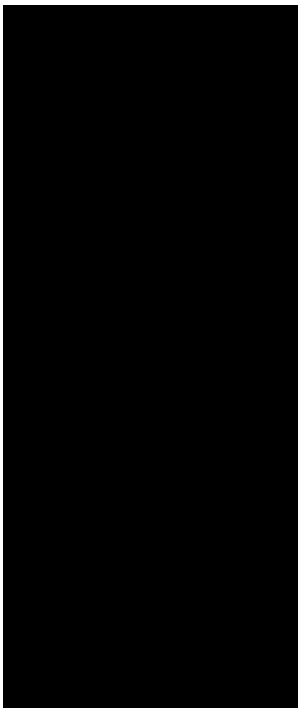
Data Integrity Reading

Data integrity refers to the reliability and trustworthiness of data throughout its lifecycle. It can describe the state of your data—e.g., valid or invalid—or the process of ensuring and preserving the validity and accuracy of data. Error checking and validation, for example, are common methods for ensuring data integrity as part of a process.

Complete, consistent, and accurate data should be attributable, legible, contemporaneously recorded, original or a ‘true copy’, and accurate **(ALCOA)**.

The following video categorizes these characteristics in five principles of Data and links them to Data cleaning; We will be covering Data Cleaning this week.



0:00

*[Video Source: Coursera - Exploring and Preparing your Data with BigQuery **(Link)** **(https://www.coursera.org/lecture/gcp-exploring-preparing-data-bigquery/examine-the-5-principles-of-dataset-integrity-VnCzN)**]*

NB: Data integrity is not to be confused with data security. Data security refers to the protection of data, while data integrity

refers to the trustworthiness of data.

ALCOA Framework

The acronym ALCOA has been around since the 1990's, is used by regulated industries as a framework for ensuring data integrity and is key to Good Documentation Practice (GDP). ALCOA relates to data, whether paper or electronic and is defined by US FDA guidance as Attributable, Legible, Contemporaneous, Original and Accurate. These simple principles should be part of your data life cycle, GDP and data integrity initiatives.

- **Attributable:** Information should be captured in the records so that it is uniquely identified as having been executed by the originator of the data (e.g. a person or computer system).
- **Legible, traceable, and permanent:** Data should be readable, understandable, and should allow a clear picture of the sequencing of steps or events in the record so that all activities conducted can be fully reconstructed by the people reviewing these records at any point during the records retention period.
- **Contemporaneous:** Data should be recorded at the time when the data are generated or observed.
- **Original (or "True Copy"):** Data should be in the format in which it was originally generated, preserving the integrity

(accuracy, completeness, content and meaning) of the record.
    - World Health Organization: Original data include the first or source capture of data or information and all subsequent data required to fully reconstruct the conduct of the activity.
- **Accurate:** Data should be correct, truthful, complete, valid and reliable.

Why is Data Integrity Important?

Imagine making an extremely important business decision hinging on data that is entirely, or even partially, inaccurate. Organizations routinely make data-driven business decisions, and data without integrity, those decisions can have a dramatic effect on the company's bottom line goals.

A new report **(https://www.zdnet.com/article/most-executives-dont-trust-their-organizations-data-analytics-and-ai/)** from KPMG International reveals that a large majority of senior executives don't have a high level of trust in the way their organization uses data, analytics, or AI. Only 35% say they have a high level of trust in the way their organization uses data and analytics. 92% are concerned about the negative impact of data and analytics on an organization's reputation. What's more, 62% of senior executives said technology functions, not the C-level and

functional areas, bear responsibility when a machine or an algorithm goes wrong.

Organizations need to go through the motions of preserving data integrity in order for C-level executives to make proper business decisions.

Case Studies

To understand the pressing ramifications of the data integrity issue, we must remind ourselves that data is the fuel of the digital economy. In the pursuit of increased efficiency, digital approaches are being used for risk reduction and greater innovation in industry life cycle processes—from discovery to commercial manufacturing—and must continue to be further leveraged.

Additionally, data must become less siloed and flow much more seamlessly throughout organizations at various stages of the product life cycle. However, just as if gasoline in a car engine is contaminated, it will damage and/or cease the engine, the same is true for “digital fuel.” Correct and uncorrupted data must flow through an organization so that correct and reliable decisions can be made. Even without sophisticated digital data management considerations, data integrity of even the most basic data systems must be assured to ensure compliance.

Below are some case studies which provide implications of non-compliance to principles outlined in the ALCOA framework. Your task will be to identify which aspect of the framework wasn't taken into consideration?

- **New England Compounding: Meningitis Outbreak 2012**
  - Pharmacy technicians instructed to prioritize production over cleaning and disinfecting.
  - Pharmacy technicians instructed to falsify cleaning records, showing rooms were properly cleaned when they had not been neglected to investigate contamination found in the clean rooms.
  - The company distributed orders even though they were still waiting for sterility test results.
  - "As alleged in the indictment, these employees knew they were producing their medication in an unsafe manner and in unsanitary conditions, and authorized it to be shipped out anyway, with fatal results," said Attorney General Eric Holder.
  - Results
    - 64 reported deaths
    - More than 800 patients sickened
    - President sentenced to 9 years in prison
    - Other employees charged with multiple criminal acts
    - Led to the passage of the Compounding Quality Act. [Link **(https://www.hhs.gov/about/agencies/asl/testimony/201**

8-01/examining-implementation-of-the-compounding-quality-act.html)]

- Source. [Link] (https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/criminal-investigations/january-31-2018-new-england-compounding-center-pharmacist-sentenced-role-nationwide-fungal)
- *Question: Which principles of ALCOA weren't taken into consideration?*

- **Takata: Auto Airbags 2015**
  - 15 deaths
  - 100 injuries
  - 100 million vehicles worldwide,
  - ~33 automotive brands Defective airbags still in cars.
  - Potential danger of spraying shrapnel caused by defective airbag inflators when the airbag goes off.
  - Takata engineers removed some test results to artificially reduce variability in air-bag inflator performance.
  - "Takata provided inaccurate, incomplete and misleading information to regulators for nearly a decade," said NHTSA spokesman Bryan Thomas. "Had they told the truth, Takata could have prevented this from becoming a global crisis."
  - Results:
  - Source. [Link] (https://www.usatoday.com/story/money/2017/06/25/takata-

air-bag-scandal-timeline/103184598/)

- *Question: Which principles of ALCOA weren't taken into consideration?*

- **Koito Industries: Airline Seats 2010**
  - Developed software that would display acceptable-looking readouts on screens whenever inspectors from the Transport Ministry came to observe the testing procedures
  - Failed to perform a critical part of one of the tests and had applied the results of previous tests to newer products that had not been made to the same standards
  - Fabricated test results for flammability and static and dynamic strength tests;
  - Failure to meet static and dynamic strength criteria could result in injuries to the flight crew and passengers during emergency landing conditions - FAA (US)
  - Results
    1. 150,000 suspect seats installed in 1000 aircraft EASA
    2. FAA (US) issued airworthiness directives to test and correct
  - Source. **[Link] (https://www.nytimes.com/2010/12/07/business/global/07iht-ravseat.html)**
  - *Question: Which principles of ALCOA weren't taken into consideration?*

- **Peanut Corporation of America: Salmonella Outbreak 2008-09**
    - In some cases, company officials falsified lab results, stating peanut products were safe to eat when tests showed otherwise, or when products had never been tested at all, according to court papers. - The Wall Street Journal
    - The company shipped the product with falsified Certificates of Analysis (COA), which attested to the purity of contaminated lots
    - CEO wrote in a March 2007 email to a plant manager about contaminated products: "Just ship it. I cannot afford to lose another customer."
    - Results
        1. 9 reported deaths
        2. More than 700 consumers sickened
        3. CEO sentenced to 28 years in prison;
        4. Others sent to prison including Plant Quality Manager
        5. Loss of nearly $1 billion in peanut sales Plant closed & company liquidated.
    - Source: [Link] **(https://www.ers.usda.gov/webdocs/publications/37835/8684_ocs10a01_1_.pdf)** .
    - *Question: Which principles of ALCOA weren't taken into consideration?*

Risk Assessment & Management strategy

You might also end up undertaking a managerial role in your organization. Thus, a management strategy for your firm would include the details of your global corrective action and preventive action plan. More specifically, your strategy would include:

- A detailed corrective action plan that describes how you intend to ensure the reliability and completeness of all of the data you generate, including analytical data, manufacturing records, and all data submitted to regulatory bodies i.e. FDA.
- A comprehensive description of the root causes of your data integrity lapses, including evidence that the scope and depth of the current action plan is commensurate with the findings of the investigation and risk assessment. Indicate whether individuals responsible for data integrity lapses remain able to influence CGMP-related or product application data at your firm.
- Interim measures describing the actions you have taken or will take to protect users and ensure the quality of your product, such as notifying your user on product updates, conducting additional testing, adding lots to your stability

programs to assure stability, and enhanced complaint monitoring.

- Long-term measures describing any remediation efforts and enhancement to procedures, processes, methods, controls, systems, management oversight, and human resources (e.g., training, staffing improvements) designed to ensure the integrity of your company's data.
- A status report for any of the above activities already underway or completed.

References

1. Cindy Ng. (2018, May 5). What is Data Integrity and How Can You Maintain it? Retrieved from https://www.varonis.com/blog/data-integrity/
2. Sharon K. Pederson. (2017, February 06). Data Integrity Issues and Concerns. Retrieved from https://www.pda.org/docs/default-source/website-document-library/chapters/presentations/missouri-valley/data-integrity-issues-concerns.pdf?sfvrsn=4
3. Peter Dellva. (2017, May 10). Why Pharmaceutical Data Integrity Is More Important Than Ever? Retrieved from www.biotechlogic.com/why-pharmaceutical-data-integrity-is-more-important-than-ever/