Correction

HMI: Human Machine Interface

ICS: Industrial Control System

IIoT: Industrial Internet of Things

IoT: Internet of Things

IP: Internet Protocol

NAC: Network Access Control

OEM: Original Equipment Manufacturer

PC: Personal Computer

PLC: Programmable Logic Controller

PMS: Power Management System

RAID: Redundant Array of Independent Disks

RTD: Resistance Temperature Detectors

SAN: Storage Area Network

SCADA: Supervisory Control and Data Acquisition

SIS: Safety Instrumented Systems

SMOC: Software Management of Change

SoS: System of Systems

USB: Universal Serial Bus

USCG: United States Coast Guard

11 References

11.1 ABS

ABS *Guidance Notes on Application of Cybersecurity Principles to Marine and Offshore Operations – ABS CyberSafety™ Volume 1*

ABS *Guide for Cybersecurity Implementation for the Marine and Offshore Operations – ABS CyberSafety™ Volume 2*

ABS *Guide for Software Systems Verification – ABS CyberSafety™ Volume 4*

ABS *Guidance Notes on Software Provider Conformity Program – ABS CyberSafety™ Volume 5*

ABS *Guidance Notes on Equipment Condition Monitoring Techniques*

ABS *Guidance Notes on Failure Mode and Effects Analysis (FMEA) for Classification*

ABS *Guidance Notes on Risk Assessment Applications for the Marine and Offshore Industries*

ABS *Guide for Hull Condition Monitoring Systems*

ABS *Guide for Integrated Software Quality Management*

ABS *Guide for Surveys Using Risk-Based Inspection for the Offshore Industry*

ABS *Rules for Building and Classing Marine Vessels*

11.3 IEEE

IEEE 1451: *IEEE Standard for a Smart Transducer Interface for Sensors and Actuators*

IEEE 1012 – 2004, *IEEE Standard for Software Verification and Validation*

11.5 ISO

ISO/IEC 27000: *Information technology - Security techniques - Information security management systems - Overview and vocabulary*

ISO/IEC 27001: *Information technology – Security techniques - Information security management systems - Requirements*

ISO/IEC 27002: *Information technology – Security techniques - Code of practice for information security controls*

11.7 Other

National Institute for Science and Technology (NIST) *Guide to Storage Encryption Technologies for End User Devices*

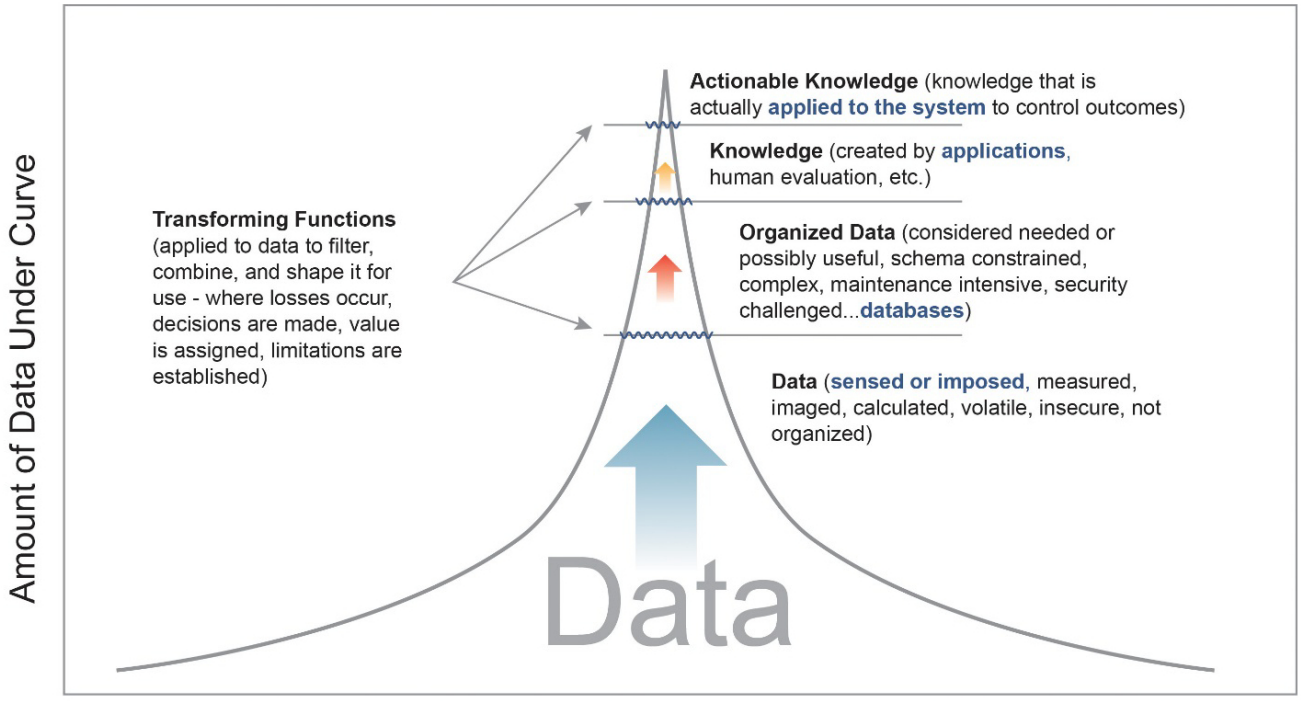
SECTION 2
Data Sources

1 General

For assets in the marine and offshore industries, data processing starts with the collection of data at the sensor level. Examples of data include location, direction and speed, hull structural stresses encountered, vessel motions, engine performance, equipment status, and environmental conditions. At this stage, both structured and unstructured data is collected. Structured data is formatted for use in specific systems, such as in transaction databases. Unstructured data may be data streams, messages, documents or other data types or artifacts that may be stored by means other than in a transaction database.

Massive quantities of "random" data of sensed behavior are electronically transformed into organized data, then into knowledge, and finally into purposeful actionable knowledge. Typically, an enterprise will organize its data across many different systems and applications. Graphically, the data source model's fundamentals are shown in Section 2, Figure 1 below.

FIGURE 1
Data Source Model



3 Raw Data Input

3.1 Data from Sensors (Raw/Unconditioned)

Sensors are the eyes and ears of ship automation and condition and performance monitoring, and are located in multiple locations and systems onboard the vessel. As the first stage of data sources, a sensor is an electronic device that produces electrical, optical, or digital data derived from a physical condition or event. Data produced from sensors is then electronically transformed into more organized output that can later be correlated as system information.

Enormous amounts of measurement instrumentation have been incorporated in modern machinery control systems, including temperature sensors, pressure sensors, flow sensors, vibration sensors, current sensors, and many more. They come in the form of mechanical gauges, electrical meters, transducers, thermocouples, resistance temperature detectors (RTD), etc. Section 2, Figure 2 below illustrates representative samples of common sensors used onboard vessels.

FIGURE 2
Various Sensors

Pressure Sensor

Temperature Sensors

Electric Meters (Current, Voltage, KW, etc.)

The information gained from the sensors is considered raw data without any conditioned process. At this stage, the raw data is neither conditioned nor organized, which means the data cannot be directly used.

5 Organized Data Sources (Information)

5.1 Data from Databases

A database is a collection of data and information that is organized so that it can easily be accessed, managed, and updated. It can be a collection of schemas, tables, queries, reports, views and other objects. The data is typically organized to model the aspects of reality in a way that supports processes requiring information. Databases are used to support internal operations of organizations, and to hold administrative information and specialized data.

5.3 Data Traces from Identified Equipment or Systems

Sensor devices provide performance or status information to operators, and they are essential for controlling equipment operation, providing alarms, or triggering equipment safety features, such as automatic shutdowns, alerting a degradation on condition or performance for action by the maintenance team. The engineering data onboard ship is used for machinery control decisions that may be performed by 'intelligent' devices or people.

The organized data is seeing a possible “tiered approach” to analytics. It can be from parameter level, to equipment level, to system level, then to the holistic level overarching at vessel. An example is found in shipboard engines. Leading engine manufacturers have transformed their operations through Engine

Condition Monitoring. From approximately one hundred sensors on each engine, they receive real-time performance data at central monitoring centers. Detailed information may be found in Norris, G. (2015, October 28), *Designing High-Tech Engines for Easier Maintenance, Aviation Week & Space Technology*.

In an electrical propulsion system, data generated from sensors is required for monitoring and control use. For example, bearing lube oil inlet pressure, voltage, frequency, current, stationary windings temperature, field voltage and current are the data collected and traced from the sensors for the propulsion generator system. Each of the data streams can be identified from certain equipment or system.

A Dynamic Positioning (DP) vessel has its own DP sensor network. It is a system comprising devices that measure vessel heading (such as gyrocompasses or inertial navigation systems), vessel motions (such as motion reference units), wind speed and direction, propulsion machinery system, and thruster system.

5.5 Data from Industrial Internet-of-Things (IIoT)

To date, the internet has primarily served as communication between humans, particularly in its application in the World Wide Web, email, and social media etc. The Industrial Internet-of-Things (IIoT), simply refers to Internet-enabled communication between non-human entities, such as devices and equipment, with data storage, applications and computers. The idea is that items as diverse as temperature sensors and washing machines can easily connect to Internet Protocol (IP) networks and communicate data about their own condition or what they are measuring, and they may also respond to external requests for data and actuation of certain commands.

The IIoT focuses strongly on intelligent cyber-physical systems. These systems comprise machines connected to computers that interpret, analyze and make decisions almost instantly based on sensor data from many widely distributed sources. IIoT-enabled devices activate when certain conditions arise and send alerts and associated data about emerging conditions which may require a response. For example, one situation might be that a thruster in a Dynamic Positioning System (DPS) may notice significantly different excursions on a semi-submersible drilling platform compared to other thrusters. This anomaly would initially be detected by communication and self-diagnostic analysis between the thrusters. When deviations beyond certain acceptable tolerances are noticed, an alert can be issued to a central control center for human oversight and intervention. Associated diagnostic data can also be provided for further analysis and the root cause of the problem confirmed and/or resolved. In a DPS, like many others, a safe limit on how much is controlled by computer must be carefully considered.

The advent of IIoT has spurred the creation of more and more intelligent machines that interact with other machines, with their environments, with data centers and with humans. As illustrated in Section 2, Figure 3, the three main components of an IIoT system are:

- *Things:* Device, sensors & actuators
- *Connections:* Local network, Internet
- *Data:* Information

FIGURE 3
Industrial Internet-of-Things (IIoT)

It is important that connections, communications and access to IP-enabled sensors and systems that are considered components of the Industrial Internet of Things (IIoT) are specifically addressed as part of the Operational Technology (OT) security measures onboard any ship, asset or facility. Remote accessibility to IIoT devices must be controlled carefully, as these devices are expected to be standalone, sealed, never-updated network participants, meaning that they can become conduits into primary networks if left exposed to unauthorized communications. These devices and similar systems are addressed in the specification of Protect Operational Technology in Section 5, Capability (17) of the ABS *Guide for Cybersecurity Implementation for the Marine and Offshore Operations – ABS CyberSafety™ Volume 2 (Cybersecurity Guide)*.

7 Conditioned Data Sources (Knowledge)

Data conditioning is the processing of data in a way that prepares it for the next stage of processing. Many applications involve environmental or structural measurement, such as temperature, pressure, level and vibration, from sensors. These conditioned data sources contain knowledge and meaningful information and feed for either machine interpretation or human interpretation.

By use of data management and optimization techniques, data conditioning will lead to intelligent routing, optimization and protection of data for data transmission or storage in a computer system.

7.1 Data Conditioned for Machine Interpretation/Use

Normally, the data conditioning process intended for machine use includes amplification, filtering, attenuation and isolation. For example, thermocouple signals have very small voltage levels that must be amplified before they can be digitized. Other sensors, such as resistance temperature detectors (RTDs), thermistors, strain gages, and accelerometers, require excitation to operate.

The machine uses the conditioned data to implement certain functions. For example, the air conditioning system uses the thermo-sensors conditioned data (such as digitized signal) to automatically control the room temperature.

7.3 Data Conditioned for Human Interpretation/Use

Algorithms, parameters, limits, stochastics, and statistics can be applied as applicable to aid operators in assessing machinery conditions; such as satisfactory operation, impending degradation, and prognostics or forecasting. Data conditioned for human use may include trend analysis visualizations, short-term performance parameter alerting, and similar reporting that can accelerate human decision making by reducing the steps required to interpret the aggregated data.

9 Actionable Data Sources (Applied Knowledge)

The last stage on the peak of the data sources model is the actionable data sources which includes the applied knowledge. The applied knowledge contained in the data is actually applied to the system to control the outcomes and for health and performance monitoring where it may be provided by a 3rd party vendor. At this stage, the data quantity has been significantly decreased compared to the previous process.

9.1 Manual Systems Control

Manual system control requires that a human operator be involved in all controlling activities in order to perform system functions. In these systems, the operator alone senses control data, makes control decisions, and implements the control actions without support by mechanical or computerized equipment. On contemporary offshore assets, very few control systems are purely manual. The most common application of actionable data is in automated control, discussed in 2/9.5 below.

Onboard the vessel, the operator can view the automated control systems from diesel engine control to power management system, from alarm and monitoring system to DP control system. Programmable Logic Controller (PLC), Distributed Control System (DCS) and Supervisory Control and Data Acquisition (SCADA) systems are commonly used in automatic control systems. The actionable data sources play key roles in these automated systems. For example, a DP control system aggregates multiple actionable data sources to implement the automatic control required in dynamic positioning operations.

9.3 Mechanized System Control

Mechanized system control commonly requires the support of computerized data. A large number of mechanized control systems are in place on contemporary offshore assets.

Mechanized control systems are simply systems that incorporate mechanical and digital machines to augment human strength, intelligence, and judgment in order to control equipment. Systems that require mechanized support to enable control commonly provide a combination of mechanical, electric, hydraulic, pneumatic, or computerized controllers. No matter the type of control augmentation, the human operator is an essential “component” of the control system. In contrast, manual systems and their control depend solely on the strength, knowledge, and judgment of an operator without the benefit of the augmentation described above. Fully automated systems can operate without the presence of a human operator.

Mechanized control system architectures that include computerized controllers, commonly rely on digital data in order to function. The integrity of digital data supporting computer-augmented mechanized control systems for critical functions is to be protected.

9.5 Automated System Control

Automated system control is increasingly present on contemporary offshore assets. A significant number of sophisticated drilling control systems have been placed in drilling platforms.

The automated control system uses control theory for regulation of processes without direct human intervention. Such systems require the least human operator involvement. Without human operator presence, automated systems and their control depend on sensors, transducers, transmitters, controllers and final control elements. In the simplest type of an automatic closed loop control, a controller compares a measured value of a process with a desired set value, and processes the resulting error signal to change the input to the process, in such a way that the process stays at its set point despite disturbances.

Automated control systems rely heavily on timely and accurate digital data in order to function. The integrity of digital data supporting software based on automated control for critical functions must be protected if the system is to operate reliably and correctly.

Some examples of automated control systems onboard vessels include diesel engine control, power management systems, alarm and monitoring systems, and DP control systems.

9.7 Analysis

Conditioned data may support data analysis applications in certain situations. Such systems come in many forms, means of implementation, comprehensiveness, and off vessel communication capability. Such applications may be as simple as a single machine Original Equipment Manufacturer (OEM)-provided diagnostics package on an engine (a common example), or they may be implemented at the system or asset level by a separate data analytics vendor who is tapping into a myriad of data streams coming from automation systems (operational data) or condition or environmental monitoring sensors or devices. Such data is then collected and analyzed at the asset level for diagnosis/prognosis of systems or vessel health and performance states.

One example in the marine and offshore world is the vessel hull and systems condition data that have been continuously used as key performance analysis indicators. Hull monitoring systems continuously collect the data used for motion, stress and voyage applications. Both unprocessed, collected data and summary trend reporting are used for immediate interpretation of processed data and for subsequent longer-term evaluation by vessels' operating personnel. Conditioned, measured data can also be used for safety assessment and analysis. Such data can lead to better drydock and survey planning. It can also provide better understanding of damage accumulation and thus the prevalence of fatigue in certain prone locations, which also feeds better survey and drydock planning.

In a propulsion system, for example, historical maintenance and failure records would be within the scope of data to be included. Engine utilization and power settings, particularly just prior to the occurrence of functional or performance issues or failures, would be useful information to understand how the failure relate correlates to the Maximum Continuous Power Rating and operating times (i.e., operating profile). Pump performance data (e.g., flow rates, temperatures, pressures), particularly issues just prior to the occurrence of failures, would be useful in developing an improved operating profile.

On legacy assets, often those systems involved the addition of sensors, communication and wiring, and the typing into a new data collection historian or similar device.

SECTION 3
Data Uses

This Section covers how collected data can be used and applied to improve asset performance.

1 Monitoring for Situational Awareness

Data monitoring can provide situational awareness. For example, all monitored real-time data is typically transferred and displayed in the ship's bridge so that the Master is aware of health measures on the vessels. These data could contain weather, engine, power and vessel health and performance conditions.

3 Monitoring for Intervention

Data monitoring can be used for intervention or corrective actions. Monitored data and its reporting may provide decision support and feedback for current operations. For example, a DP control system monitor takes action to increase the speed or change the system thrust direction to optimize performance. A Power Management System (PMS) control can reduce or increase power based on vessel demand.

In some cases, human intervention is required. The workflow for intervention is to be sensibly considered and implemented within a system. For example, condition monitoring of equipment may indicate that a failure is imminent, therefore requiring notification to maintenance personnel of the potential failure modes and maintenance needs. Maintenance processes should be updated to incorporate this insight and realize the benefits, which include more accurate troubleshooting and more efficient maintenance.

5 Monitoring by Regulatory Bodies

Data monitoring is often used by marine regulatory bodies, such as United States Coast Guard (USCG), or the Bureau of Safety and Environmental Enforcement (BSEE). Critical data monitoring helps regulatory bodies develop and enforce proper policies and requirements.

For example, BSEE has issued the final well control regulations. The regulations are BSEE, 30 CFR Part 250, *Oil and Gas and Sulfur Operations in the Outer Continental Shelf-Blowout Preventer Systems and Well Control*. The final rule addresses the full range of systems and equipment related to well control operations, with a focus on blowout preventer (BOP) requirements, well design, well control casing, cementing, real-time monitoring, and subsea containment. Real-time data from well control equipment is required to be transmitted to BSEE for monitoring.

7 Input to Control Systems

As described above, sensor data is used as inputs to control machines and equipment. An increasing number of control system functions have been implemented onboard vessels and assets with some examples listed below:

- DP control
- Propulsion remote control

- Engine control
- Power management system
- Cargo control

As the industry moves towards autonomy and autonomous systems, data use becomes a vital part of operations.

9 Input to Analysis and Patterning

For the marine and offshore industry, data analysis begins with the collection of data at the vessel level, such as location, direction and speed, hull structure stresses encountered, vessel motions, and engine, equipment, and environmental conditions. Then the data is transmitted securely by satellite to a response center, which can offer basic services such as vessel tracking and structural integrity monitoring. This process is still at the early adoption stages in the marine and offshore industries. It indicates that data analysis will play a significant role the near future for improving the performance of individual systems or vessels as a whole.

Significantly more value can be added by applying data analysis to the entire fleet or business unit, where data monitoring objectives go beyond optimizing the performance of individual assets. It is at this level that data from disparate, multidimensional sources may be analyzed to secure new macro-level insights.

Weather patterns and wider operating conditions, such as sea currents and temperatures, etc. could have an effect on vessels operations, and data about those conditions may be gathered to provide insight on vessels operations. Historical and transactional data trends can help identify risks and opportunities associated with current data (e.g., gradually increasing vibrations of a certain frequency in a particular type of pump may indicate imminent failure).

11 Input for Maintenance

As the demand for efficiency increases, maintenance plays a big role in improving asset operations. Data can be used for maintenance to reduce unexpected or scheduled down time. Condition monitoring can include both hull condition monitoring and equipment condition monitoring. Condition-monitoring tasks are scheduled or continuous activities used to monitor machine condition and detect a potential failure in advance so that action can be taken to prevent that failure. Condition Based Maintenance (CBM) is a maintenance plan, which is based on the use of Condition Monitoring to trigger the relevant maintenance task and corresponding part replacement or other corrective action. This process involves establishing a baseline and operating parameters, then frequently monitoring the machine and comparing any changes in operating conditions to the baseline. Maintenance tasks and overhauls are then carried out before the machinery fails.

Further details regarding condition-based monitoring can be found in the following ABS publications:

- ABS *Guide for Hull Condition Monitoring Systems*
- ABS *Guidance Notes on Equipment Condition Monitoring Techniques*

SECTION 4

Data Importance

Most data collected and applied aboard marine and offshore vessels are important to vessel safety, security, performance, and handling. Even so, data that support systems carrying out Integrity Level 2 (IL2) (details provided in *ISQM Guide* and also for quick reference within Section 5, Table 2 below) and Integrity Level 3 (IL3) functions aboard a vessel are considered critical for protection of human life, the vessel, and possibly the maritime environment. By contrast, systems and data supporting IL0 and IL1 functions are important to onboard business and personnel comfort and convenience, but have minimal impact on safety and security. This distinction leads to a need to determine the relative importance of data when making decisions about applying limited resources to data integrity management. Data that supports essential onboard functions deserve special integrity protections. This section provides a simple risk-based model for decision making when applying resources to data integrity protection.

1 Data Integrity and Vessel Operations

While a number of functions aboard marine vessels operate independent of human intervention, others require exceptional event monitoring and decision making by shipboard personnel and/or safety instrumented systems (SIS). Sensors provide the data required for event monitoring and possible operator intervention and resolution. The data-driven controls also provide the capability for operators to take corrective action when failure or reduced-performance events occur. The data needed to guide actionable knowledge for these corrective actions – especially when applied to IL2 or IL3 functions – merit consideration as critical data within special integrity protection.

Additionally, most systems with control consoles and human-machine interface (HMI) situational awareness dashboards support essential onboard functions. Simple examples are the navigation systems on vessels and the consoles provided with driller's chairs on drillships. Failure of these functions due to loss of data integrity also represents a risk to vessel and crew safety. Therefore, these systems also merit consideration for special data integrity protection.

Although the above concepts focus on exceptional event handling, all data that supports normal vessel operation is important. The vessel's data management system architect is responsible for defining and documenting the risk hierarchy of managed data, and is to base data integrity protection decisions on that hierarchy. The key concept is that software providers, asset owners, and asset operators are to base data integrity protection decisions on a rigorous risk management process that is grounded in an assessment of failure risks and consequences. The foundation of this risk management process is the realistic assignment of integrity levels as provided in the *ISQM Guide*. More risk assessment guidelines can be found in the following ABS publications,

- ABS *Guide for Surveys Using Risk-Based Inspection for the Offshore Industry*
- ABS *Guide for Surveys Based on Machinery Reliability and Maintenance Techniques*
- ABS *Guidance Notes on Risk Assessment Applications for the Marine and Offshore Industries*
- ABS *Guidance Notes on Failure Mode and Effects Analysis (FMEA) for Classification*

3 Data Integrity and Business Strategy Development

In the above section, data integrity is discussed in terms of operational performance and decisions – decisions that support improved vessel and crew safety. Additionally, the same data is used in large part to analyze patterns of vessel lifecycle characteristics that can be applied to strategic business decisions.

As discussed above, vessel data is used to improve design, vessel fleet maintenance and supplied services decisions, staffing decisions, and supplier selection and supply chain management decisions. While these decisions may not be accompanied by the immediacy of incident response actions, they can have equally important long-term impact on the health and continuity of a marine enterprise. Data pulled from onboard systems are used to both guide enterprise product and service development decisions, and drive commitments to develop or implement emerging operational processes and technologies. Therefore, the data used to build the knowledge needed to support these strategic decisions also merits consideration for special data integrity protections.

5 Data Integrity and Integrated System Support

Marine systems are by nature highly integrated. As these systems become more functionally capable, they also become more complex. These conditions combine to make the integrity of data transferred among linked systems especially important. In standalone or discrete systems, a data integrity loss or timing failure can cause a failure in that system. In highly integrated systems, similar issues can cause a cascading failure that impacts numerous subsystems, some of which are likely to be linked to critical or essential systems (e.g., IL2, IL3, and SIS subsystems). The data management system architect is responsible for understanding the data integrity interdependencies and the risks to the major functions associated with an integrity failure of data flows. Further, this understanding is to be made apparent in an asset's functional description documentation and the resulting data integrity protections.

7 Data Integrity and Compliance Reporting

Marine systems associated with condition or status monitoring for reporting to outside authorities should also be considered as high-priority systems. Regulatory reporting may rely upon those systems monitoring energy management, pollution emission control, oily discharge prevention, safety instrumented systems, and others. Individual systems require security and measurable data integrity to represent their results, monitoring or reports in ways that can be certified by the regulatory agency or compliance authority. These systems may not have operational impacts on the safety, security, performance or handling of a ship or offshore asset, but they may still be high integrity level systems due to their monitoring or reporting tasks. Data verification also plays critical role in supporting owner compliance to regulations.

SECTION 5
Data Security

Data security is defined as protecting data from destructive forces and from the unwanted actions of unauthorized users.

Data security risks can include unlawful control of device/machine, abusive insertion, update and deletion of data, or negligent or inadvertent data losses that impact system functions.

Just as an organization would not leave a vessel unguarded, they must also protect their data from unauthorized access. As hackers have gotten more sophisticated, so too have their traditional, data heavy targets. As a result, hackers have been turning to targets such as those in the Marine and Offshore industry. Even without hackers, uncontrolled data can easily be exposed. Confidential information is frequently shared too broadly through internal reports distributed outside of their intended audience, emails forwarded outside of an organization, or even in annual reports. Uncontrolled data is insecure data. For data security, data is to be controlled while at rest, in motion (transit), and in use (processing).

This Section outlines the basic principles of how to identify and classify the data, and how the data may be protected.

1 Data Types/Protocols

In these Guidance Notes, data types are defined according to the way data is transmitted during communication. Serial Data, Bus Data and Streaming Data are the three types described in this document. Section 5, Table 1 provides typical features of each data type.

TABLE 1
Data Types

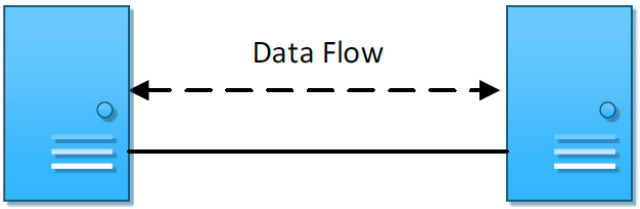
Data Types	*Features*
Serial Data	• Point-to-Point • Physical Layer
Bus Data	• Multi-point • Physical – Network – Physical Layer
Streaming Data	• Sequence of messages continuously transmitted • Transport Layer

1.1 Serial Data

Serial Data refers to the data been transmitted in serial communication between point-to-point interfaces. The data flow diagram between two nominal devices is shown in Section 5, Figure 1. Often, the receiving

end of one host is connected to sending end of the other and vice versa. Serial data is transmitted in the physical layer and relatively easy to secure because the transmission medium is specified by standard, and known by location.

FIGURE 1
Serial Data Illustration



Serial data transfer has the following advantages and benefits.

- It requires only a limited number of wires to exchange data between devices, thereby simplifying transmission path engineering. Signal strength can be augmented easily with repeaters and amplifiers.
- Serial communications require low interface pin counts. Serial communications can be performed with just one I/O pin, compared to eight or more for parallel communications. Many common embedded system peripherals, such as analog-to-digital and digital-to-analog converters, LCDs, and temperature sensors, support serial interfaces.
- Serial buses can also provide inter-processor communication networks. This allows large tasks that would normally require larger processors to be tackled with several inexpensive smaller processors. Serial interfaces allow processors to communicate without the need for shared memory and semaphores, and the problems they can create.

Perhaps the most successful serial data standard for native computer and telecommunications applications is the RS-232. Similarly, the RS-485 and RS-422 are among the most successful standards for industrial applications.

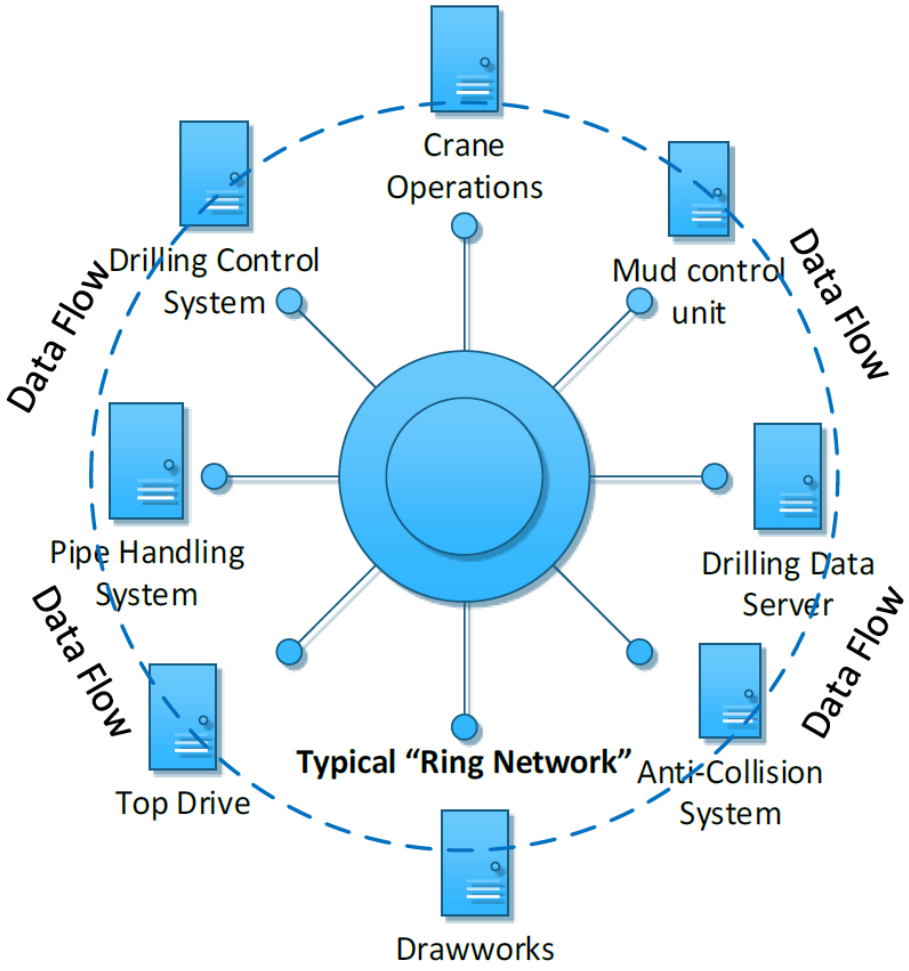
Typical marine implementations include RS232/485 serial data interfaces or fiber optic cables and hubs. Each of the I/O modules is physically wired in a ring bus with unique IDs (detailed information available in the *ISQM Guide*) to allow data transfer to/from each module. The wire protocol and messages utilized are determined by the Original Equipment Manufacturer (OEM) of the control systems (automation, control, and monitoring). There are multiple process protocols that exist which describe the messages transmitted on the serial bus including MODBUS, CANbus, and PROFIBUS. A vessel's serial bus thus provides the data pathway between physical sensors and the hardware/software interfaces to enable automated data collection.

Serial data is fairly easy to be secured since it transmits data in a physical layer (i.e. cables). There is danger of electromagnetic interference (EMI) against serial data, but this risk can be reduced with cable shielding, cable or interface armoring and grounding, or other physical solutions to reduce system noise.

1.3 Bus Data

Bus Data refers to the data transmitted among multi-point networks. Here multi-point networks can mean bus, star, ring or mesh topology. An illustrated diagram of a multi-point ring network is shown in Section 5, Figure 2.

FIGURE 2
Bus Data Illustration



Transmission Control Protocol (TCP)/Internet Protocol (IP) is the most common protocol set used on modern bus transmission systems. Security measures and controls for these common systems may be found in the *Cybersecurity Guide*.

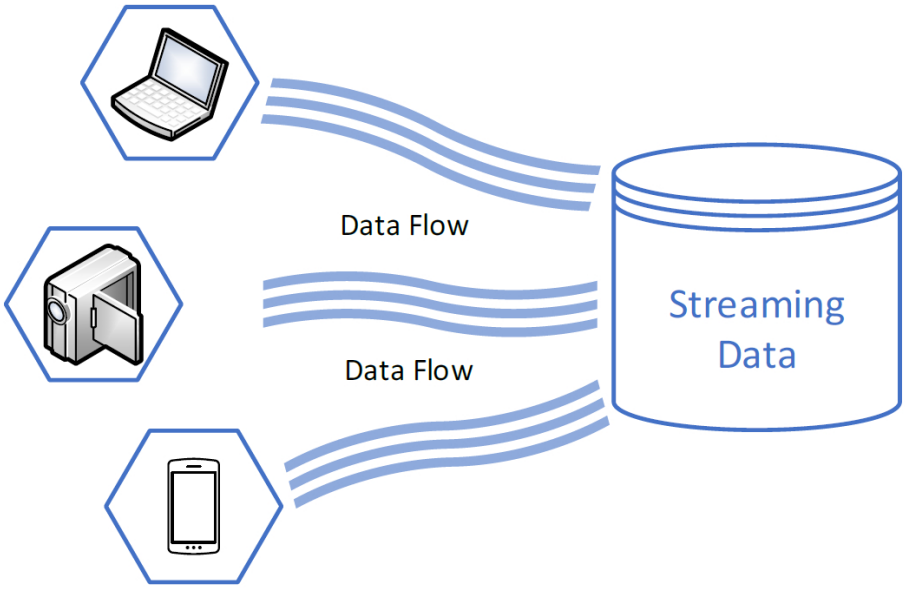
1.5 Streaming Data

Streaming data is another type of data found in the marine and offshore industries. Streaming data refers to a sequence of message-oriented data in-sequence transport used to transmit or receive information in a real-time application among the networks. Three major types of streaming data are listed below:

- Automatically-generated machine data streamed from connected devices. Such devices can be sensors or Internet of Things (IoT)/Industrial Internet of Things (IIoT) devices. An example is streaming data generated from a security network camera.
- Human-generated data from social media, such as feeds originating from collaboration systems or social media like Facebook and Twitter.
- Automatically generated human data. For example, the streaming data generated due to actions while performing online web browsing.

Conceptual streaming data is illustrated in Section 5, Figure 3 below.

FIGURE 3
Streaming Data Illustration



In the marine and offshore industries, the most common streaming data is automatically generated machine data. The major driving force for such data is the appearance of IoT/IIoT scenarios, such as real-time remote management and monitoring. Some benefits of the streaming data include

- Gain meaningful, time-sensitive insights into operational systems;
- Perform real-time analytics for accelerated decision making; and
- Achieve mission-critical reliability and scale with continuous system adjustments based on real-time/ near-real-time data reporting.

The IoT/IIoT connects bidirectional-communicating devices with one another in real time. With computational systems becoming more ubiquitous as processors become cheaper and more capable, more and more real-time data will be available among networked devices.

Streaming data has been designed for improved security. An example is a 4-way handshake to protect against synchronized flooding attacks, and large "cookies" for association verification and authenticity. Multi-homing and redundant paths increase resilience and reliability.

3 Data Classification (Protection Level)

Data Classification is the process of organizing data into categories for its most effective and efficient use. Data Classification facilitates assignment of protection level based on its importance. As noted in Section 4, data has different criticality based on its supporting systems.

Once a data-classification scheme has been created, security requirements that specify appropriate handling practices for each category and storage standards that define the data's lifecycle requirements are to be defined.

In this document, data is classified in terms of criticality to the mission of the vessel. Four categories have been defined as follows and shown in Section 5, Table 2. The data classification has immediate relevance to Integrity Level (IL) (defined in the *ISQM Guide* and the *Cybersecurity Guide*). The relationship is also represented in Section 5, Table 2:

- *Mission Critical:* Data is critical to the vessel. For example, data in safety system/equipment for main propulsion, power management data for the electric power generating system.
- *Mission Essential:* Data that is essential to the vessel supporting IL2 functions systems. For example, the vessel management system.
- *Non-Mission Essential:* Data supporting IL1 functional systems is non-mission essential. Such data is generally used for business essential purposes, including such examples as Enterprise Resource Planning (ERP) data (orders, purchase orders, human resources data and employee payroll).
- *Non-Vessel:* Data supporting IL0 functions systems is not for vessel operation. For example, personal email and web data may be classified as IL0.

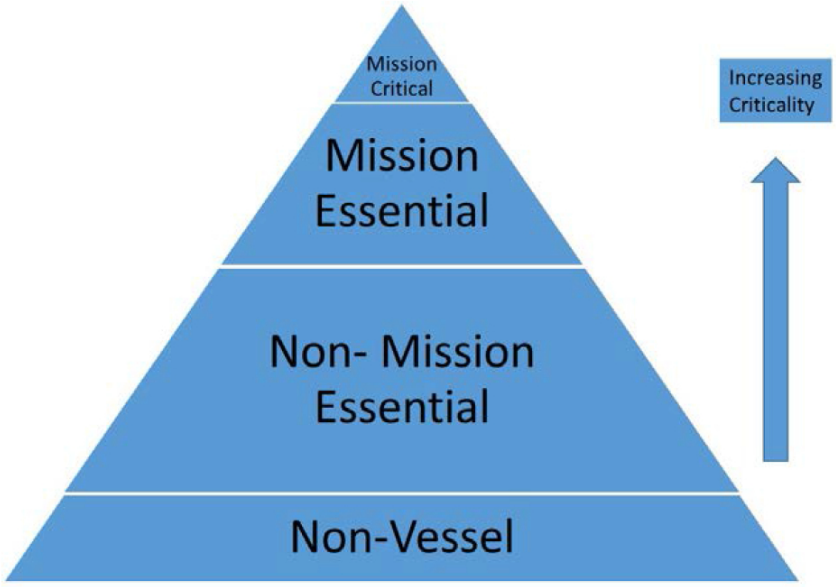
TABLE 2
Data Classification

Data Classification	*Integrity Level (IL)**	*Potential Consequences Functional*	*Examples, not inclusive*
Non-Vessel	0	Minor impact on operation. Might affect supporting process system but not main process system	Entertainment System, Administrative computer systems, office network, Data Collection system (non-Authority required)
Non-Mission Essential	1	Might lead to maintenance shutdown of non-critical system. Main process continues to operate.	Non-essential control of systems, BPCS, Non-essential communication systems, Vessel Management System. Enterprise Resource Planning (ERP)
Mission Essential	2	Shutdown of main system, excessive time for repair.	Drilling control system, BPCS, SIS systems (minimum rating), PMS, essential systems, DP control system, main engine control system, safety systems, cargo control system, navigation system, new or unproven essential technologies minimum rating.
Mission Critical	3	Significant repair time or loss of the marine or offshore asset.	Drilling Blowout Preventer control system, SIS or safety control systems, boiler firing control system, etc.

Note:
*Integrity Level for the control systems are defined in *ISQM Guide*. Please refer to *ISQM Guide* for the detailed procedures.

The four data classification relationships are shown in Section 5, Figure 4.

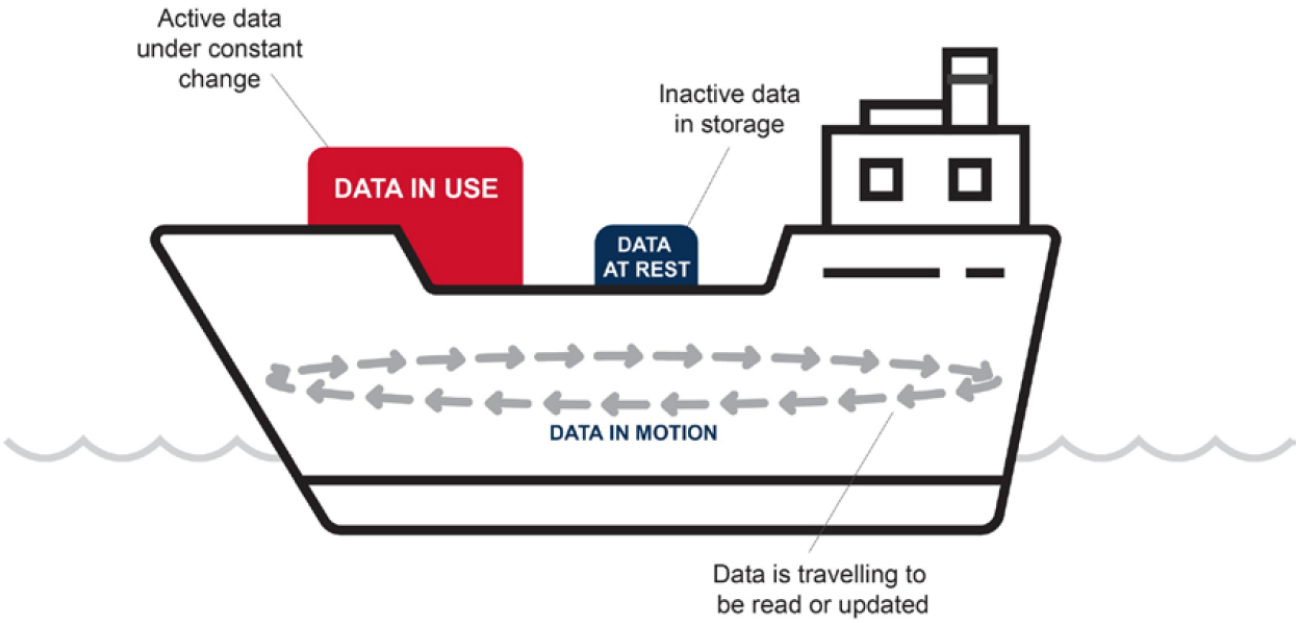
FIGURE 4
Data Classification



- Mission Critical data has the most criticality but the least quantity. Among all the data generated and collected onboard, this data only occupies a very low percentage. This data is to receive the highest protection to keep the data secure. The data protection method will be selected based on the data states (explained in the following section).
- Mission Essential data has the second highest level of criticality. These data are of a higher quantity than Mission Critical data.
- Non-Mission Essential data has the biggest quantity with fairly low criticality. These data are of the highest quantity. Most of the data generated onboard are under this category.
- Non-Vessel data has the lowest criticality and is trivial to the asset. The quantity of such data varies on different assets.

As noted in the definitions in Section 1, three basic states of data characterize data: Data At-Rest (DAR), Data In-Motion (DIM), and Data In-Use (DIU). Understanding the different data states can help to select the methods of security measures and encryptions that appropriate for protecting the data. The three states of data onboard a vessel are illustrated in Section 5, Figure 5.

FIGURE 5
Three States of Data



5 Data At-Rest (DAR)

Data At-Rest refers to the data that resides in storage (a device or backup medium in any form) but excludes any data frequently transferred in the network or residing in temporary memory. DAR is in an inactive and stable state that is not currently being transmitted across a network or actively being read or processed. It is not travelling within the system or network, and not being acted upon by any application or the CPU.

Here are some examples of DAR:

- Data in data shares or repositories
- Data on endpoints (i.e., PC or laptop devices) that is not accessed or changed frequently
- Archived data
- File stored on hard drives or USB thumb drives
- Files stored on backup tape and disks
- Files stored off-site or on a storage area network (SAN)

7 Data In-Motion (DIM)

Data In-Motion (DIM) refers to the data currently in transit, traveling across a network or sitting in a computer's RAM ready to be read, updated, or processed. It is data in the process of moving through, or crossing over networks from local to cloud storage or from a central server to a remote endpoint. In Marine and Offshore operations, DIM includes the data traveling in vessel's local communication paths, or between vessel and vessel, or between vessel and shore (on shore, communicated between vendors, flags, class etc.). The data moving could be through wire or wireless transmission.

9 Data In-Use (DIU)

Data In-Use (DIU) refers to data that being processed by one or more applications. Such data is not under storage status or during the transmission. DIU is in the process of being generated, updated, changed, or deleted. It also includes data being viewed by users accessing it through various endpoints.

9.1 Authorized Access Only

The endpoints are the most vulnerable points for DIU since they are the place where users can access and interact with the data. The data set can potentially have multiple users from multiple endpoints. Protecting DIU starts from access control to the data.

First, the data is to be protected by authorized access only, which includes user authentication, identity management, and profile permissions. This means only the individuals with the proper permission, qualification and knowledge are able to access and manipulate the data.

Second, most employers have their employees sign legal agreements that they will not share data with anyone that does not have permission to view it.

9.3 Penetration Protection

DIU is also to be protected from penetration by corruptive agents or incidents. Generally, the phrase "penetration" describes an unauthorized or malicious intrusion into a computer system; however, an unauthorized accidental intrusion is also a penetration that can have serious consequences. A successful intrusion that is detected and characterized may reveal weaknesses in the system's protective capabilities. Intrusion characterization is also intended to reveal the actual and potentially corruptive nature of the intrusion upon hardware, firmware, processing applications, networks, and data. For these reasons penetration protection is to be implemented in the interest of protecting data integrity during data creation, transport, processing, and storage.

Penetration testing is commonly used to attempt to penetrate a digital system environment in order to test the capabilities of protective measures (system architecture, hardware, software, and procedures). A penetration test is performed to determine sufficiency of protective measures. It is also performed to detect and characterize weaknesses in protective measures. Based upon the outcomes of the test, protective measures can be corrected and enhanced as necessary to prevent an attack or accidental data corruption incident.

Routine penetration testing is recommended for systems that collect, process, store, and transport/transfer protected data.

9.5 Handling Policies and Procedures

Detailed data handling policies and procedures are to be defined for secure management of protected data. Data handling policies and procedures are provided in the *Cybersecurity Guide*, and are to minimally include:

Documented data protection policies and procedures pertinent to each phase of the data management lifecycle include:

i) Procedures documenting authorization process for accessing or handling protected data, including read/write privileges, use locks, blocking devices, and strong login credentials;

ii) Procedures documenting process for updating or modifying protected data; and

iii) What type of operation can be run on the data

11 Security Measures and Controls

Security measures for DAR, DIM, and DIU are found in the *Cybersecurity Guide*.

SECTION 6
Data Integrity

1 General

As more sophisticated automation systems become common in marine and offshore assets, growing concerns regarding data integrity have been raised. The concerns cover various topics in data integrity, such as basic concepts of data integrity in marine and offshore assets, how the data integrity is best managed, and what are the measurement tools for data integrity. Traditional definitions and concepts may not adequately support specialized data integrity applications in marine and offshore assets. This section adopts a holistic view to define the data integrity concept, the systematic process of data integrity, and the verification of data integrity.

1.1 Definition

As defined in ISO/IEC 27000 *Information technology - Security techniques – Information security management systems – Overview and vocabulary*, Integrity means the "property of accuracy and completeness". Integrity can have a large number of meanings depending on the context. For example, it can refer to accuracy, functionality, uncorrupted data, and absence of interference, restriction of access, code structure, and calculation accuracy.

In these Guidance Notes, Data Integrity refers to the accuracy, consistency (validity), and completeness of data over its lifecycle. Compromised data is almost no use to the marine and offshore asset, and could cause a dangerous situation to human safety, safety of the vessel, and/or an environmental threat. For this reason, maintaining data integrity is a core focus of marine and offshore asset cybersecurity. Maintaining data integrity helps improve recoverability and searchability, traceability (to origin), and connectivity.

Data integrity may be compromised in various ways. Some representative failures that can affect data integrity include:

- Physical broken hardware devices (such as sensors or disk crash)
- Human error, including unintentional actions and malicious intent
- Transfer errors, including unintended alterations or change during the transfer process
- Cyber threats, including viruses/malware, hacking

Data integrity is a fundamental component of information security. It can be used to describe a state, a process or a function and is always used as a representation for data quality. The three descriptions are listed as follows:

- *As a state:* Data integrity defines a data set that is both valid and accurate and maintains fidelity.
- *As a process:* Data integrity verifies that data has remained unchanged in transit from sending to receiving.

- *As a function, related to security:* Data integrity maintains information accuracy, is auditable, and supports reliability.

1.3 Data Integrity and Data Security

Data security basic principles and criteria have been described in Section 5 of this document. Data integrity and data security are closely related terms, and they both play important complementary roles.

Data integrity is the desired result of data security, mostly achieved by the act of data protection (securing data). Efforts for data integrity focus on validity, accuracy, and completeness of data. Work in data security focuses on the act of protecting data. Whether unintended modification or malicious intent, data security is a critical part of maintaining data integrity.

For modern marine and offshore operations, data integrity plays an essential role in the accuracy and efficiency of a vessel's normal operation as well as business process.

As described in Section 4, the integrity level of data is to match the integrity level of the function it supports.

Data integrity can be implemented in a variety of ways. The data is to remain unchanged while it is being handled, transferred or replicated. These Guidance Notes will describe data integrity from the following three areas:

- Confirm data integrity from traceable and trustworthy data source
- Protect data integrity from unintended and malicious modification
- Monitor, verify, validate and measure data integrity

3 Maintain Data Integrity from Traceable/Trustworthy Data Source

3.1 Origin: Supplier and Sensor Accuracy

3.1.1 Pedigree Verification

In this context, the word "pedigree" describes the status of data origin. The definition provided by the Oxford Dictionary for pedigree is "The record of descent of an animal, showing it to be purebred". The same meaning applied to origin of data source is that data has the record to demonstrate it was born pure and accurate, plus it has the clear trace back to its root. The pedigree of the data source is to be verified to maintain the data traceability. The supplier is to deliver a system having an approved and accepted level of integrity which means the sources have been selected appropriately and certified.

3.1.2 Chain of Custody Verification

The term Chain of Custody (CoC) has been commonly used in legal contexts which refers to the chronological documentation or paper trail, showing the seizure, custody, control, transfer, analysis, and disposition of physical or electronic evidence. We are adopting this term in data integrity to explicitly demonstrate the properly documented CoC is an essential element of data integrity. Correctly performed custody procedures allow that data to be validated.

Data generation, management, transmission, storage and processing must all trace through logs to show a clear CoC. An example of the necessity for data tracking is ballast water treatment system data. Due to the new regulatory requirements, ballast water data is required to be collected periodically and reported to the regulatory agencies. Accuracy of reported data is crucial to maintain correct reporting.

Data-generating systems must present data with a traceable pedigree and CoC traceability to support both versioning and auditable record trail:

- Versioning is to be an automated procedure to check data in and out to maintain the latest version of the data.
- Record trail is to be generated by the system instead of by users. System generated date and time stamps and other related information such as location can be included in the record trail. It will provide the log for both internal and external use.

3.1.3 Sensor Verification & Validation

As described in Section 2, the most common data source is from onboard reporting sensors. Massive quantities of data may be generated and collected from hundreds, and perhaps thousands of sensors installed on the vessel.

There are two potential reasons for non-trustworthy or erroneous sensor data: Unintentional errors and intentional misbehavior.

- Unintentional errors may be caused by hardware malfunctions (broken or obstructed sensors), poor positioning of the node (unconnected or incorrectly attached node) or depleted batteries. Loss of sensor may also include loss of a monitoring or aggregator system, excessive electronic noise on the line, or transmission line breaks.
- Intentional misbehavior may be caused by crew actions or external attackers, exploiting security vulnerabilities for unexpected purposes. Standalone sensors and reporting nodes may be at risk, given IoT or IIoT devices will be network-addressable, generally isolated in unobtrusive areas, and infrequently updated. Most IoT or IIoT devices are expected to be sealed appliances that will serve a purpose and be replaced before they will be maintained or upgraded.

The best practices to resolve the above two issues include:

- Testing of the sensor software/hardware to meet hardening standards applicable to the industrial (marine or offshore) environment;
- End-to-end communications testing, from sensor to data capture device; and
- Proper security actions implemented to prevent intentional misbehavior. The detailed procedures can be found in the *Cybersecurity Guide*.

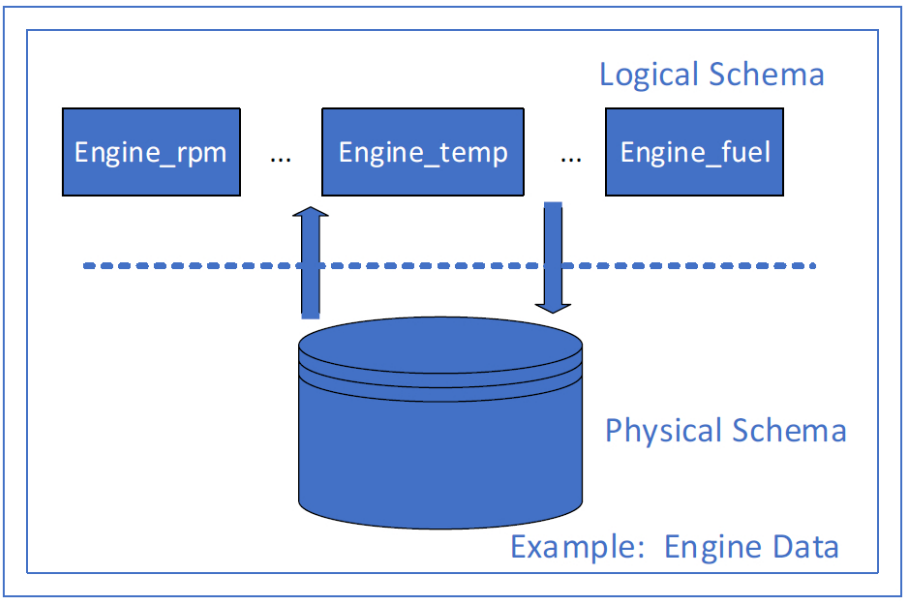
3.3 Organization: Data Schema Transformation Accuracy

Data schema is the skeleton structure that represents a logical view of the entire database. It defines how the data is organized and how the relations among data elements associate.

As data is generated and collected from its source, it will go through a data organization process. Data schema is the architectural description process that helps data to be organized for appropriate use.

Data schema is illustrated in Section 6, Figure 1.

FIGURE 1
Data Schema

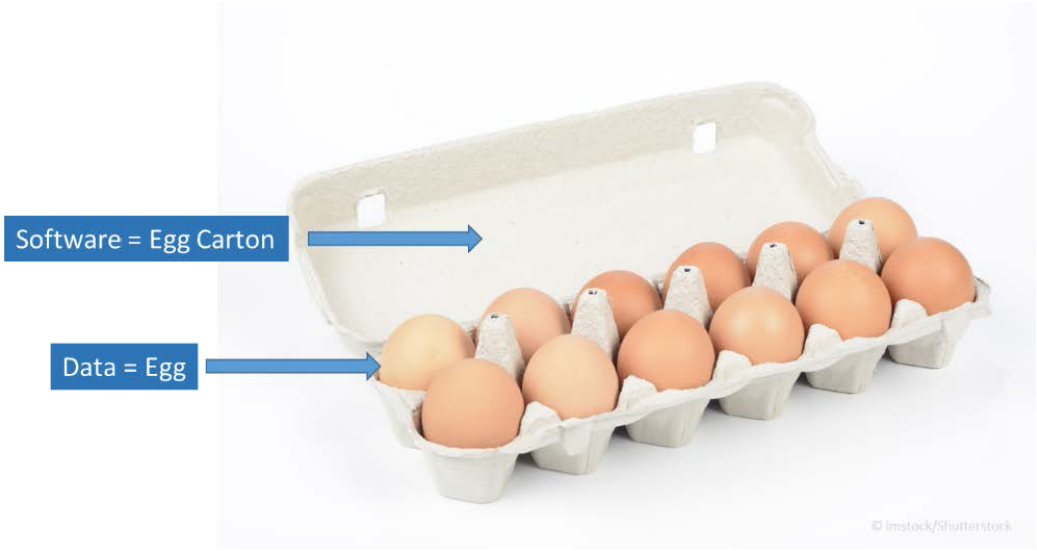


Data schema transformation may raise the following issues:

- Has the data been changed after data schema transformation?
- Does the data handler alter its attribute?
- Does the data organization process change the data?
- How can an operator verify that the data maintain accuracy after the database insertion/organization process?

In order to prevent data changes by the application software during data schema ingest, the software itself must well designed, thoroughly tested, and kept updated. The simple illustration for the relationship between data and software is shown in Section 6, Figure 2. Data is the product of software as software and data always complement each other. Software is the protection for data integrity just like an egg carton protects the egg inside.

FIGURE 2
Data and Software Relationship



The *ISQM Guide* provides guidance on quality software development for integrated control systems. This well-known method of Software Development Lifecycle (SDLC) management guides developers to write quality software that meets stated and documented requirements. During the system design phase, a designer considers data architecture and structure (data schema) in the requirements specification. The developer codes the system software and tests the work against the requirements specification periodically. Such work iterates through the whole lifecycle development phases until the software meets the requirements and it is then delivered.

3.5 Transmission: Transfer from Point of Use Accuracy and Timing

During data transmission, data travels and is transmitted from source to users (machine or human). Previous discussion in Section 5 above provided context, and this section will focus on error detection and correction, network physical and logical security.

3.5.1 Error Detection and Correction

During data transmission noise can be introduced and lead to data errors.

Error detection refers to the detection of error caused by noise or other impairments introduced into data while it is transmitted from source to destination. The most common error detection schemes are listed below:

- Repetition codes
- Parity bits
- Checksums
- Cyclic redundancy checks (CRCs)
- Cryptographic hash functions
- Error-Correcting Code (ECC)

Error correction refers to the reconstruction of data to be error-free. The most common error correction methods are listed below:

- Automatic repeat request (ARQ)
- Error-Correcting Code (ECC)

- Hybrid Schemes

Error detection and correction methods are listed below:

- Error-Correcting Code (ECC)
- Forward Error Correction (FEC)
- Redundant Array of Inexpensive Disks (RAID) Level 2

Both error detection and correction techniques enable the reliable delivery of data over the communication network. These error detection and error correction are used in different applications, from television cameras, RAID and distributed data stores to digital money transfers.

The key technology for data quality enforcement is ECC. Error-correcting codes are commonly used in lower-layer communication, as well as for reliable storage in media such as CDs, DVDs, hard disks, and RAM. An ECC is a process of adding redundant data, or parity data, to a message, such that it can be recovered by a receiver even when a number of errors (up to the capability of the code being used) were introduced, either during the process of transmission, or on storage. Since the receiver does not have to ask the sender for retransmission of the data, a backchannel is not required in forward error correction, and it is therefore suitable for simplex communication such as broadcasting. The standard ECC memory used in systems today can detect and correct the error without user input or action.

Therefore, data accuracy may be expected under most conditions by use of error detection and error correction mechanisms in the receiver device or in the database.

3.5.2 Network Physical and Logical Security

Network security detection methods provide for protection of data against physical threats. . The detailed practices and process specification requirements have been clarified in the *Cybersecurity Guide*.

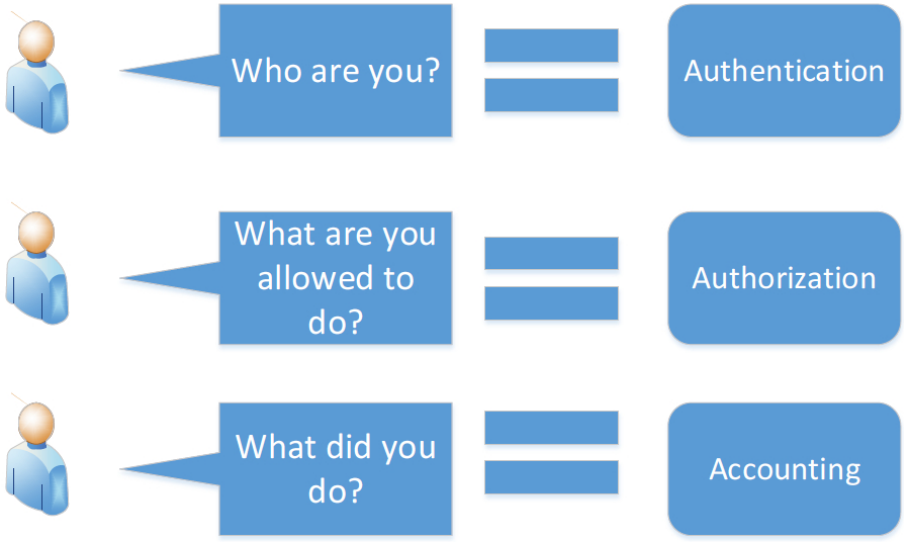
For logical security, software safeguards can provide for data protection. Such software safeguards include user identification, password access, access rights, authentication, and authority level. The final goal is to allow only authorized users to access the organization's data or perform actions on the data.

Authentication, authorization, and accounting (AAA) is a term for a framework for controlling access to information resources. These combined processes are considered for effective network logical security. AAA is a technology that has been in use before the Internet as we know it today.

- *Authentication:* As defined in ISO/IEC 27000 *Information technology – Security techniques – Information security management systems – Overview and vocabulary*, "Provision of assurance that a claimed characteristic of an entity is correct". Authentication provides a way of identifying a user. Usually this is implemented by having the user enter a valid username and password to gain access to resources.
- *Authorization:* A process to define the access policy that gives someone permission to do something. Following authentication, a user must gain authorization for doing certain tasks. This process determines whether the user has the authority to perform the requested task.
- *Accounting:* Measures the resources a user consumes during access. The amount of system time or the amount of data that a user has sent and/or received during the session are typically considered as accounting.

The relationships among these characteristics are depicted in Section 6, Figure 3.

FIGURE 3
Authentication, Authorization and Accounting



With both network physical and logical security protection, the data source is to be kept pure, accurate and reliable.

5 Protect Data Integrity from Unintended and Intended Modification

Data integrity is also protection from data modification. Modification means the data has been changed, updated, altered due to any of multiple reasons. It can be unintended (accidental) or intended (beneficial and malicious). The following subsection is based on the type of data modification.

5.1 Protection against Accidental Integrity Loss

5.1.1 Using the *ISQM Guide*: Data Management Application "Quality" (Adherence to Specifications) and V&V

As described in the earlier two sections, following the traceable and trustworthy data source and secure data transmission, the data has been delivered by suppliers (i.e. sensors) with an approved and accepted level of integrity and reliability. New concerns become, "How do I prevent and detect the loss of integrity due to accidental operation?" and "How can we avoid accidental integrity loss?"

In order to address the new concerns, we follow the same concept as data schema transformation in 6/3.3. The requirements of data integrity must be included in the software requirements to give the best opportunity for correct implementation. The ISQM process methods provide foundational guidance for protecting data integrity through prevention of unintended modification:

i) *Requirements Specifications:* During software requirements specification in the design phase, the circumstances of data changes/modifications are to be considered. This requires all possible scenarios for data modification to be considered and included in the requirements for software development as either deliberately considered acceptable or unsatisfactory conditions. Accidental integrity loss conditions are to be included. Further detailed requirements can be found in the *ISQM Guide* Section 4 Requirements and Design (R&D) Phase.

ii) *Testing Procedures/Protocols:* A complete testing procedure or protocol set is to be developed for each control system. This is also part of ISQM process; test procedures are expected to handle system or component level integrity management to prevent data from being changed or modified in any conditions outside expected function of the control

system. Further detailed requirements can be found in the *ISQM Guide* Section 6 Verification, Validation & Transition (V V&T) Phase.

iii) *Software Management of Change (SMOC) Process:* SMOC is the process for managing authorized software changes in both critical and non-critical systems. SMOC is a systemic means of managing software versions, integrity and interoperability conditions by preventing untested, unauthorized software or data changes. Further detailed requirements can be found in the *ISQM Guide* Section 7 Operation and Maintenance (O&M) Phase.

5.1.2 Using the *Cybersecurity Guide*: Logical Access Authorization Control

Logical access authorization control is another way to prevent accidental integrity loss. It is often needed for remote access of hardware, and it is contrasted with the term "physical access". It includes methodologies such as password protocols, personal identification numbers (PINs), biometric scans, etc. The requirements have been set in the *Cybersecurity Guide* and represented as follows:

i) *Capability (5): Provide Perimeter Defense*. "*OT5-2:* The Company provides isolation of logical access to OT/ICS to confirm traffic to control systems can only originate from authorized sources within the Company's environment."

ii) *Capability (8): Execute Access Management*. "*P8-5:* The Company tracks operational or process control assets that do not have either physical or logical access control mechanisms, substituting access process controls as required to confirm positive knowledge of personnel or machine access to control systems or components." "*OT8-2:* The Company defines and implements role-based business rules for logical access to any Company ICS. Authorization of access is based on job function requirements and risk assessment processes."

iii) *Capability (9): Maintain Asset Management*. "*OT 9-4:* The Company protects against unauthorized logical access to proprietary operational and protective systems using mixtures of logical and physical methods, including (but not limited to) segregated communications paths, screening mechanisms, access control processes, strong passwords, and/or multifactor authentication."

iv) *Capability (20): Provide Unified Identity Management*. "*OT20-2:* The Company defines and implements role-based business rules for logical access to any Company ICS. Authorization of access is based on job function requirements and risk assessment processes." "OT20-8: The Company protects against unauthorized logical access to proprietary operational and protective systems using segregated communications paths, screening mechanisms, access control processes, identity system enrollment and role designation, strong passwords, and/or multifactor authentication."

v) *Capability (35): Implement Secure Software Development*. "*IT35-11:* The Company confirms that software code repositories are closely managed, carefully restricted for personnel access, and strictly limited in avenues of logical access through any network, with accountability for any access event."

5.3 Best Practices for Intended (Beneficial) Integrity Loss

Intended (beneficial) integrity loss may be due to modification which is intended and based on cleansing routines that enable better analytics to be performed. The essential goal of data cleansing process is to find a suitable balance between fixing dirty data and maintaining the data as close as possible to the original data from the source, and be ready for better data analytics. Therefore, such modification is routine based, planned and beneficial. Generally the approaches such as data auditing, workflow specification, workflow execution, and post-processing and controlling are the common practices to maintain data integrity against intended (beneficial) integrity loss.

5.5 Preventive Actions Against Intended (Malicious) Integrity Loss

Intended (malicious) integrity loss may be due to purposeful actions by human or automated entities, and data quality and security may suffer as a result of data theft or data destruction. Access control and penetration control are two major security techniques for protection against malicious integrity loss.

5.5.1 Using the Cybersecurity Guide: Access control

As defined in *ISO/IEC 27000 Information technology - Security techniques - Information security management systems - Overview and vocabulary*, "Access Control is a means to ensure that access to assets is authorized and restricted based on business and security requirements". It is a security technique that can be used to regulate who have the authorization to access the resources. It allows only the authorized users touch the system. Anyone without correct identity should not able to access the system.

This security technique prevents the malicious attacks from early stage of data integrity. The detailed requirements have been set in the *Cybersecurity Guide* and the high level requirements are listed here,

i) Access Control Lists (ACLs) for physical possession or contact with system assets (devices, systems, workstations, servers, PLCs, network protocol translators, network connections, etc.) are established and kept up to date.

ii) ACLs are to be established in any remote access methods for all operational technology or process control system components, systems, modules, applications or appliances.

iii) ACLs are one of the perimeter defense security methods to be established against unauthorized access.

iv) Physical access control devices are to be installed and monitored as designed and described in the system Functional Description Document (FDD).

v) Network access control (NAC) is to be used to enforce uniform enterprise system policies and hygiene across all access points.

vi) Access control is required on all software.

5.5.2 Using the *Cybersecurity Guide*: Penetration control

Penetration control and information security methods are security techniques that use penetration testing to find potential security weaknesses in the IT/OT system. Design and production of these controls should commence in the architecture design phase of the system.

Examples of penetration control are listed below:

- Segregation of the system/network
- Blocking unauthorized access methods, such as use of unauthorized thumb drives
- Security implemented by using firewall or malware

Detailed requirements for exercise penetration testing (capability 27) have been set in the *Cybersecurity Guide*.

7 Monitor, Verify, Validate and Measure Data Integrity

7.1 Data Integrity Monitoring (Overall System of Systems)

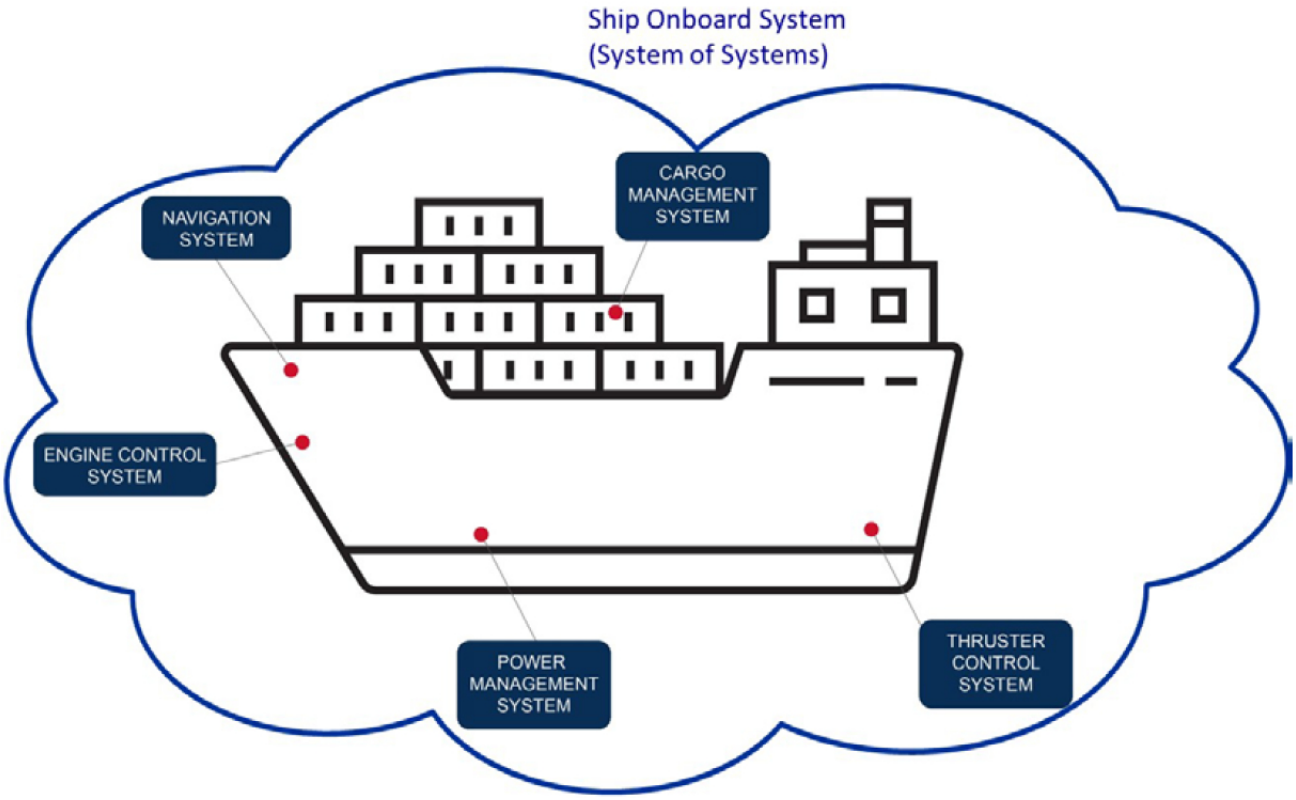
Monitoring of data integrity is to provide detection capability against otherwise undetected integrity effects or losses. Critical data is to be monitored continuously, with alarms or notifications set for fault conditions. Fault conditions are established for monitoring as part of requirements specification, as noted in 6/5.1 above. Normally data managers or data analysts check unknown data against good data (reference data) so

they can recognize differences in data content and trends, especially when comparing new data against reported norms.

Data integrity monitoring starts with a single component or single system. Once ‘normal’ is recognized and codified in monitoring procedures, the monitoring process can extend to the system level and to the overall System of Systems.

An integrated System of Systems is a way to conceptualize the boundary defining a system. This concept was introduced to marine and offshore industry due to more and more complex and integrated systems onboard the vessel. This is most appropriate for working with software-intensive control systems as they propagate across ships or offshore platforms. Section 6, Figure 4 illustrates one example of ship network systems onboard.

FIGURE 4
Notional System of Systems Onboard Vessel



In the maritime industry, System of Systems means ship, offshore platform, or other maritime asset with multiple existing standalone and networked systems. The data associated with a given individual system may not be limited to that particular system. It may be accessed and used by many other systems that connect to it, and special attention is required to understand how data can be exposed to the overall system of systems. Monitoring of data integrity will also be extended not only to component or single system level but also to the overall system of systems.

The standard regarding Data Continuous Monitoring have been set in the *Cybersecurity Guide*.

7.3 Verification and Validation of Data Integrity

Using the definitions of Validation and Verification (V&V) from the ABS *Guide for Software Systems Verification – ABS CyberSafety™ Volume 4 (SSV Guide)*:

Validation: Determination that an item (system) is suitable for the intended service.

Verification: Determination that an item (system) meets the specified criteria.

The V&V of a system for purpose of data integrity management can be accomplished by system performance testing. These tests include initial testing, final product testing, and periodic testing. The initial testing is always performed during development; final product testing provides verification of requirements satisfaction prior to deployment; and periodic testing is done at intervals to verify continued correct operation. These system test phases should include data verification during the extent of the tests.

The *SSV Guide* has requirements for verification. The following are the general guidelines:

- Validation and Verification plan
  - Scope of verification (what is to be tested)
  - Details of testing including test cases & test methods
- Validation and Verification report
  - Each test phase is to generate a report
  - Report is expected to include traceability, test cases, test results, comparison

7.5 Measurement of Data Integrity

Data integrity is of pivotal importance to marine system safety and functionality. Conceptually, the simplest form of data integrity is accuracy. If sensed data accurately describes or represents the characteristics of a physical or virtual system accurately, then that data representation is said to possess integrity. Previous sections of this document describe methods for analyzing initial data integrity and detecting corruption during transport.

Even with well-conceived and executed methods for preventing and detecting losses in data integrity, corruptions and integrity losses do occur. Since these losses can and do result in system failures, another way to view the data integrity preservation process is to consider methods for analyzing observed system failures to determine if a root or contributing cause of the failure is data corruption (i.e., data integrity loss).

One method is to use the Corruption Vector Index (CVI) for gauging system integrity. This approach gauges the relative importance of a Corruption Vector. Corruption Vector is generally characterized by "How", "Where/When", "How much", and "Who/What".

- The "How" characteristic is the observed behavior of the fault.
- The "Where/When" characteristic is its' corruption position in the SDLC lifecycle and the control system architecture. Further detailed information regarding SDLC may be found in the *ISQM Guide*.
- The "How much" characteristic is the relative impact on the control system.
- The "Who/What" characteristic refers to the source of the corruption.

The numeric values are assigned to each of the systemic conditions that determine the CV's position in the control system, and the technology conditions that determine the CV's impact. Adding Systemic Value and Incident Value yields Corruption Vector Index.

As shown in Section 6, Figure 5 below, Systemic Value is produced by multiplying the values rated in each category. A low result represents a low corruption value and indicates the best system integrity maintaining. In contrast, a high Gauged System Value result indicates a high corruption value, and represents the system at a high risk of corruption.

FIGURE 5
Gauged System Value

A	B	C	D	E	F	G	H
Supplier X, Incident n		Location of Corruption in SDLC		Systemic Integrity Maintaining Measures - Actions Gauge			C*D*E*F*G
Incident	Corruption Vector Description (1)	SDLC Discovery Point (2)	SDLC Remediation Point (3)	V&V (4)	Cyber Security (5)	MOC (6)	Gauged Systemic Value
1	Failure Incident ID	2	1	1.0	1.0	1.0	2
2	Failure Incident ID	1	3	1.1	1.1	1.1	4
3	Failure Incident ID	1	1	1.0	1.0	1.0	1
4	Failure Incident ID	2	5	1.1	1.0	1.1	12
5	Failure Incident ID	2	4	1.0	1.1	1.0	9
6	Failure Incident ID	3	3	1.1	1.0	1.0	10
7	Failure Incident ID	3	5	1.1	1.0	1.1	18
8	Failure Incident ID	4	3	1.1	1.1	1.0	15
9	Failure Incident ID	4	4	1.1	1.1	1.1	21
10	Failure Incident ID	5	5	1.0	1.1	1.0	28

Notes:

1 The SDLC penetration point (corruption source) is determined by analysis of discovery point, remediation point, and values of system integrity maintaining measures.

2 Incident discovered at SDLC point: 1 - Concept Stage; 2 - Requirement/Design; 3 - Construct/Dev; 4 - V V&T; 5 - Operation/Maintenance.

3 Incident corrected at SDLC point: 1 - Concept Stage; 2 - Requirement/Design; 3 - Construct/Dev; 4 - V V&T; 5 - Operation/Maintenance.

4 During V&V, Supplier provides description and a pre-installation test report with the change: 1 - Yes; 1.1 - No

5 Cybersecurity measures implemented: 1 - Yes; 1.1 – No

6 A rigorous detailed MOC process is used: 1 - Yes; 1.1 - No

As shown in Section 6, Figure 6, Incident Value is produced in a similar way as the Systemic Value. Incident technology penetration point (description of function) is listed in the first column of the table, following columns are filled with Integrity Level (IL) of Technology Discovery Point, Corrective Change Impact, and Failure State Impact (detailed explanation shown inside the table). The final Gauged Incident Value is generated by multiplying the values rated in each category. The low result represents the low incident possibility and indicates the individual function has high integrity level, low (or no) change impact and low (or no) failure state. The high result represents the opposite.

FIGURE 6
Gauged Incident Value and Corruption Vector Index

J	K	L	M	N	O	P	Q
Supplier X, Incident n		Impact Measures - Gauge			L*M*N		C*D*E*F*G + K*L*M
Function	Technology Penetration Point[1]	IL of Technology Discovery Point[2]	Corrective Change Impact[3]	Failure Sate Impact[4]	Gauged Incident Value[5]		Corruption Vector Index
1	Failure Incident ID	1	1	1	1		3
2	Failure Incident ID	2	2	3	12		16
3	Failure Incident ID	3	4	2	24		25
4	Failure Incident ID	4	3	4	48		60
5	Failure Incident ID	1	2	3	6		15
6	Failure Incident ID	2	3	2	12		22
7	Failure Incident ID	3	4	1	12		30
8	Failure Incident ID	4	2	3	24		39
9	Failure Incident ID	2	1	4	8		29
10	Failure Incident ID	4	4	2	32		60

Notes:

1 Description of the specific Industrial Control System (ICS) function at which the corruption incident is discovered.

2 Integrity Level (IL) of the function at which the incident was discovered: 1 - IL0, 2 - IL1, 3 - IL2, 4 - IL3

3 1 - No change, 2 - Minor Change, 3 - Major Change, 4 - Emergency Change

4 1 - Redundant, immediate recovery, 2 - Stop safe-Reset-Resume, 3 - Stop safe, repair required, 4 - Stop unsafe

The final Corruption Vector Index is generated by adding the Systemic Value and Incident Value. By tracking functions and suppliers over time, a trend chart will be produced. This chart is to gauge Corruption Vectors and its impact to its respective control system integrity.

When system failures happen, an incident discovery point is to be identified, describing the location of failures in the SDLC stage, such as in SDLC early stage, middle or later stage. The supporting data corresponding to its respective function Integrity Level becomes an important cause of in system failures. For example, Mission Critical data drives the system with IL3 functions. When any failure occurs in the system with IL3 functions, it is possible for Mission Critical data to be being corrupted. Thus, it helps determine whether the source of the system failure is due to data corruption. Further failure analysis helps to determine if data corruption is a root or contributing cause of a failure. The CVI approach also helps to determine if the level risk to the system created by data corruption.

The CVI and the associated failure analysis results indicate important information, one of which is the data corruption source. The data corruption source could be from supplier, installer, or updater. For example, we can confirm the number of times the supplier touched the system, which could open a corruption path. Every time someone touches the system, it can be corrupted. Therefore we can determine the relative quality of supplier by checking their practices on handling the data, some as whether the original software was well delivered, database has been maintained well, and the update has been successfully performed without affecting data integrity. If any of these steps are on track, the final risk of data corruption will be decreased significantly.

The CVI approach also helps improve data integrity by implementing methods that would prevent similar data corruptions in the future. Through the analysis result, the data corruption has been identified and action is to be taken to prevent similar data corruption from reoccurring. Also by tracking the data

corruption over time, measurable patterns of data corruption could be attributed to specific practices or data handlers. These should be targeted for remedial action.