`

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
Data	All original records and true copies of original records, including source data and metadata, and all subsequent transformations and reports of these data which are generated or recorded at the time of the GMP activity and which allow full and complete reconstruction and evaluation of the GMP activity. Data should be accurately recorded by permanent means at the time of the activity. Data may be contained in paper records (such as worksheets and logbooks), electronic records and audit trails, photographs, microfilm or microfiche, audio or video files or any other media whereby information related to GMP activities is recorded. ➢ The values and information generated by processing, calculating or transcribing from the raw data. This may include computer printouts. ➢ Data is a set of values of qualitative or quantitative variables ➢ pieces of data are individual pieces of information.	**Data must be**: • A - attributable to the person generating the data ✓ Clearly indicates who recorded the data or performed the activity ✓ Signed / dated ✓ Who wrote it / when • L - legible and permanent ✓ It must be possible to read or interpret the data after it is recorded ✓ Permanent ✓ No unexplained hieroglyphics ✓ Properly corrected if necessary • C - contemporaneous ✓ Data must be recorded at the time it was generated

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
	➢ In case of computers data is the numerical, quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media. ➢ Data is Information derived or obtained from raw data. (e.g. a reported analytical result)	✓ Close proximity to occurrence • O - original (or 'true copy') (un-tampered) ✓ Data must be preserved in its unaltered state ✓ If not, why not ✓ Certified copies • A - accurate ✓ Data must correctly reflect the action / observation made ✓ Data checked where necessary ✓ Modifications explained if not self - evident **Original record**: Data as the file or format in which it was originally generated, preserving the integrity (accuracy, completeness, content and meaning) of the record, e.g. original paper record of manual observation, or electronic raw data file from a computerised system ➢ Data is measured, collected, reported and analysed. ➢ Data can be visualized using graphs or images.

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
		➢ Data as an general concept refers to the fact that some existing information or knowledge is represented or coded in some form suitable for better usage or processing. ➢ Data represents things known or assumed as facts, making the basis of reasoning or calculation.
ALCOA+	A commonly used acronym for "attributable, legible, contemporaneous, original and accurate" which puts additional emphasis on the attributes of being complete, consistent, enduring and available throughout the data life cycle for the defined retention period.	**ALCOA Plus** Complete: All data is included (e.g. testing, repeat, or analysis) Consistent: Consistent generation of records and application of time stamps Enduring: Data recorded on controlled worksheet or invalidated systems Available: Data available for review, audit, inspection for the life of record
Attributable	This term refers to the need to be able to determine who performed each action. If there were changes made, who made the change and why.	**Examples:** ✓ Failure to have audit trails for the system and to ensure that they are working all the time. ✓ For manual systems, it is the failure to have traceability

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
	In all cases, there needs to be a strong link to the original source data.	throughout the documents to show who performed tasks, changed information and the rationale for changes, e.g., cross-outs that are not showing the reason for change and/or who made the change. ✓ Using other people's passwords or a single password for a group as this limits the ability to determine who exactly did the work and who made the changes. Another variation is the posting of passwords, so that anyone can use them.
Legible	The data must be documented using permanent ink and should allow for the reader to identify all entries. (Manual systems) Printouts should be legible, e.g., not fade over time.	✓ Handwriting cannot be identified as to meaning upon review. ✓ Failure to make corrections in a way that allows the original information to be read, e.g., not crossing out with a single line through the entry. ✓ Use of thermal printouts and not making copies so that the print will not fade out over time.
Contemporaneous	All data should be recorded at the time the work is performed. All date and time stamps should be in order (based upon date and time).	Writing notes on post-its to enter into the log book or data form at a later date.

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
Original	Is the document the original (raw) data? This should be the first time the information is recorded. In some cases, the original may not be available but a "certified true copy" is available e.g., a copy may be from a thermal printer and photocopied to preserve the printing. It should be signed and dated with wording that this is a certified copy.	✓ Fabricating of data for which there is no original data ✓ Copying data and claiming it to be new and original data ✓ Unauthorized access to computer system to modify, delete, or not save electronic files is not prevented
Accurate	This refers to the data being entered without errors or editing. If editing occurred, it must be properly documented, e.g., audit trail, traceable to original data.	Data which has been falsified, e.g., failure to run experiments to generate the data. Backdating the dates on the data as if it occurred earlier Coping existing data as if it occurred this time Turning audit trails off to hide data changes Unauthorized access to computer system to modify, delete, or not save electronic files is not prevented Releasing product that does not meet the product specifications (i.e., failing product)
Complete	All of the data generated is included in the analysis. This includes all runs, whether good or bad. In some cases, data may not be used in an	✓ Failure to maintain all of the data generated for a test, e.g., eliminating bad testing data and only keeping part of the data

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
	analysis, but it is addressed in a deviation or investigation and shown to be invalid.	✓ All of the analysis (good or bad) for the sample was maintained ✓ For chromatographs, not including data for all injections ✓ Testing samples into compliance ✓ Unauthorized access to computer system to modify, delete, or not save electronic files is not prevented ✓ Re-running samples to get desired results
Consistent	This refers to the consistent use of date and time stamps and that the data is collected/reported in the proper sequence (as expected).	✓ Evidence that data is backdated ✓ Date and/or time markings are out of order
Enduring	The original data is recorded in controlled records, e.g., controlled (numbered) worksheets, laboratory notebooks (bound) or electronic media.	✓ Use of post-its or scrap paper to record data and later transfer the data without identifying it as transcript data ✓ Failure to save electronic or manual records
Available	One can access the data throughout the lifetime of the record (and the associated retention period required).	Failure to maintain the records for the product lifecycle (and any other applicable retention periods) Failure to save electronic or manual records

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
Record	➢ A record is a collection of elements, typically in fixed number and sequence, indicating chronology of the actions and typically indexed. ➢ In computer science, a record (also called structure or compound data) is a basic data structure. A record is a data type that describes such values and variables.	➢ Records can exist as paper document or electronic in any storage medium, including main memory and mass storage devices such as magnetic tapes or hard disks. ➢ Records should be traceable to the doer (name / Id of the person), date and time; it's also an evidence of activity performance. ➢ Record should be retained as per specified retention period and hence shall not be destroyed before this period.
Report	Report is document which is prepared for any informational work (usually of writing, or an electronic entry) made with the specific intention of relaying information or recounting certain events in a presentable manner.	➢ Reports are often used to display the result of an experiment, investigation, or inquiry. ➢ Reports use features such as graphics, images, voice, or specialized vocabulary in order to pursue to undertake an action.
GMP Documentation	Any procedure, instruction, logbook, record, raw data, manual, or policy associated with the development, manufacture, testing, marketing, and distribution of a medicinal product required to demonstrate compliance with the GMPs and other regulatory requirements.	A GMP document is any written record associated with the manufacture, control and distribution of the API or pharmaceutical product.

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
Raw data	Original records and documentation, retained in the format in which they were originally generated (i.e. paper or electronic), or as a 'true copy'. Raw data must be contemporaneously and accurately recorded by permanent means. In the case of basic electronic equipment which does not store electronic data, or provides only a printed data output (e.g. balance or pH meter), the printout constitutes the raw data. Original records The original record (data) which can be described as the first capture of information, whether recorded on paper or electronically. Raw data is synonymous with source data.	Raw data must: ➢ Be legible and accessible throughout the data lifecycle. ➢ Permit the full reconstruction of the activities resulting in the generation of the data. Information that is originally captured in a dynamic state should be available in that state.
Certified true copy or true copy.	A copy (irrespective of the type of media used) of the original record that has been verified (i.e. by a dated signature or by generation through a validated process) to have the same information, including data that describe the context, content, and structure, as the original.	
Meta data	Metadata is data that describe the attributes of other data, and provide context and meaning. Typically, these are data that describe the	Example: Date/time stamp, User ID, Instrument ID, Audit trails, etc. ➢ Metadata forms an integral part of the original record.

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
	structure, data elements, inter-relationships and other characteristics of data. It also permits data to be attributable to an individual. Metadata are data that provide the contextual information required to understand other data. These include structural and descriptive metadata, which describe the structure, data elements, interrelationships and other characteristics of data. They also permit data to be attributable to an individual.	➢ Without metadata, the data has no meaning. ➢ Structured information that describes, explains, or otherwise makes it easier to retrieve, use or manage data Metadata that are necessary to evaluate the meaning of data should be securely linked to the data and subject to adequate review. For example, in the measurement of weight, the number is meaningless without metadata, such as, ✓ the unit, ✓ milligram, ✓ gram, ✓ kilogram, and so on. Other examples of metadata include ✓ the time or date stamp of an activity, ✓ the operator identification (ID) of the person who performed an activity, the instrument ID used, ✓ processing parameters, ✓ sequence files,

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
		✓ audit trails and ✓ other data required to understand data and reconstruct activities.
Data Integrity	The extent to which all data are complete, consistent and accurate throughout the data lifecycle. The degree to which data are complete, consistent, accurate, trustworthy, reliable and that these characteristics of the data are maintained throughout the data life cycle.	The Data integrity shall be maintained throughout the lifecycle of the document. The data should be collected and maintained in a secure manner, so that they are attributable, legible, contemporaneously recorded, original (or a true copy) and accurate. Assuring data integrity requires appropriate quality and risk management systems, including adherence to sound scientific principles and good documentation practices. The data should comply with ALCOA+ principles.
Data Lifecycle	All phases in the life of the data (including raw data) from initial generation and recording through processing (including transformation or migration), use, data retention, archive / retrieval and destruction. All phases of the process by which data are created, recorded,	The SOPs on data retention shall specify the retention period and the archival and retrieval and destruction procedure. There should be a planned approach to assessing, monitoring and managing the data and the risks to those data, in a manner

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
	processed, reviewed, analysed and reported, transferred, stored and retrieved and monitored, until retirement and disposal.	commensurate with the potential impact on patient safety, product quality and/or the reliability of the decisions made throughout all phases of the data life cycle.
Primary Record	The record which takes primacy in cases where data collected or retained concurrently by more than one method fail to concur.	In situations where the same information is recorded concurrently by more than one system, the data owner department should define which system generates and retains the primary record, in case of discrepancy. The 'primary record' attribute should be defined in the standard operating procedures, and should not be changed on a case by case basis.
Computer system transactions	A system including the input of data, electronic processing and the output of information to be used either for reporting or automatic control. A computer system transaction is a single operation or sequence of operations performed as a single logical 'unit of work'. The operation(s) that make up a transaction are not saved as a	Computer systems should be designed to ensure that the execution of critical operations are recorded contemporaneously by the user and are not combined into a single computer system transaction with other operations. A critical processing step is a parameter that must be within an appropriate limit, range, or distribution to ensure the desired product

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
	permanent record on durable storage until the user commits the transaction through a deliberate act (e.g. 1. Pressing a save button. 2 .The metadata (i.e., user name, date, and time) is not captured in the system audit trail until the user commits the transaction. 3. In Manufacturing Execution Systems (MES), an electronic signature is often required by the system in order for the record to be saved and become permanent)	quality. This is controlled through the parameters in the BMCR/ BPCR
Audit Trail	Secure, computer-generated, time-stamped electronic record that shows reconstruction of events relating to the creation, modification, or deletion of an electronic record. GMP audit trails are metadata that are a record of GMP critical information. The audit trail is a form of metadata containing information associated with actions that relate to the creation, modification or deletion of GxP records.	Chronology: who, what, when, and sometimes why of a record. Audit trails capture: Backdating, Aborting runs, Testing into compliance Deleting, Altering data. An audit trail provides for a secure recording of life cycle details such as ✓ creation, ✓ additions, ✓ deletions or alterations of information in a record,

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
		either paper or electronic, without obscuring or overwriting the original record. An audit trail facilitates the reconstruction of the history of such events relating to the record regardless of its medium, including the **"who, what, when and why"** of the action.
Computerised system user access / system administrator roles	System rights (permitting activities such as data deletion, database amendment or system configuration changes) should not be assigned to individuals Where this is unavoidable in the organisational structure, a similar level of control may be achieved by the use of dual user accounts with different privileges.	System Administrator rights (permitting activities such as data deletion, database amendment or system configuration changes) should not be assigned to individuals with a direct interest in the data (data generation, data review or approval).
Data retention	True copy of the original data that is maintained securely throughout the records retention period. Should include associated metadata.	Raw data (or a true copy thereof) generated in paper format may be retained for example by scanning, provided that there is a process in

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
		place to ensure that the copy is verified to ensure its completeness. Data retention may be classified as archive or backup. Data and document retention arrangements should ensure the protection of records from deliberate or inadvertent alteration or loss.
Archive	Long term, permanent retention of completed data and relevant metadata in its final form for the purposes of reconstruction of the process or activity. Archiving is the process of long-term storage and protection of records from the possibility of deterioration, and being altered or deleted, throughout the required retention period.	Archive records should be locked such that they cannot be altered or deleted without detection and audit trail. The archive arrangements must be designed to permit recovery and readability of the data and metadata throughout the required retention period. Archived records should include the complete data, for example, ✓ paper records, ✓ electronic records including associated metadata such as audit trails and electronic signatures. Within a GLP context, the archived records should be under the control of independent data management personnel throughout the required retention period.

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
Backup	A copy of current (editable) data, metadata and system configuration settings (variable settings which relate to an analytical run) maintained for the purpose of disaster recovery. The copying of live electronic data, at defined intervals, in a secure manner to ensure that the data are available for restoration.	Backup and recovery processes must be validated.
Data Review process	To look over, study or examine / To consider retrospectively, look back again / To examine with the eye to criticism of correction.	Data review shall be done at each instance & by expert/ trained person on relevant activity.
Cross-out	A cross-out indicated a correction has been made. This is accomplished by drawing a single straight ink line through information which has been entered inadvertently or incorrectly.	All cross-outs must be initialled and dated. It is required to mention justification for correction/cross out. Example: The redlined document prepared in word with track changes under review menu.
Label	Label should give the contents of the item, when it was created and who is responsible.	Example: The log book/ Register is issued with the Label to give the tile and brief particulars register like: no of copies, issued by, previous log book number, department.
Logbook	Logbook contains the records of a performance, a list of how variable change over time especially on pieces of equipment.	Log book should be completed and should have chronological number in papers.

Term	Definition / Description	Expectation / guidance (where relevant) / Examples of DI Issues
Overwriting	Overwriting refers to writing over previously recorded information to make a change.	Overwriting is never allowed on any cGMP document.
Date	Date the actual day on which information is entered or printed on a document.	As per procedure "Good Documentation Practices"
Backdating:	Backdating is the practice of going back to a previously completed task that has not been properly initialled and dated and placing the date that the task was completed on the date line, as thorough filling in the date had been done in a timely fashion.	This practice is not allowed in any cGMP document.

Other types of situations which trigger data integrity concerns include:

- Trending data that has no microbial excursions ever
- Physical spot checks of reported data versus the actual data results in the incubator
- Fewer tests are recorded than the number of tests documented to be performed
- Averaging of microbiology data to yield results that are within limits
- Performance of unofficial test injections, e.g., chronographic analysis
- Incomplete records for released tests and products

- Failure to routinely follow protocols.

Data Integrity: General Examples

- Human errors
  - ✓ data entered by mistake
  - ✓ ignorance (not being aware of regulatory requirements or poor training)
  - ✓ Wilfully (falsification or fraud with the intent to deceive)
- Selection of good or passing results to the exclusion or poor or failing results
- Unauthorised changes to data post acquisition
- Errors during transmission from one computer to another
- Changes due to software bugs or malware of which the user is unaware
- Hardware malfunctions
- Technology changes making an older item obsolete – old records may become unreadable or inaccessible

Data Integrity Citations for 24 Companies

Data Integrity Citations	Number of Observations from FDA Warning Letters and EU Statement of Noncompliance for 24 Companies
No adequate controls to prevent unrecorded data changes	11
Missing Data	5
General Failure to maintain complete and accurate records	5
Common Login Information Used by multiple individuals	5
Lack of Critical Raw Data	4
Altered Files	4
System setup to automatically discard negative test results or readings	2
Falsified data	1
Printing or recording critical information on personal computers	1
Lack of validation of computer systems	1
Concerns about the quality of clinical trial data	1
Electronic records do not meet quality of paper records	1

Other Citations

- Overwriting of electronic raw data until acceptable results were achieved
- OOS not initiated
- Falsification of data to support regulatory filings
- Standalone GC systems without adequate controls
- Falsification of batch records (re-writing clean records)
- Non-contemporaneous recording of lab data
- Recording of sample weights on scraps of paper
- Missing raw data
- Unofficial testing of samples (trial samples)
- OOS results not investigated
- Retesting completed but not justified
- No restriction/protection of electronic data
- Chromatographic software was not validated to ensure rewriting,
- deletion of data prohibited
- IPQC performed without batch record present

- Unexplained 'trial' samples run before analysis
- Deletion of HPLC data -lack of data security
- Missing stability samples
- Lack of records demonstrating who performed analysis
- Raw data not recorded contemporaneously nor by the performing analyst
- Failed injections of QC standards (SS) deleted, repeated and inserted into the analytical sequence without explanation.

CONSEQUENCES OF DATA INTEGRITY ISSUES

- ✓ Patient Safety and Lives Can be Lost
- ✓ Warning Letters and Consent Decrees
- ✓ Import Embargos, Recalls and Seizure of Products
- ✓ Need to Hire Third Party Consultants for Data Integrity
- ✓ Application Integrity Policy
- ✓ Loss of Regulatory Trust
- ✓ Criminal Charges Can be Cited and Enforced (Arrests and Court Trials)