4.2.4.2 Data Integrity

Return to Securability

About

Return to Top

Data Integrity is the completeness, accuracy and consistency of data throughout the entire data lifecycle of the data as well as when the Data is at Rest, Data-in-Motion and Data-in-Use.[1)]

Figure 1 shows the five levels Automation Pyramid and the functionality usually associated with each one. There is Data-at-Rest at each level of the pyramid. As the data transitions up and down from level to level within the pyramid, the Data is in Motion. Within each level, the data will most likely be accessed therefore, the Data is in Use.

Table 1: The five levels of the Automation Pyramid.

Automation Level	Description
Field Level	The Field Level where products are produced. In other words, this is where the physical work plus monitoring occur. Electric motors, hydraulic and pneumatic actuators to move machinery, proximity switches used to detect that movement or certain materials, photoelectric switches that detect similar things will all play a part in the field level.
Control Level	The Control Level uses the the control devices to “run” the devices in the Field Level. The Control Devices make decisions based on information provided by sensors, switches, and other input devices to complete the programmed task.
Supervisory Level	The Supervisory Control and Data Acquisition (SCADA) is combines the Field and Control Levels to provide oversight from a single location. This is usually accomplished using Graphical User Interface (GUI), or Human-machine interface (HMI), to remotely control operations. For example, water plants often employ this technology to control remote water pumps.
Planning Level	The Planning Level uses Manufacturing Execution System (MES) to monitor the entire manufacturing process. For example, in a factory to plan for everything from raw materials to the finished products. This allows management to visualize the current state of operations and aids them in making decisions and adjust raw material orders or shipment plans based on real data received from Supervisory, Control and Field Levels.
Management Level	The management level uses the companies integrated management system such as as Enterprise Resource Planning (ERP). Corporate management visualize and control operations. This level allows the businesses monitor all levels (i.e., manufacturing, to sales, to purchasing, to finance and payroll). The integration of an ERP promotes efficiency and transparency within a company by helping to communicate the levels.

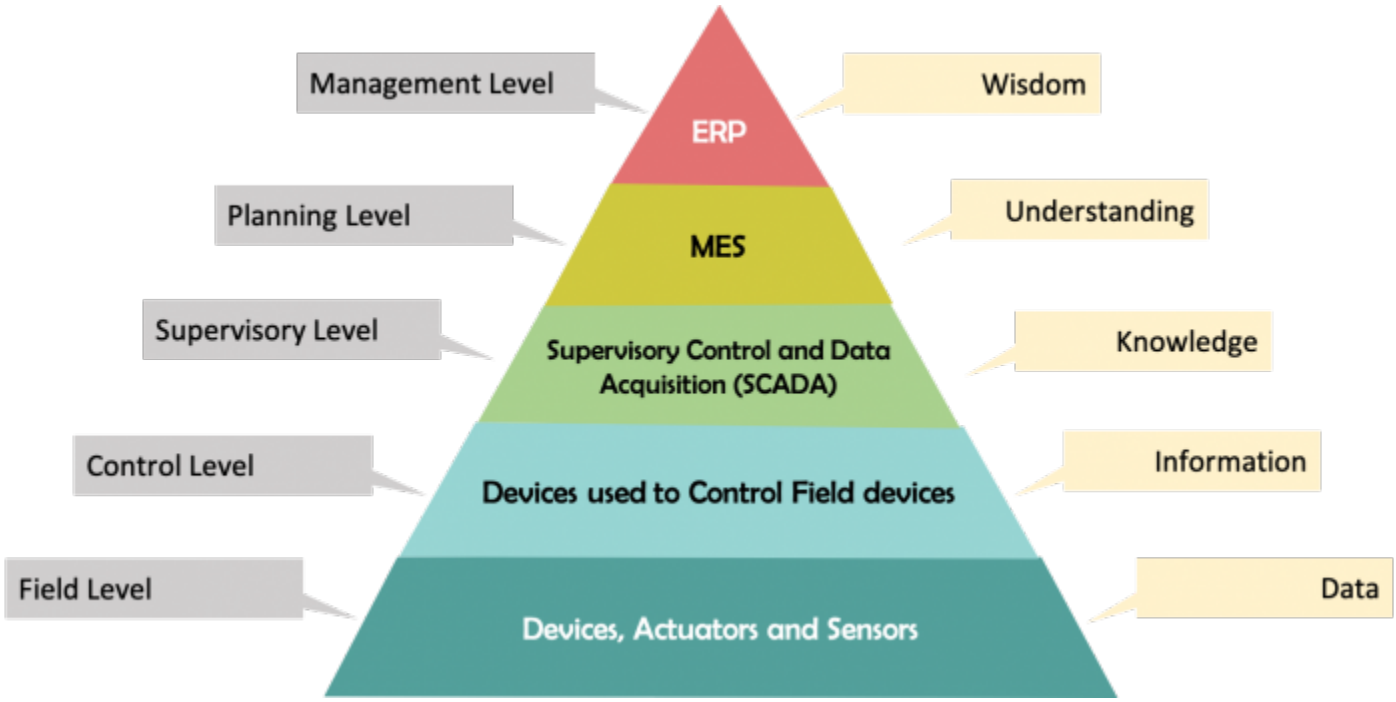

Figure 1: Automation Pyramid

At each level, the Data at Rest can be categorized as one of two kinds of data integrity both of which are a collection of processes and methods intended to enforce Data Integrity.

- Physical Integrity protects data's wholeness and accuracy as it's being used. When expected or unexpected down times occur (i.e., natural disasters strike, power goes out, or hackers disrupt database functions) physical integrity is compromised. Some other issues which can compromise the integrity of the data are Human error, storage erosion, or a host of other issues making it impossible for data processing managers, system programmers, applications programmers, and internal auditors to obtain accurate data.

- Logical Integrity keeps data unchanged as it is accessed. Logical integrity protects data from some of the same issues as Physical Integrity (i.e., human error and hackers as well) but in different ways. There are four types of logical integrity.
  1. Entity Integrity - supports unique values that identify any particular data entry and that the key is not null.
  2. Referential Integrity - ensures that references to other data entries exists.
  3. Data Integrity - ensures that domain rules (i.e., data restrictions) are enforced for the data within the Data Structure. For example, minimum, maximum, number of decimals, nullable, etc. are enforced.
  4. userdefintegrity - ensures that business rules are enforced. For example, if a value is set, another value must also be set (unset); if a value exceeds a threshold, a notice must be sent.
  - **Note:** Data integrity is not Data Security and is not Data Quality.
    - Data Security defines the steps taken in the prevention corruption and from corruption by outside attacks by people or processes.
    - Data Integrity defines the steps taken to keep the data intact and accurate from internal people and processes and for the entirety of the data's existence.

DDS Specifics

Return to Top

data_distribution_service_dds Systems are Distributed Systems and are generally architected so Data Integrity is almost an after thought.

It is important to remember that systems using DDS are distributed by definition. These systems are actually a network of DDS Nodes. Each node provides functionality at an appropriate automation level (i.e., Field Level through Management Level) of the Automation Pyramid). For example:

1. The Field Level is the production flow that does the physical work and monitoring
2. The Control Level the information from all then sensors is collected level and is used to make decision
3. The Supervisory Level that is the human layer or Supervisory Control and Data Acquisition (SCADA) the information that is used to access the data to control the systems from one single location and plus it usually adds some graphical user interface for plant managers aspect.
4. The Planning Level (Manufacturing Execution System (MES)) it monitors the entire manufacturing process in a plant or a factory
5. The business logistics (Cloud) layer - it monitors or levels of from manufacturing to sales to purchase and to finance

A well defined system also provides for multiple nodes that provide redundant functionality. For example, in a Chemical Mixing Process, there may be multiple Sensors (i.e., nodes) providing temperature for the mixing vat and publishing to the same Topic. There are also multiple nodes listening (i.e., Subscriber to the data being published). If the data being published at the sensors were not identical in structure, format, and content then the redundancy creates more work for the subscribers. In DDS, if one of the publisher were to fail, the subscribers would automatically get the data from the second sensor.

In a like way, there can be multiple subscribers at any particular automation level, or the subscribers could be at different levels in the automation pyramid. For example, one subscriber at the Field Level, and one at the Control Level, etc. Each level may be using a different programming language. For example, the Field Level might be using C/C++, the Control Level might use Java, the Supervisory Level might use Python or R, the Planning level might use SQL or Javascript, and the Business Logic might use executable Business Process Model and Notation (BPMN). As long as there is a mapping between the programming language and the idl data structures, Data Integrity is preserved and there is no problem with each language using the same Topic.

1)
What is Data Integrity, Accessed 8 July 2020, https://www.talend.com/resources/what-is-data-integrity/

From:
https://www.omgwiki.org/dido/ - **DIDO Wiki**

Permanent link:
https://www.omgwiki.org/dido/doku.php?id=dido:public:ra:1.4_req:2_nonfunc:25_security:04_data_integrity&rev=1605490493

Last update: **2020/11/15 20:34**