



GOOD PRACTICE GUIDE:

Data Integrity – Key Concepts



Documents
Document Sharing Hub

GOOD PRACTICE GUIDE:

Data Integrity – Key Concepts

Disclaimer:

The *ISPE GAMP® RDI Good Practice Guide: Data Integrity – Key Concepts* provides detailed practical guidance to support data integrity within a regulated organization. This Guide is created and solely owned by ISPE. It is not a regulation, standard or regulatory guideline document. ISPE cannot ensure and does not warrant that a system managed in accordance with this Guide will be acceptable to regulatory authorities. Further, this Guide does not replace the need for hiring professional engineers or technicians.

Limitation of Liability

In no event shall ISPE or any of its affiliates, or the officers, directors, employees, members, or agents of each of them, or the authors, be liable for any damages of any kind, including without limitation any special, incidental, indirect, or consequential damages, whether or not advised of the possibility of such damages, and on any theory of liability whatsoever, arising out of or in connection with the use of this information.

© Copyright ISPE 2018. All rights reserved.

All rights reserved. No part of this document may be reproduced or copied in any form or by any means – graphic, electronic, or mechanical, including photocopying, taping, or information storage and retrieval systems – without written permission of ISPE.

All trademarks used are acknowledged.

ISBN 978-1-946964-11-3

Preface

ISPE GAMP® RDI Good Practice Guide: Data Integrity – Key Concepts explores areas presented in *ISPE GAMP® Guide: Records and Data Integrity* in further depth. This key concepts Guide incorporates tools such as Cultural Excellence and critical thinking skills into data integrity practices to aid companies in meeting regulatory requirements and expectations.

Numerous examples of good data integrity practices along with ways to identify risks and detect issues are included to assist organizations in developing or raising their data integrity awareness.

This Guide is positioned under the *ISPE GAMP® Guide: Records and Data Integrity*, and is aligned with *ISPE GAMP® 5: A Risk-Based Approach to Compliant GxP Computerized Systems*.

Acknowledgements

The Guide was produced by a Task Team led by Lorrie Vuolo-Schuessler (GlaxoSmithKline, USA) and Charlie Wakeham (Waters Corporation, Australia). The work was supported by the ISPE GAMP Community of Practice (CoP) and sponsored by Michael Rutherford (Syneos Health, USA).

Core Team

The following individuals took lead roles in the preparation of this Guide:

Sam Andrews	Integrity Solutions Ltd.	United Kingdom
Erika Ballman	Albemarle Corporation	USA
George Bass	GGB Services	USA
Ivan Diamond	Bio Products Laboratory Ltd.	United Kingdom
Robert Dillman	Eli Lilly & Co.	USA
Sophie Zhiyao Ding	Ernst & Young LLP	USA
George Evgrafov	PAREXEL International	Germany
Kira Ford	Eli Lilly & Co.	USA
Tami Frederick	Perrigo Company	USA
Elmar Harringer	CoProCo Ing.-Büro Harringer	Germany
Volker Hattwig	Coconeo Ltd.	Germany
Oliver Herrmann	Q-FINITY Quality Management	Germany
Paul Labas		USA
Heather Longden	Waters Corp	USA
Anthony Margetts	Factorytalk Co., Ltd.	Thailand
Barry McManus	Empowerment Quality Engineering	United Kingdom
Leslie A. Paul, MS	Perrigo Company	USA
Siegfried Schmitt	PAREXEL International	United Kingdom
Markus M. Schröder	Coconeo Ltd.	Germany
Doug Shaw	Azzur Group	USA
Rob Stephenson	Rob Stephenson Consultancy	United Kingdom
Michelle Vuolo	Sanofi	USA

Contributors and Reviewers

The Leads wish to thank the following individuals for their valuable contribution during the preparation of this Guide.

Nuala Calnan, PhD	Biopharm Excel	Ireland
James Canterbury	Ernst and Young	USA
Tom De Rudder	Novartis NTO Aseptics	Belgium
Morten Friis	Epista Life Science	Denmark
Isabel Munoz-Willery	NL42 Consulting Paperless Lab Academy	Spain
Anders Vidstrup	NNIT A/S	Denmark

Subject Matter Expert Input and Review

The Team Leads wish to thank the following for their significant contribution to the document.

Monica Cahilly	Green Mountain Quality Assurance, LLC	USA
----------------	---------------------------------------	-----

Regulatory Input and Review

Particular thanks go to the following for their review and comments on this Guide:

Gaye Camm	Therapeutic Goods Administration (TGA)	Australia
Stephen Grayson	Medicines and Healthcare products Regulatory Agency (MHRA)	United Kingdom
Karl-Heinz Menges	Regierungspräsidium Darmstadt	Germany

Special Thanks

The Leads would like to give particular thanks to Bob McDowall (R D McDowall Ltd., United Kingdom), Mark Newton (HeartlandQA (Eli Lilly – retired), USA), Maximilian Stroebe, PhD (GSK Vaccines, Netherlands), and ISPE Technical Advisor, Sion Wyn (Conformity Ltd., UK) for their efforts during the creation process of this Guide. The Team would also like to thank ISPE for technical writing and editing support by Jeanne Perez (ISPE Guidance Documents Technical Writer/Editor) and production support by Lynda Goldbach (ISPE Guidance Documents Manager).

The Team Leads would like to express their grateful thanks to the many individuals and companies from around the world who reviewed and provided comments during the preparation of this Guide; although they are too numerous to list here, their input is greatly appreciated.

Company affiliations are as of the final draft of the Guide.



600 N. Westshore Blvd., Suite 900, Tampa, Florida 33609 USA
Tel: +1-813-960-2105, Fax: +1-813-264-2816

www.ISPE.org

AVA Documents
Document Sharing Hub

Table of Contents

1	Introduction	9
1.1	Background	9
1.2	Purpose.....	9
1.3	Scope.....	10
1.4	Structure of This Guide	11
2	Data Governance	13
2.1	Data Integrity Culture	13
2.2	Roles and Responsibilities.....	23
2.3	Good Documentation Management Practices	26
2.4	Data Classification	31
2.5	Gap Assessments as Part of a Corporate Data Integrity Program	33
3	Data Life Cycle.....	41
3.1	Data Definitions and Requirements	41
3.2	Data and System Life Cycle Interrelationships	45
4	Risk Management Approaches.....	59
4.1	Focus of Risk Management	59
4.2	Supplier and Third-Party Management	59
4.3	GxP Computerized Systems	65
4.4	System Interfaces	80
4.5	Access Controls	83
5	Critical Thinking	87
5.1	Auditing	87
5.2	Use of Analytics to Detect Data Integrity Issues	96
6	Appendix 1 – Data Integrity Gemba Checklist in the Laboratory	103
7	Appendix 2 – IMPACT Tool Applied to Data Integrity	105
8	Appendix 3 – Corporate Data Integrity Program Case Study	107
8.1	Background.....	107
8.2	Program Objectives	107
8.3	Governance	108
8.4	Program Action Plan	109
8.5	Conclusion	109
9	Appendix 4 – Culture and Continuous Improvement Capability Road Map	111
10	Appendix 5 – Regulatory Definitions of Data Terminology.....	113
11	Appendix 6 – Requirements Planning.....	121
11.1	Introduction	121
11.2	Requirements.....	122
11.3	Requirements Analysis	123

12 Appendix 7 – Requirements Specification and Data Integrity Risks for Interfaces.....	141
12.1 Interface Requirements Specification	141
12.2 Typical Data Integrity Issues Related to Data Interfaces	143
13 Appendix 8 – Example of a Four-Tier Classification System of a Life Science Company	145
14 Appendix 9 – Security Controls	147
14.1 Security Controls.....	147
14.2 Review of Controls.....	148
15 Appendix 10 – Case Study: DBA and Security Controls for an RTSM System in a GCP Environment.....	149
15.1 Background.....	149
15.2 Infrastructure Controls	149
15.3 Account Controls.....	150
15.4 Segregation of Duties	150
15.5 Periodic Reviews	150
15.6 Internal Audit.....	150
16 Appendix 11 – Case Study: DBA and Security Controls for an ERP System in a Medical Device Manufacturing Environment.....	151
17 Appendix 12 – Case Study: Laboratory Computerized System.....	153
17.1 Typical Use Scenario	153
17.2 Records Risk Assessment and Controls Considerations.....	154
17.3 CDS Example	154
17.4 Remediation Plan.....	155
18 Appendix 13 – Case Study: Uncontrolled Spreadsheet	157
18.1 Scenario.....	157
18.2 Records Risk Assessment and Controls Considerations.....	157
18.3 Spreadsheet Example.....	157
18.4 Remediation Plan.....	158
19 Appendix 14 – Case Study: Process Control System	159
19.1 Scenario.....	159
19.2 Records Risk Assessment and Controls Considerations.....	159
19.3 PCS Example.....	160
19.4 Remediation Plan.....	161
20 Appendix 15 – Case Study: Business Application System.....	163
20.1 Scenario.....	163
20.2 Records Risk Assessment and Controls Considerations.....	163
20.3 IT Systems Example	163
20.4 Remediation Plan.....	164
21 Appendix 16 – Reviewing Laboratory Systems.....	165
21.1 General Requirements.....	165
21.2 Access Roster Review	165
21.3 Data and Transfers	166
21.4 Data Processing.....	166
21.5 Laboratory System Audit Trails	166

22 Appendix 17 – Reviewing IT Systems	169
22.1 IT System Overview	169
22.2 User Access	169
22.3 IT Audit Trails	170
22.4 IT System Validation	171
22.5 IT System Data Flow	173
22.6 IT System Data Storage	173
23 Appendix 18 – Reviewing Supporting Data	175
23.1 Time card or Badge-in vs. Data or Batch Approval	175
23.2 Maintenance Records vs. Data in Historian	175
23.3 Batch Records vs. Component or Material Records	175
23.4 Concealing Things in a Parallel System: The Numbers Game	175
23.5 Timing: Determining the Real Sequence of Events	176
24 Appendix 19 – Auditing Access Controls	177
25 Appendix 20 – Regulatory Guidance Regarding Classification of Deficiencies	179
26 Appendix 21 – Detecting Aberrant Results	181
26.1 Detection Methods	181
26.2 Assumptions	182
26.3 Grouping, Normalizing, and Profiling Data	182
27 Appendix 22 – References	185
28 Appendix 23 – Glossary	189
28.1 Abbreviations and Acronyms	189
28.2 Definitions	191

1 Introduction

1.1 Background

The *ISPE GAMP® RDI Good Practice Guide: Data Integrity – Key Concepts* provides detailed practical guidance intended to support data integrity within a regulated organization.

In considering a holistic approach to data integrity, some simple assumptions can help to focus effort:

- Data integrity is usually lost early in the process, typically at data collection or during initial processing.
- Once data integrity is lost, it can never be restored; however, it is important to be able to detect the loss of integrity.
- Enterprise-level systems may provide effective controls, but issues can occur if the data is corrupted before it reaches the system (per first bullet).
- Every human interaction is a potential opportunity for a data integrity issue.
- Every interface is a potential opportunity for a data integrity issue.
- Data must be protected throughout the data life cycle including passing through interfaces between systems and stages.

These data integrity considerations need to be identified and the resulting risks mitigated to remove or reduce as many of the potential threats as possible. It is important to be able to identify when data integrity could be compromised so that methods of detection or prevention can be implemented.

There are four primary elements required to ensure data integrity at all stages:

- Cultural Excellence, embedding behaviors and leadership in support of data integrity.
- Quality Risk Management approach to designing and managing GxP data processes such that risks are assessed and mitigated, and controls are communicated and iteratively reviewed and improved (and encompassing the Data Integrity by Design paradigm).
- Technical and procedural controls are in place to prevent unauthorized changes.
- Evidence to provide assurance of data integrity in support of business processes.

1.2 Purpose

This Guide provides detailed guidance in four core areas:

- Data Governance
- Data Life Cycle
- Risk Management Approaches
- Critical Thinking

These core areas represent four of the six data integrity key concepts introduced in the *ISPE GAMP® Guide: Records and Data Integrity* [1]. Of the remaining two key concepts from that Guide:

- GxP Computerized System Life Cycles is extensively detailed in *ISPE GAMP® 5: A Risk-Based Approach to Compliant GxP Computerized Systems* [2].
- ALCOA and ALCOA+ are fundamental to data integrity approaches, with well-documented guidance contained in the WHO TRS No. 996 Annex 5: Guidance on good data and record management practices [3].

In addition, Section 2.3 of this Guide presents user behaviors supporting ALCOA+.

1.3 Scope

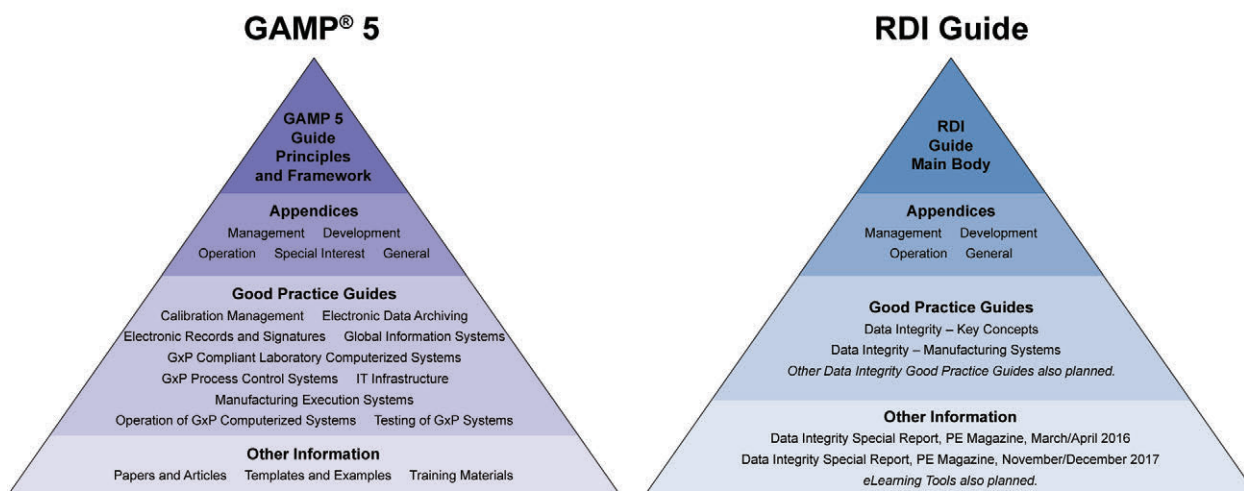
The scope of this Guide is data required under one or more GxP regulations (often referred to as the predicate rules). In preparing this Guide, the following regulatory guidances have been taken into account:

- WHO Technical Report Series No. 996 Annex 5: Guidance on good data and record management practices (2016) [3]
- FDA Data Integrity and Compliance with CGMP – Draft Guidance for Industry (April 2016) [4]
- MHRA 'GXP' Data Integrity Guidance and Definitions (March 2018) [5]
- EMA Questions and Answers: Good Manufacturing Practice – Data Integrity (August 2016) [6]
- PIC/S PI 041-1 (Draft 2) Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments (August 2016) [7]
- China FDA Draft Guidance on Drug Data Management Practice (October 2016) [8]
- Health Canada GUI-0001: Good manufacturing practices guide for drug products (February 2018) [9]
- TGA Data Management and Data Integrity (DMDI) (April 2017) [10]

1.4 Structure of This Guide

This Guide is positioned under the *ISPE GAMP® Guide: Records and Data Integrity* [1] and is intended to provide additional and specific detail on data integrity, in the same way as *ISPE GAMP® 5* [2] has Good Practice Guides that provide additional and specific detail on computerized systems validation.

Figure 1.1: Structure Basis of This Guide



The main body of this Guide is supplemented by appendices containing examples and case studies.

2 Data Governance

2.1 Data Integrity Culture

2.1.1 Introduction

The definition of data governance from MHRA is [5]:

“The arrangements to ensure that data, irrespective of the format in which they are generated, are recorded, processed, retained and used to ensure the record throughout the data lifecycle.”

Data governance encompasses the people, processes, and technology required to achieve consistent, accurate, and effective data handling. The *ISPE GAMP® Guide: Records and Data Integrity* [1] examines elements of the data governance framework including the technical, procedural, and behavioral controls that underpin data integrity. Appendix M3 [1] within that Guide discusses the impact of human factors on data integrity by highlighting the following considerations:

- Understanding and mitigating the impact of corporate and local culture
- Understanding the classification and underlying root cause of incidents (from minor lapses to acts, either intentional or unintentional) that result in material impact to patients and/or products and/or application integrity
- Implementing mechanisms to minimize human error rates
- Reducing motivation, pressures, and opportunities for data falsification and fraud
- Promoting impartiality in quality-related decision making
- Applying effective behavioral controls by influencing behaviors and attitudes

In other words, the *ISPE GAMP® Guide: Records and Data Integrity* [1] highlights key factors that should be considered relative to how an organization's culture can influence data integrity outcomes.

In this Chapter, the behavioral controls are examined in more detail in the context of a cultural excellence approach.

The *ISPE Cultural Excellence Report* (issued in April 2017) [11] aligns with and complements these key concepts, stating that:

“Culture can be described using many different terms, but the key is to define, emphasize, and support the demonstration of desired behaviors and results.”

The *ISPE Cultural Excellence Report* [11] explores the term “cultural excellence,” proposing that within any given organization there is not a separate quality culture, safety culture, data integrity culture, etc. Rather, one primary corporate or organizational culture exists that influences the behaviors and actions giving rise to quality, data integrity, and safety outcomes that matter to the patient and the business.

The *ISPE Cultural Excellence Report* [11] recommends moving away from the traditional “culture of compliance” toward a “culture of excellence,” stating:

“This signals a shift from reliance solely on regulatory compliance to an emphasis on continuous improvement in which there is deep understanding throughout an organization of the elements critical to product quality.”

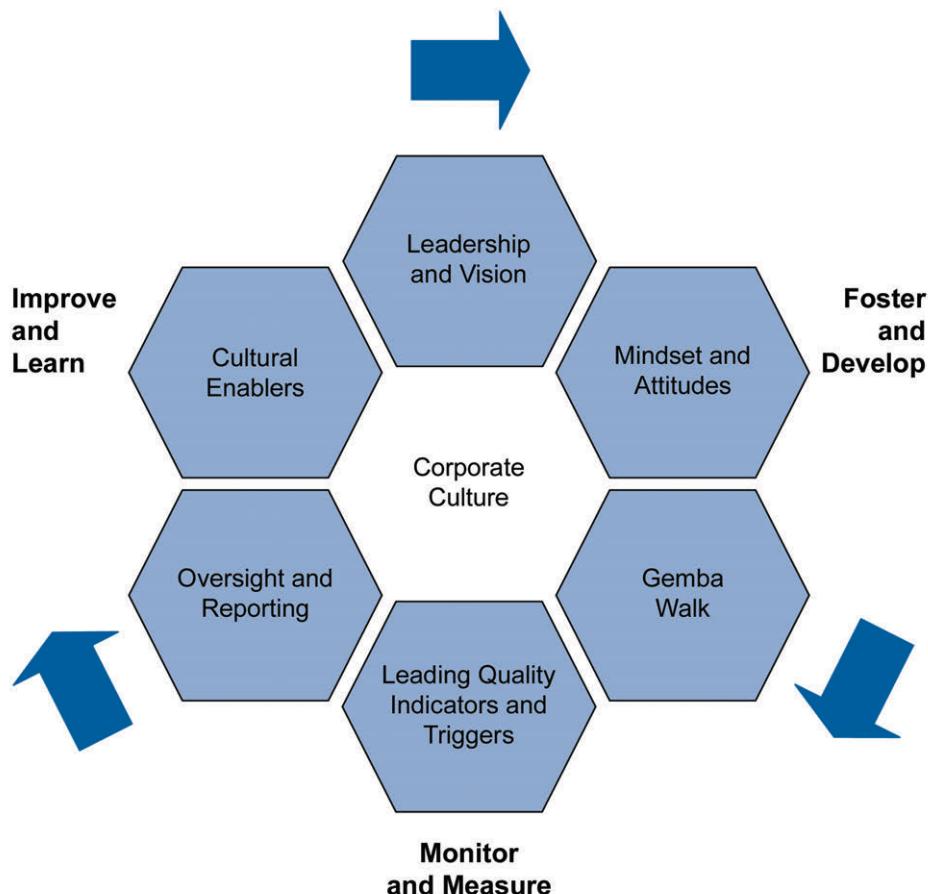
The cultural excellence framework in Figure 2.1 focuses on six dimensions that together foster, develop, monitor, measure, learn, and ultimately improve an organization's culture of excellence [11]. They are:

- Leadership & Vision
- Mindset & Attitudes
- Gemba Walk
- Leading Quality Indicators (LQIs¹) & Triggers
- Management Oversight & Reporting
- Cultural Enablers

All are discussed in more detail later in this Section.

The *ISPE Cultural Excellence Report* [11] outlines a series of practical and powerful approaches, practices, and tools to support implementation of the cultural excellence framework and promote behavioral change. The following Section on data integrity culture demonstrates how some of these approaches, practices, and tools can be applied to ensure the integrity of the data an organization produces.

Figure 2.1: Six Dimensions of Cultural Excellence Framework [11]



¹ LQIs are a variant of quality metrics that include a measure of behavioral attributes, and are discussed in more detail in the *ISPE Cultural Excellence Report* [11].

2.1.2 Leadership and Vision

2.1.2.1 The Importance of Leadership to Data Integrity

Leadership plays a key role in establishing a healthy corporate culture. Attitudes toward safety, quality, and operational excellence can be shaped, both positively and negatively, by leader actions and behaviors. Leaders can also set the desired tone and direction relative to data integrity practices and expectations.

2.1.2.2 Leadership Model for Data Integrity

The first step for success is to ensure that corporate and site leadership are aware of, and fluent with, the recent regulatory findings and concerns around data integrity, and expectations for best practices. Only then are they likely to successfully influence their organization toward needed actions.

A clear data integrity direction or plan enables the entire organization to understand the desired state and the importance of compliance with the requirements, enabling employees to work in alignment with the specified goals and standards.

Leaders must clearly articulate the importance of data integrity as a regulatory standard for patient safety as well as a business success factor. Leaders must share the message broadly and frequently within the organization; this can be accomplished both formally and informally. It is essential, however, that leadership returns often to the message to maintain agreement and reaffirm its importance.

2.1.2.3 Modeling Leader Behavior

Leaders must make decisions consistent with the company quality expectations. As patient safety, product quality, and data used to support regulated decisions are paramount, actions must align with the company's data integrity plan, and data integrity risks and gaps should be viewed as potential patient safety risks.

If data integrity violations arise, all aspects must be considered to make the right decisions for patient safety, product quality, and data used to support regulated decisions. As leader actions are scrutinized, it is vitally important that leaders understand the impact of their words and actions; this will help ensure that the right messages related to the data integrity plan cascade throughout the organization.

With respect to data integrity, leadership values of transparency and humility are important. Being open to understanding the vulnerabilities within any given system and transparency about improvements needed are necessary to improve and develop data integrity cultural maturity.

Often executive-level leaders are not the technical Subject Matter Experts (SMEs) as it relates to data integrity. Site leaders, therefore, must enable those within the technical areas (IT, Laboratories, Quality, Operations) to coach and mentor the SMEs in the desired leader roles and empower them with the tools and resources to execute data integrity action plans effectively.

Leaders at all levels should clearly establish acceptable data integrity practices and show, by their actions, sincere and serious efforts to improve the overall data integrity maturity.

Leaders who employ both formal and informal communication can gain more exposure and reach different organizational levels more effectively. Their behavior helps to shape the thoughts and actions of other employees; therefore, leaders must actively model the principles, values, and vision desired by the company, personalizing the overall data integrity message with their own leadership style to ensure credibility.

Leaders can also promote an environment that is open to change, one in which ideas to improve quality and data integrity compliance are welcome and in which employees are not afraid to voice data integrity concerns. This emphasis on "Speak Up" is key, as employees who discover vulnerabilities should not be fearful to raise and address issues.

Many companies provide anonymous phone lines that allow employees to share confidential concerns of quality, safety, or other topics. Leaders should actively seek to develop Speak-Up cultures (viewed as ideal for enabling cultural excellence) in which employees feel comfortable sharing all concerns, including those that relate to data integrity.

2.1.3 Key Leader Actions

Leaders must consistently promote and seek within their organization information and suggestions to drive improvement.

Additionally, leaders should invest in their own development and evaluate their own effectiveness in order to determine areas in which they can improve as leaders.

They need to vigilantly monitor and display key performance metrics that hold the organization accountable for continuous improvement goals. (See more in Sections 2.1.6 Key Performance Indicators, and 2.1.7 Management Oversight and Review.)

Vigilance requires monitoring the entire organization to assess and reassess the ongoing process of data integrity maturity, not just as a remediation project or initiative.

2.1.3.1 Shaping Data Integrity Mindsets and Attitudes

As discussed in greater detail in the *ISPE Cultural Excellence Report* [11], leaders can influence mindsets and attitudes by:

- Clarifying behavioral expectations for employees
- Monitoring behavioral performance
- Providing positive cultural learning experiences to reinforce desired behavioral expectations
- Assessing cultural performance to address gaps

Cultural excellence is an ongoing commitment by leaders and individuals to model desired behaviors, promote transparent, proactive ownership of quality by all, and hold themselves and others accountable to standards.

2.1.3.2 Behavioral Criteria and Improvement Actions

The *ISPE Cultural Excellence Report* [11] includes a Cultural Excellence Assessment Tool, which is designed to help organizations assess the maturity of 21 desired key behaviors as part of their quality culture maturity program. This tool provides a behavior-based framework to understand, assess, and develop excellence in quality culture within organizations as described below.

Accountability

Employees consistently see quality and compliance as their personal responsibility. This is key to applying self-behavioral controls in the area of data integrity.

Establishing clear individual accountability for data integrity compliance is a foundational step in helping to shape the collective mindset for cultural excellence. Accountability should be communicated consistently through job descriptions, onboarding practices, GxP training, and performance goals, and be supported by coaching, capability development programs, rewards, and recognition. Leaders should hold themselves and others accountable for performing to quality and compliance standards.

This concept of employee accountability does not supersede or replace the need for management leadership and accountability for overall quality.

Ownership

Employees must have sufficient authority to make decisions and feel empowered to do their jobs well.

Individual ownership of quality and compliance is a primary driver for shaping the quality mindset. When individuals are fully engaged, empowered, and taking action to improve product quality, patient safety, and the quality of data used to support regulated decisions, organizations typically benefit from continuous improvement and faster decision making.

Action Orientation

Employees committed to cultural excellence regularly identify issues and intervene to minimize potential negative effects on quality and compliance.

Establishing the expectation that individuals demonstrate a proactive orientation helps shape the quality mindset and foster cultural excellence. Leaders should promote and leverage proactive efforts (e.g., risk assessments, Gemba walks (see Section 2.1.5), employee suggestions) to reinforce support for the desired behavior. Additionally, it is important that rewards and recognition be aligned to support these efforts, rather than reinforce reactive “fire-fighting” efforts.

Speak Up

Employees are not afraid to speak up, identify quality issues, or challenge the status quo for improved quality; they believe management will act on their suggestions.

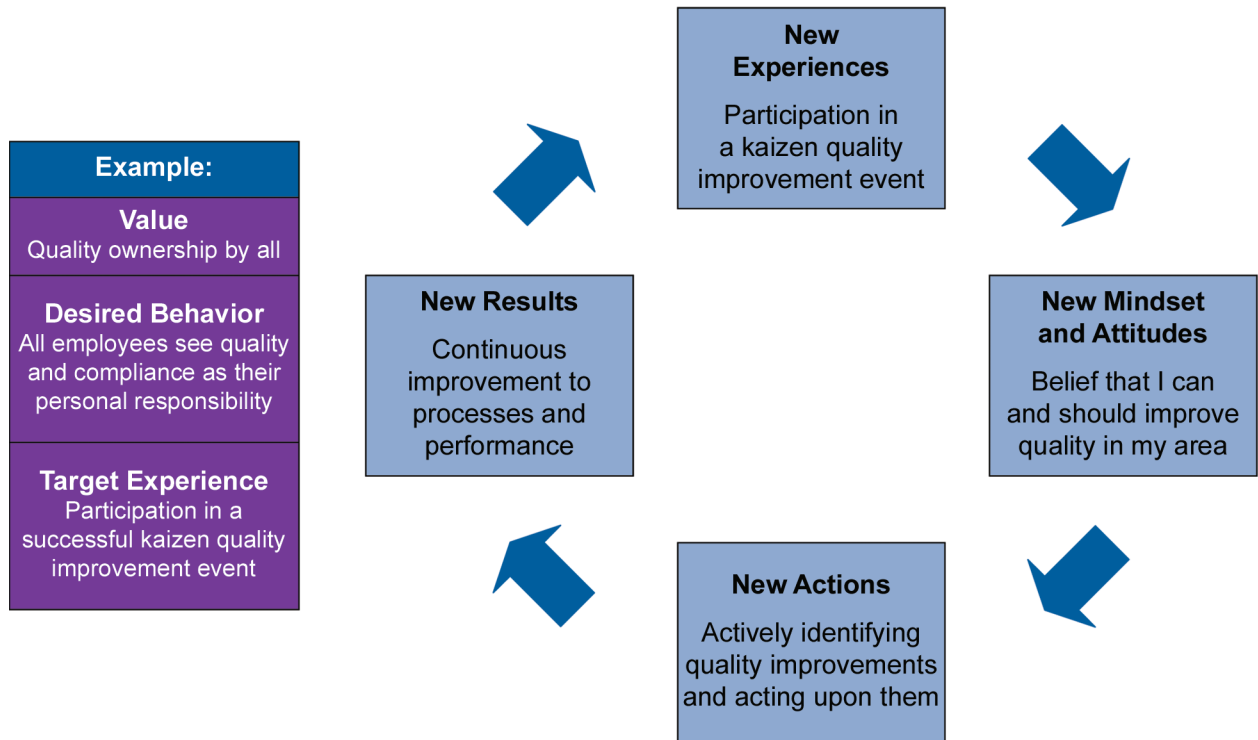
Empowering individuals to speak up and raise quality issues helps foster the quality mindset. Leaders should support this by modeling the desired behavior, building trust, and creating an environment in which individuals feel comfortable raising quality issues, engaging frontline personnel in problem solving, and involving employees in continuous improvement activities.

There are geographical cultures that may inhibit employees from speaking up; the impact of local cultures is discussed in more detail in the *ISPE GAMP® Guide: Records and Data Integrity, Appendix M3* [1].

2.1.4 Mindset and Attitudes

A company's ability to monitor and shape mindset and attitudes can greatly improve results and performance. Figure 2.2 shows an experience-based process in which mindset influences attitudes, which influence behaviors, which directly influence actions, which influence results, which ultimately influence performance. This becomes a continuous improvement loop as new results and performance further form desired mindsets.

Figure 2.2: Shaping Quality Mindset: Experience-based Approach [11]



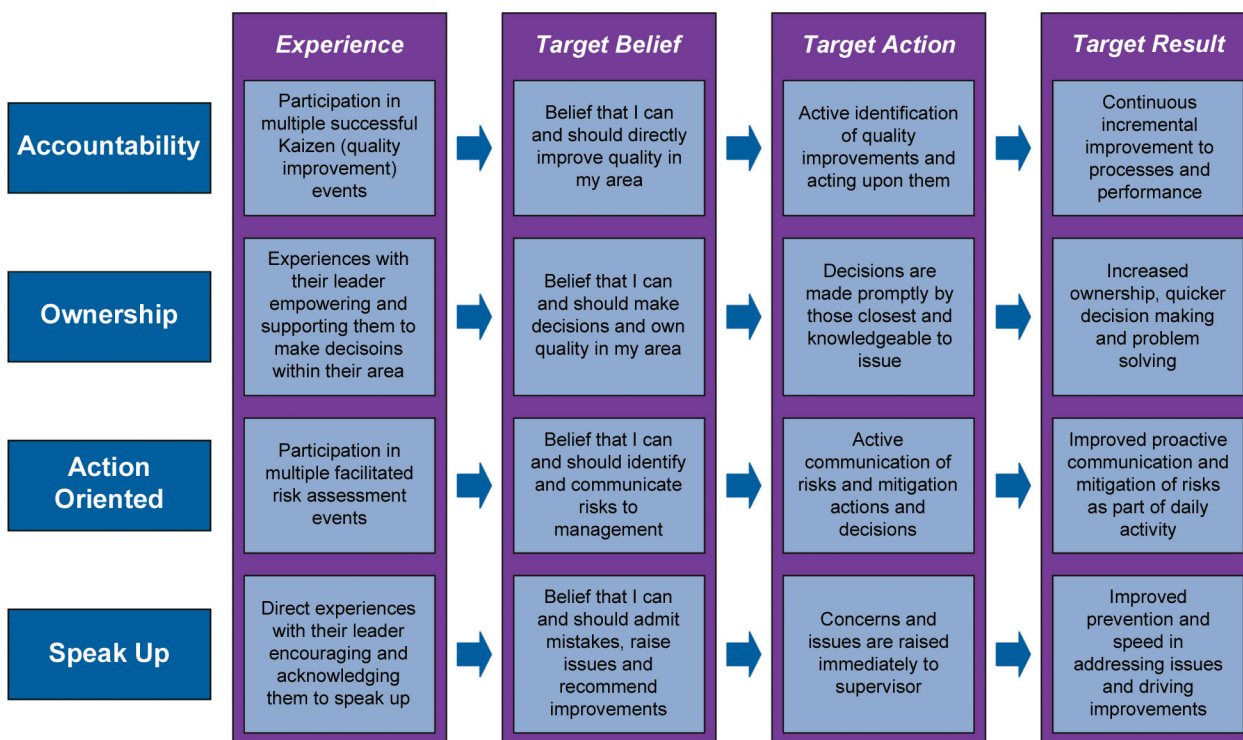
Individual leader actions and behaviors clearly contribute to site and corporate culture. There are commonalities among industry leaders related to behavior, actions, and traits that facilitate greater employee engagement, attainment of site goals, and a corporate culture of excellence.

To more effectively shape the desired mindsets and attitudes leaders should:

- Share a vision about the importance of data integrity frequently and broadly within the organization, including case studies, regulatory observations or citations, and other shared industry knowledge.
- Demonstrate decision making and behaviors that align with the company's stated quality vision, including specific data integrity action items.
- Value operational excellence above a focus on regulatory compliance for compliance sake.
- Shape employee experiences and mindsets through formal and informal quality discussions where site metrics are reviewed, data integrity action plans are revisited, and quality issues are raised.
- Establish Gemba walks as a best practice activity for the shop floor, laboratories, and other functional areas; develop Gemba guidelines or checklists to aid a data integrity walk through.
- Develop key site metrics, implement leading quality metrics, and monitor proactive measurements to drive continuous improvement.
- Provide organization-wide structural enablers to support improvement and inspire an environment of continual learning.
- Challenge the organization to drive for excellence and create a culture in which patients and employees benefit.

Figure 2.3 lists experience-based examples across the key areas of accountability, ownership, action orientation, and Speak Up that can help shape mindset and attitudes, emphasizing targeted, positive experiences that can assist in delivering the desired actions and results.

Figure 2.3: Shaping Quality Mindset: Experience-based Examples [11]



Management awareness of mindset and attitudes, combined with an understanding of their effect on behavior, results, and performance, can increase an organization's capability to improve cultural performance over time.

2.1.5 Gemba Walks

Engaged employees build strong processes and systems that are continuously improved. These provide sustainable growth, reduce costs, increase profitability, and create a happy, collaborative team.

This Section contains examples of cultural enablers to enhance the data integrity and cultural excellence of the system through key performance and key behavioral indicators via the use of a Gemba approach. It also provides a sample road map that may be used for each key role in the system to standardize the plan for process improvement, role-based behaviors, training, and sustainability/control.

Enablement requires the development of Leading Behavioral Indicators (LBIs) at all levels within an organization. Cultural excellence culminates with all six cultural dimensions operating at peak maturity: leadership and vision, mindset and attitudes, Gemba walks, LQIs and triggers, oversight and review, and cultural enablers [11]. Each of these elements demonstrate a direct linkage to data integrity. The following provides an overview of the power of Gemba.

2.1.5.1 Gemba Definition

The Japanese term Gemba means "actual place." Jim Womack, author of *Gemba Walks*, expands this definition to call Gemba the place in an organization "where humans create value." [11]

Gemba is a well-defined element of the Lean process improvement concept and, as such, an accepted operational excellence tool in industries that have adopted Lean principles. The Toyota Production System, which originated the Lean concept, has used Gemba walks for decades. Within the pharmaceutical industry, however, the concept of Gemba has not yet been widely implemented [11].

It is important when establishing a Gemba walk to develop LBIs that align with the desired performance targets. Employees understand that we act upon what we measure.

Gemba walks **are** [11]:

- An enabler for cultural change in management style and philosophy
- A role-modeling opportunity for leaders
- A way to empower personnel
- An enabler for continuous improvement through problem solving with the people who experience them
- An opportunity to find the root cause of issues, spot waste and quality risks, and for leaders to remove obstacles
- A coaching/mentoring opportunity to build and/or enhance capabilities and behaviors, and recognize and reinforce desired behaviors
- An enabler for communication of site priorities/challenges and how the unit's performance contributes to the overall success of the site
- An opportunity to learn from the shop floor, foster quality mindset; encourages informed decision making for leaders
- An opportunity for the operators/technicians to show their pride and excellence in their jobs

Gemba walks **are not** [11]:

- An audit (neither quality/compliance nor environmental health and safety)
- A general complaint or venting session
- A debate to defend individual viewpoints without facts
- A troubleshooting exercise in which participants focus exclusively on areas with (technical) issues

2.1.5.2 Gemba Walk Results

In the pharmaceutical industry there may be complaints that supervisors and management rarely make time to go out onto the shop floor or into the laboratories where they could interact with employees and observe what is really going on.

Gemba walks demonstrate a visible commitment from the leadership to all members of the organization. They allow site leadership and supervisors to spread clear messages using open and honest dialogue, to coach and mentor employees, and to get a good indication of the progress of behavioral change at all levels.

Gemba empowers employees, as their contributions to site results are recognized and their ideas for continuous improvements are heard and acted upon.

As the key purpose of Gemba is to identify continuous improvement opportunities, it is critical to record commitments and agreed actions. One of the easiest ways to do this is to display the agreed actions on visual boards in the area, promoting local ownership of progress and providing a central point for the Gemba to commence [11].

Cultural enablement of the leader, supervisor, and technician level employees can be demonstrated through standardized Gemba walks. Appendix 1 demonstrates an example of a Gemba checklist for the laboratory, focusing on the behaviors and activities in support of data integrity.

2.1.6 Key Performance Indicators

Metrics can serve as a powerful tool in driving behaviors in both positive and negative ways. As discussed in the *ISPE GAMP® Guide: Records and Data Integrity* [1], metrics that encourage fraudulent practices (pressure, opportunity, and rationalization) can create data integrity issues.

However, well-designed metrics help raise awareness of best practices, identify issues, and ultimately improve data integrity within the organization. Within a cultural excellence framework supporting data governance, Key Performance Indicators (KPIs) may be referred to as LQIs or LBIs.

The *ISPE Cultural Excellence Report* [11] leverages the “ABC” model of behavioral science outlined by Leslie Braksick [12], which states that:

- A. Antecedents trigger behavior
- B. Behaviors are followed by consequences
- C. Consequences determine if behaviors will recur

Braksick’s work also included the consequence rule that states that consequences have a four times greater impact on behavior than antecedents, yet organizations often leave consequences largely unmanaged [12]. As such, how management responds to metrics has a significant impact on the behaviors espoused, and therefore, the results achieved.

To create well-designed metrics that will drive the desired behaviors, Braksick presents an IMPACT model for developing behavior-based LQIs or LBIs [12].

The IMPACT model [12] requires the following steps:

- Identify the desired quality-improvement goal
- Establish the appropriate **M**eaasure to deliver the goal
- **P**inpoint the “desired” behavior to deliver the goal
- **A**ctivate the **C**onsequences to motivate the delivery of the goal
- **T**ransfer the knowledge across the organization to sustain the performance improvement

The *ISPE Cultural Excellence Report* [11] presents a tool based on Braksick’s ABC and IMPACT models [12] called the LQI design tool. This report highlights that the strength of the tools not only come from pinpointing behaviors that matter, but also from designing positive consequences to deliver the desired results [11].

See Appendix 2 for an example of the IMPACT Tool [12] applied to data integrity.

2.1.7 Management Oversight and Review

Per ICH Q10 [13]:

“Leadership is essential to establish and maintain a company-wide commitment to quality and for the performance of the pharmaceutical quality system.”

Robust oversight and review, engaging both management and employees, reinforces a strong culture of excellence by demonstrating transparency, fostering trust, and facilitating dialogue. All of this enables learning, brings attention to issues so they can be addressed, and highlights best practices so they can be replicated.

One manner in which leadership can achieve this is by instituting a corporate data integrity program. The *ISPE GAMP® Guide: Records and Data Integrity* [1] highlights the following key implementation considerations for a corporate data integrity program:

- A documented rationale
- Executive sponsorship and governance process
- Management accountability
- Levels of training
- Metrics to measure performance
- Program reporting to communicate progress
- Audit and assessment processes

See Appendix 3 for a case study of one organization’s corporate data integrity program.

Companies that have a healthy or mature quality culture are often excellent learning organizations that value and share knowledge across the organization. This learning characteristic is a guiding principle for effective management oversight and review programs as they:

- Align quality objectives
- Monitor for continuous improvement
- Ensure leadership involvement
- Engage in effective external party oversight and reporting

2.1.8 Cultural Enablers

Finally, enabling cultural excellence requires more than just setting goals and KPIs. As emphasized in the preceding Sections, actions and behaviors must also be aligned across the organization.

To that end, the *ISPE Cultural Excellence Report* [11] includes a powerful road map tool that uses a systems approach to identify possible improvements and the necessary associated behaviors.

The Culture and Continuous Improvement Capability Road Map is a tool designed to perform informal continuous improvement within a work area, to provide technical expertise to support cultural excellence, to assess training needs, to align KPIs to role-based behavioral indicators, and to provide information for activities and controls in a structured roadmap [11].

This road map tool encompasses the following key elements: [11]

- Assuring appropriate mentorship and training for all staff
- Building a roadmap for improvement and sustainability to assure cultural excellence
- Using a Plan-Do-Check-Act-Monitor approach to enhance current systems and programs for each role
- Developing **Role-Based Behavioral Indicators** modeled for sustainable data integrity

Appendix 4 provides an example using this tool as a template.

Ultimately, cultural excellence recognizes quality not as an operational burden or compliance requirement, but as a necessity that allows companies to make decisions that best benefit the patient. The ISPE Cultural Excellence Report [11] is an additional resource available to organizations that seek to define, emphasize, and support the demonstration of desired behaviors as a means to consistently deliver enhanced quality outcomes.

2.2 Roles and Responsibilities

Within any organization it is important that senior management establishes and supports an organization built on quality and integrity. They must define the appropriate roles and responsibilities and assign appropriate personnel.

Data integrity must be driven into every aspect of the organization and form the basis of everything the organization does.

Table 2.1 lists the typical roles and responsibilities in an organization. These may not align with job titles in many organizations because often individuals assume multiple roles within the company. In the table below, the roles should not be limited to a focus on an individual system but rather apply across the data life cycle and the wider business process.

Table 2.1: Typical Roles and Responsibilities

Role	Responsibilities
Senior Management	• Establish and maintain the data integrity assurance culture throughout the organization • Set corporate data integrity policy and standards • Ultimately responsible for quality throughout the organization • Allocate appropriate resources to support and sustain good data integrity management • Provide oversight, risk management, and monitoring for data integrity • Define and communicate personnel responsibilities relevant to data integrity practices • Commit to quality process improvement and modernization
Chief Data Integrity/ Governance Officer	• Representative of Senior Management responsible for defining and implementing the strategic approach to organization-wide data integrity assurance • Responsible and accountable for organizational data integrity initiatives • Depending on the scale and complexity of the organization this may be fulfilled as a full-time role, responsibilities assigned as part of an individual's role, or through a committee

Table 2.1: Typical Roles and Responsibilities (continued)

Role	Responsibilities
Process Owner	- Typically, a senior member of the functional unit using the system - Ultimately responsible for the business process or processes being managed - Controls access to the system, in conjunction with the system owner - Responsible for ensuring that the computerized system and its operation is in compliance and fit for intended use in accordance with applicable Standard Operating Procedures (SOPs) - Responsible for providing adequate resources to support development and operation of the system, including continually trained personnel in operations, oversight, and process documentation, as well as financial resources to support the system - Accountable for the existence of appropriate SOPs for operation, change control, and periodic review, and that these SOPs are followed - Accountable for continuous assessment of the system and responding to reported or uncovered findings, observations, or data integrity risks with appropriate remediation and controls - May also be the System Owner and/or Data Owner, depending on the size and complexity of the organization
System Owner	- Ultimately responsible for the availability, and support and maintenance of a system, and for the security of the data residing on that system - Supports the Process Owner to assure continuous assessment of the system and to respond to reported or uncovered findings, observations, or data integrity risks with appropriate remediation and controls - Controls access to the system, in conjunction with the process owner - Acts on behalf of the users and is typically from IT or Engineering functions - Responsible for the availability, support, and maintenance of a system, and for the security of the data residing on that system - Responsible for ensuring that the computerized system is supported and maintained in accordance with applicable SOPs - Accountable to ensure IT support SOPs exist and are followed, provide for training of support staff, and promote system life cycle management including change management, system upgrades/replacement, and availability of system inventory and configuration management
System Support and Administration Personnel	- Support system operations in alignment and compliance with applicable policies and procedures - Responsible for maintaining system access controls - those administering user ID and passwords should not also have responsibility for, or report through, those responsible for the data within the system (segregation of duties) - Depending on the size and complexity of the system and/or organization this may also be the System Owner
Database Administrator	- Responsible for internal and external controls over databases - Depending on the size and complexity of the system and/or organization this may also be the System Owner Internal Controls - Assures installation qualification procedures include internal security controls e.g., database password maintenance, system and user administration controls - Assures back end database changes and data migrations are made in accordance with established procedures that include traceability to source data and ALCOA+ principles (see <i>ISPE GAMP® Guide: Records and Data Integrity, Section 1.5 Key Concepts</i> [1])

Table 2.1: Typical Roles and Responsibilities (continued)

Role	Responsibilities
Database Administrator (continued)	External Controls - Establishes procedures to assure the inclusion of data governance principles to protect databases from external corruption, including risk management procedures on SQL inserts, and other external data corruption methods - Assures that database maintenance can identify between internal and external threats or changes
Data Steward	- Typically a member of the functional unit using the system - Responsible for tactical coordination and implementation of data usage, management, and security policies as determined by data governance initiatives - Responsible for ongoing data review, data audits, and/or continuous improvement initiatives
Data Owner	- Typically a member of the functional unit using the system - Ultimately responsible for the integrity and compliance of specific data at various stages of the data life cycle in accordance with applicable policies and SOPs - May also be the Process Owner
Quality Function	- Responsible for independent oversight and review to assure integrity of data throughout the data life cycle - Accountable for quality policy and standards - Provides current regulatory agencies and industry guidance to the organization - Approves key documents and records including procedures, standards, specifications, plans, and reports - Reviews critical to quality data, whether paper or electronic, as part of QMS activities including qualification and validation, change management, and batch release - Audits systems, processes, governing documents, and records to assure data integrity
Subject Matter Expert	- Individuals with specific expertise in a particular area or field - Responsible for planning and defining verification strategies, defining acceptance criteria, selection of appropriate test methods, execution of verification tests, and reviewing results - Provide expertise in their area or field to determine, define, and implement data integrity measures that prevent, detect, and respond to data integrity concerns
End User	Responsible for collecting, analyzing, reviewing, reporting, and using data and information in a manner that accurately, truthfully, and completely represents what actually occurred in either paper or electronic format
System Developer	- Provides best practices and identifies technical limitations during all phases of the system life cycle - Responsible for developing the system in accordance with established procedures so that all user requirements are met - Responsibilities vary depending on systems and services provided (e.g., COTS vs SaaS Cloud)

2.3 Good Documentation Management Practices

Good Documentation Practices (GDocP) are the combination of technical, procedural, and behavioral controls needed to ensure that GxP data meets the requirements of ALCOA+ for data integrity, irrespective of the format (paper or electronic) [1].

2.3.1 Technical Controls and Procedural Controls

The regulatory expectations for technical and procedural controls for both paper and electronic records to meet ALCOA+ are comprehensively and clearly discussed in WHO TRS No. 996 Annex 5 Appendix 1 [3], and therefore, are not reproduced here.

Additionally, specific guidance on establishing and defining data integrity requirements for computerized systems is described in the *ISPE GAMP® Guide: Records and Data Integrity* [1] and covered in further detail in Appendix 6 of this Guide.

2.3.2 Behavioral Controls

Table 2.2 lists the behaviors needed to ensure GDocP in addition to the technical and procedural controls documented in the WHO TRS No. 996 Annex 5 Appendix 1 [3]. In isolation, these behaviors are not sufficient to ensure data integrity.

Table 2.2: Behaviors Supporting ALCOA+ Requirements

ALCOA+ Requirement	Behaviors Essential to Support the Requirement
Attributable	Data must be attributable to the person or system generating the data, or performing an activity. Users shall choose passwords that are not obvious and are sufficiently complex that the password cannot be easily guessed by a colleague. Users shall keep their passwords secret from other users, and if they do keep a record of their password it must be in a secure location not accessible to others. Ensuring data is attributable can be complicated when an automated process is started by one shift worker and continues to run autonomously through other shifts under the supervision of other operators. Clear timesheets, or similar records, are needed to document the shift changes, and to maintain a record of which operators made interventions to the process during that extended operation.
Legible	Where a user is required to record a reason for change, for example in an audit trail, the reason recorded should clearly explain why the change was necessary and reference any supporting documentation such as an approved change request.* This will facilitate reconstruction of the event and understanding of the motivation for the change such that a reviewer or auditor can decide if the change was procedurally or scientifically justified. Any errors in recording data shall be conserved in the original documentation, with a clear identification of the error or reason for change; the original data must remain visible; and the correct data must be recorded along with the user name, date, and time. Where entries are handwritten, they should be clear and readable by anyone, and written in indelible ink. *Audit trail requirements are inherent within Attributable, Legible, and Complete. Different regulatory guidances discuss audit trails in each of these requirements.

Table 2.2: Behaviors Supporting ALCOA+ Requirements (continued)

ALCOA+ Requirement	Behaviors Essential to Support the Requirement
Contemporaneous	Users shall record data at the time it is generated and shall use the time shown on a networked computer or official company-maintained clock to provide the date and time if manually recorded (instead of using a personal wristwatch that may not be synchronized to company time). Users must always record the current date and time, irrespective of when the task should have been completed. Where the complexity or limitation of the process physically prevents a user from recording contemporaneously (for example a user is working with biohazardous substances within an isolator), the WHO TRS No. 996 Annex 5 [3] and PIC/S PI 041-1 (Draft 2) [7] make allowance for the use of a scribe as follows: <i>“...the supervisory recording should [must] be contemporaneous with the task being performed and should identify both the person performing the observed task and the person completing the record. The person performing the observed task should [must] countersign the record wherever possible, although it is accepted that this countersigning step will be retrospective. The process for supervisory (scribe) documentation completion should be described in an approved procedure which should also specify the activities to which the process applies.”</i>
Original	Data must be recorded directly into the authorized media. Users must record the first reading/measurement/calculated result, even where that value may be deleterious to the overall pass/fail status of the task. In the event that a user makes an error in completing a paper form or the form sustains damage in some way, the original paper form must be kept and cannot be discarded for a new form. Users must not take unauthorized copies of blank controlled forms with the intent of using them to replace any forms containing mistakes or undesirable results. Consideration should be given to incorporating watermarks and iridescent inks into forms to allow easy detection of unauthorized photocopying, over and above the procedural requirement for controlled issuance of forms. All documents should have a unique identification number, a form design that provides sufficient space for manual data entries, and make clear what data should be recorded and where. Users must take care to preserve the original record, which would be the first location in which the information is recorded. For example, when manually writing down a balance reading the user must write it directly into their laboratory notebook or controlled form; a temporary recording onto a scrap of paper would make the scrap of paper the original record. All unused forms should be accounted for, and either defaced and destroyed, or returned for secure filing, with access to the (electronic) templates strictly controlled. Reviewers need to ensure they review the original records (or true copies thereof) including complete data. For dynamic data, the review should be done on the electronic data. See Appendix 5 for the regulatory definitions of these terms.

Table 2.2: Behaviors Supporting ALCOA+ Requirements (continued)

ALCOA+ Requirement	Behaviors Essential to Support the Requirement
Accurate	Users must rigorously follow written procedures governing their daily tasks and apply their training and skills to ensure that any data they generate is accurate. Care must be taken when manually writing or typing data to avoid transcription errors; self-checking of recorded or calculated data should be encouraged as part of routine working practices. In a laboratory environment this would include ensuring that equipment is calibrated and has the correct range and resolution for the intended measurements, that any reagents are within their expiry date, and that the test method and parameters are appropriate for the sample under test. In a clinical study the data must represent the real facts of the study, including accurate measurement of patient vital signs using calibrated blood pressure monitors, weigh scales, etc., and recording all responses.
Complete	Complete data relies on the user collecting and preserving all of the original data, derived data and results, metadata, and audit trail. Complete data requires the user to collect all records created or modified and not selecting only those records that meet specifications. Where multiple results are generated, the user must ensure that all results are formally addressed and reported. All operator actions must be recorded and maintained. The user must follow a clearly defined procedure to invalidate any results where there is sound scientific justification to do so, and must preserve all data, metadata, and audit trail entries related to the results, invalidated or not. A user must not delete electronic GxP data or destroy GxP paper records that are necessary to preserve the data, content or meaning as required under the predicate rules. Such data can only be deleted at the end of the retention period (excepting data under legal hold), under a controlled procedure.
Consistent	Due care and attention must be given to any task undertaken by the user that generates GxP data. The user may develop a consistent set of key phrases for reasons for change and error corrections to facilitate searching and reviewing. Users reviewing data must first determine “can the data be trusted” before making an approve/reject decision on the data as compared to its governing specification. This may involve examining audit trails and metadata, verifying operational sequences using the time stamps recorded, searching for duplicate or modified data, and checking the original data upstream of the activity (e.g., checking the sample preparation data recorded in the laboratory notebook as part of reviewing a chromatography result, or comparing the Electronic Health Record to the data entered in the eCRF system).

Table 2.2: Behaviors Supporting ALCOA+ Requirements (continued)

ALCOA+ Requirement	Behaviors Essential to Support the Requirement
Enduring	Data must be recorded in a permanent, maintainable form for the retention period. Procedures should confirm that archived data, including relevant metadata, is available and human readable. Any GxP data created electronically must be stored in a location that is backed up. For this reason, it is not recommended to store GxP data on local computer hard drives. Users of controlled forms and worksheets must take care to protect the forms throughout their use and then to return them to the original issuer or archivist for safe archival for the retention period.
Available	A user must ensure that all of the GxP data they have generated, whether recorded on paper or electronically, is filed in a logical and labeled secure location for ease of retrieval. It is good practice that filenames for electronic data are chosen to reflect the contents of the file to facilitate search and retrieval.

2.3.3 Overview of Controls

Table 2.3 outlines the holistic approach, encompassing People, Process, and Technology considerations, essential to support the data integrity key concepts.

Table 2.3: Mapping Data Integrity Key Concepts against the People, Process, and Technology Aspects of the Holistic Approach

Key Concept	People	Process	Technology
Risk Management Approach	Data Integrity risk management and assessment process in place	Data integrity risk assessment documented, approved, and periodically reviewed for both systems and records	System configuration optimized to reduce data integrity risks via automated scheduling of backups, audit trails always on, etc.
Data Governance	Cultural excellence inherent in leadership and vision, with strong employee engagement	System training matrix includes data integrity and computer system validation, and evidence of system stakeholder training records	Granular access controls to limit access to system functionality according to responsibility and competency
Data Life Cycle	Data Owners and Data Stewards trained and in place across the data life cycle for GxP records	System records and data life cycle defined, including Creation, Processing, Review-Use, Retention, Destruction	Data audit trails are linked to GxP data and operational activities

Table 2.3: Mapping Data Integrity Key Concepts against the People, Process, and Technology Aspects of the Holistic Approach (continued)

Key Concept	People	Process	Technology
ALCOA+	Data Integrity policy and ALCOA+ principles incorporated within system stakeholders, supplier support services, contractors job descriptions	System records and data classified including primary record, metadata, master data, system access, security procedures, and evidence of control practices in place	System security includes unique user identification and passwords with expiration
Critical Thinking	Periodic Data Integrity audit process established with quality function, system stakeholders and Process Owner	Data audit trail procedure and routine review practices in place supporting GxP decisions and processing	Data audit trails can be accessed and sorted for review purposes. Technical system logs are linked to system updates, changes, backups
GxP Computerized System Life Cycle	Supplier/service provider assessment and ongoing service level agreements. Computerized system validated as fit for purpose and all system users trained.	System incident management procedure in place including assessment of accident/malpractice Routine use and system administration SOPs in place and followed	System validation status and testing correctly reflects system changes, use, updates, and shows that the system is fit for purpose

2.3.4 Considerations for Moving from Paper to Electronic Records

The guidance documents issued by the regulators [3, 5, 7] promote the use of computerized systems and integrated technical controls to reduce data integrity risks.

This can be perceived as a driver to move from paper to electronic records; however, simply transferring a paper process “as is” to an electronic equivalent (the so-called “paper on glass” approach) may carry some of the paper system weaknesses into the electronic system, and may not achieve all of the efficiency gains and benefits that a computerized system is able to offer.

Key activities in planning the paper to electronic transfer include:

- Start by defining the business process, identifying the critical records generated within this process, and determining the corresponding data life cycle for those records. (See Section 3.2 on Data and System Life Cycle Interrelationships for more detail.)
- Define the scope of paper records that would be replaced with electronic records (in one or more computerized systems).
- Ensure all of the data sources are identified for the existing paper records and that the record routings and approval processes are captured in a data flow diagram (e.g., search for those forgotten spreadsheet calculations). It may be useful to reanalyze the business process to determine if all of the record routing and approvals are required.

- Consider the controls and procedures required for data security and confidentiality. (See Section 2.4 and *ISPE GAMP® Guide: Records and Data Integrity* [1] for more information on data classification.)
- Ensure all interfaces (between computerized systems and also between business processes) are captured.
- Strive to minimize manual processes as they pose the highest potential risk.
- Consider, in a multidisciplinary team, how the process for the paper records could be optimized, such as:
 - Merging similar forms (for example, one for each product or test type) into a single form with embedded logic to offer fields determined by the selected type.
 - Interfacing forms to equipment such as balances to automatically document weights to ensure the first (original) readings are captured, any subsequent readings are audit trailed, and transcription errors are eliminated.
 - Interfacing forms to equipment, reagent inventories and/or raw material inventories to ensure the equipment used is in calibration and properly maintained, and the reagents and raw materials are within shelf life.
 - Simplifying the approval routing and implementing email notification to ensure timely approvals.

Fundamental to this process is ensuring that the team has all of the critical stakeholders represented including: business users, GxP regulatory experts, IT systems experts, and configuration and architecture experts.

Also important to this process is visible management commitment, as well as management support for the time and resources needed to perform this analysis. The time invested in the planning stage to define and optimize the existing business process will yield returns in the efficiency of the electronic process, and the knowledge gained will drive the specification of the requirements for the computerized systems to be implemented.

2.4 Data Classification

2.4.1 Data Classification Definition

Data classification directly affects how organizations understand and manage business processes at the most basic level. It is a fundamental element in data protection and a crucial element in enterprise data management and data governance.

Data classification is the process of assigning a value or rating to data. Data classification includes identifying and rating sensitive databases, tables, or columns, identifying restricted or confidential information in database storage, etc.

2.4.2 Why Data Classifications?

With an increase in big data and the proliferation of data from emerging technologies (such as the Internet of Things and connected devices), current challenges for organizations include:

1. Lack of data inventory and data owners
2. Lack of understanding of data, its value, and the different levels of protection needed
3. Lack of framework that mandates different levels of control implementation based upon a structured approach

Well-defined classification of data within a framework can assist in resolving such challenges.

2.4.3 Data Classification Objectives

Data classification as the first step toward data protection includes objectives such as:

- Classifying data to allow a tiered protection scheme and handling
- Encouraging proper labeling and handling of sensitive data for compliance
- Preventing unauthorized access to sensitive information
- Compliance with legal and regulatory requirements

2.4.4 Data Classification Framework

The overall goal is simple: to create a data classification framework that enables the organization to identify, label, and protect sensitive data in different databases. Such a framework consists of several components, including:

1. Data classification policies – define scope, responsibilities, and other governance requirements
2. Data classification scheme – determines tiers of data and associated levels of protection. Three or four-tier schemes are common, although five tier ones exist, especially in industries that produce intellectual properties.
3. Data labeling guidelines – give instructions for labeling and enables automated protection tools
4. Data handling guidelines – provide specific requirements for different data classes based on classification level
5. Data classification mapping – keeps current mapping that links data types or data sets to classification levels

To identify sensitive structured or semi-structured data in each type of database, data classification exercises can start with scanning or eDiscovery with Data Loss Prevention tools, or a walk through with the business and IT to identify locations that might hold sensitive data. Appendix 8 contains an example of a four-tier classification system.

The concept of classification and detailed information for applying it are comprehensively covered in ISO/IEC 27001 Information Security Management [14]. **Note:** the standard uses the term “information” in place of the term “data” used throughout this Guide.

2.4.5 Creating and Maintaining the Data Classification Framework

There is no “one size fits all” approach to building a data classification framework or system. Decisions should be made based on the type of organization and the overall data protection strategy. At a minimum, a high-level policy specifying the requirements for protecting sensitive data should exist and clearly link to the data classification policy.

Classification guidelines (e.g., labeling and handling guidelines) and compensating controls must be linked to each classification level. Data classification processes should be defined for consistent and repeatable execution.

More importantly, data changes over time, which means its value, sensitivity, and protection needs change as well. Data or database owners must be involved to keep database data classification guidance up to date.

2.5 Gap Assessments as Part of a Corporate Data Integrity Program

The importance of performing data integrity assessments as part of the corporate data integrity program to understand the state of data integrity controls is described in the *ISPE GAMP® Guide: Records and Data Integrity, Appendix M1* [1].

2.5.1 Assessment and Remediation Planning

An approach to a data integrity system assessment and remediation program is presented below. The program may be executed at a local departmental level or may be part of a company-wide assessment and remediation initiative:

- Create an assessment and remediation program team
- Develop program tools and procedures
- Provide training in data integrity and ALCOA+ principles to all involved personnel
- Appoint or confirm existing area/department System Owners and Data Stewards
- Review System Inventory to determine critical GxP systems (and to confirm completeness)
- Perform a high-level risk ranking exercise to establish overall priorities for review
- Perform data integrity risk assessment
- Identify mitigation and control actions
- Develop a detailed risk-based Assessment and Remediation Plan
- Track and review progress
- Perform ongoing monitoring of compliance and data integrity controls against changing requirements and evolving working practices

In terms of timescale for remediation of data integrity risks, it is important to consider the criticality of the business process and the data and the mitigating control strategies currently implemented, and then consider the following:

- EU regulations [15] first required an audit trail in 1991
- US regulations [16] first required an audit trail in 1997
- UK MHRA 2018 guidance [5] requires “demonstrated progress” toward addressing system deficiencies

2.5.1.1 Assessment and Remediation Program Team

As described in the *ISPE GAMP® Guide: Records and Data Integrity, Appendix M1* [1], a data integrity program depends on senior management support and leadership; thus, the remediation program team needs to be formed at that level within the company.

Suggested members minimally include: the Chief Data Officer, a senior Quality representative, a senior business representative, and at least one data integrity SME. The team is responsible for setting strategies and direction for the data integrity initiative, and for providing the drive and resources to enable the assessment and remediation program to succeed.

2.5.1.2 Develop Program Tools and Procedures

A set of standard processes and associated templates is required to ensure a compliant and sustainable outcome from the assessment and remediation program, which will need to be developed by the program team or their delegates. The following procedures are required:

- Management of computerized system inventory
- Records and data integrity analysis
- Risk assessment process in line with quality risk management approaches
- Identification and management of remedial actions and controls

2.5.1.3 Provide Training

The provision for training in good record and data integrity practices to all personnel potentially involved in system assessment and remediation is performed as part of the overall record and data integrity governance initiative as described in the *ISPE GAMP® Guide: Records and Data Integrity* [1].

This training is critical to the success of the program to ensure it is not a checkbox activity but a meaningful evaluation of the computerized systems in light of an understanding of the business process supported by the systems and the necessary controls required.

2.5.1.4 Define Roles and Responsibilities

The program team is responsible for the oversight of the project including determining the priorities.

Lower organizational-level teams are likely to perform the assessment/analysis and remediation tasks. To help facilitate the program in each area or department, it is important to have a local SME in records and data integrity (the Data Steward) who can consult with other Data Stewards and remediation program team members, to ensure a consistent approach to the performance of the analysis and the identification of remedial controls. Each system has a system owner who must work closely with the process owner to ensure the ongoing integrity of data produced in that system.

A full description of Roles and Responsibilities is contained in Section 2.2.

2.5.2 Identify and Prioritize GxP Systems

2.5.2.1 Review System Inventory to Determine Critical GxP Systems

As described in EU and PIC/S Annex 11 §4.3 [17, 18], there should be a system inventory for the site/department that lists all the systems in use and summarizes critical information about each system including its GxP impact:

“An up to date listing of all relevant systems and their GMP functionality (inventory) should be available.

For critical systems an up to date system description detailing the physical and logical arrangements, data flows and interfaces with other systems or processes, any hardware and software pre-requisites, and security measures should be available.” [18]

If the system inventory is incomplete or outdated, there is a risk that one or more systems requiring remediation may be overlooked. Consideration should be given to walking through the business process in each area and following the data through each activity, to ensure that all systems are captured and compiled into an updated, current, and accurate system inventory.

It is important to evaluate systems for their impact on the data integrity of the entire business process. A holistic review or walk through of the business process can capture paper-based records and/or hybrid situations where the level of manual intervention increases the potential for data integrity risk.

Additionally, it will identify areas where there is the misconception that printing the data for retention and archiving is an acceptable alternative to securing and maintaining dynamic data as electronic records, as required by the regulations.

2.5.2.2 Establish Review Priorities

A risk ranking exercise considering factors relevant to the overall record and data integrity within each system is performed. To ensure consistency, conduct the assessment needs according to a set of criteria agreed to by the program team.

The assessment ratings for each system are determined and the systems ranked in priority according to their total scores. Weighting factors may be used to reflect the relative importance of the factors if appropriate.

Factors to consider in this exercise are:

- Business criticality
- Overall system GxP impact
- Impact on patient safety, product quality, or data supporting regulated decisions
- Impact on distributed/commercial product
- Original validation approaches (e.g., were the data integrity requirements adequately considered and verified?)
- Level of known data integrity issues
- Frequency of use of system
- Other considerations as appropriate (e.g., audit findings)

The system ranking resulting from this exercise allow the systems to then be divided into groups, for example, high priority, medium priority, and low priority. These groupings are used to determine the timeline and resource allocation for assessment and remediation within the overall site or company-wide Assessment and Remediation Plan.

2.5.2.3 Perform a Data Integrity Gap Analysis

A record and data integrity analysis is performed to identify systems requiring remediation to address risks in the system or process.

This Section discusses common controls and issues within a system that could lead to increased risk to patient safety, product quality, and data integrity. Identified risks are documented as part of the gap analysis.

An effective way of achieving this in a formal and robust manner is to prepare a standard list of verification checks in a questionnaire format; this could be based on *ISPE GAMP® Guide: Records and Data Integrity, Appendix D1* [1].

The results of these evaluations should be formally documented to verify both the existence and the adequacy of the current controls. Documented evidence that the verification has been performed (for example during the original system validation) should be referenced and evaluated.

For systems where this verification is not available it may be necessary to produce evidence that the system meets the verification check, that is, testing as part of the mitigation activities.

Security and Access Control

Some systems may have a limited number of user accounts (due to application limitations rather than a financial reluctance to purchase user licenses); such a restriction can necessitate the use of generic passwords and shared accounts, preventing actions from being attributed to an individual. Where the installed version of the application may restrict the number of user accounts and access levels, it is expected that where available, users should upgrade their systems to a version that provides the necessary functionality or consider replacing the system.

Another typical shortfall is a limited number of available role profiles resulting in inadequate segregation of duties, such as between user and administrator. Where there are limited personnel available, one person may have to fulfill the administrator role in addition to their routine tasks. While that person may therefore be granted both user and administrator roles, it is important that they always log in with the role appropriate to the task: user role for routine tasks, and administrator role for administration duties [7].

In many cases the administrator account is fixed with all privileges assigned, including the ability to generate data within the system, in contradiction of the principle of least privileges and segregation of duties. There have been numerous regulatory citations of lower access-level users being assigned administrator access [19, 20] resulting in inappropriate granting of delete privileges.

It is important to understand the business process supported by the computerized system and ensure that individuals are granted access only to the functionality and/or data appropriate to their job role. Often, the focus of the roles and responsibilities definition is on write and change rights; however, for research organizations, clinical systems, and systems containing unblinding or Personally Identifiable Information (PII) data, read access needs to be as strictly limited as write access.

Physical security is a regulatory requirement [17, 18] and also an area of potential failure by a company during implementation. The storage media for original data, reported results, metadata, and audit trails must be physically isolated (e.g., server in a locked server room not in the main office area), with additional logical security to prevent remote access to the storage media for the purposes of deleting files at the operating system level.

Standalone systems are particularly vulnerable since there is a temptation to store regulated data locally on the hard drive, which is then vulnerable to accidental or malicious attack. Data stored on an accessible hard drive can be subject to deletion, renaming, and re-use as part of data manipulation.

Backup and Restore

Standalone systems are especially prone to lack of data backup given that this typically requires manual user intervention to connect suitable media and start the backup of the hard drive.

Some applications provide an option not to include the audit trail in the backup or archive copy. Where such an option is present, it should be clear that all GxP-relevant audit trail(s) should to be captured in all backups and archive versions.

Another common failing, created during system implementation, is to store backup files on the original server or on a separate server located in the same server room. In the event of a server room disaster, both the original files and the backup files are lost with an associated loss of regulated data. During system implementation or remediation, the backup scripts should be set up such that the backup files are automatically copied to a remote location [7].

Audit Trails and Logs

Systems may have some or all of the following issues regarding data audit trails:

- GxP-relevant activities not captured
- Merging entries of GxP-relevant changes with general event log entries
- Insufficient information captured in an entry, missing any of: user identification, date and time, a description of the action, and the new and old values
- Overwriting (first in, first out) of audit trail entries
- Inability to read, sort, search, or filter audit trail entries for review
- Entries are not independently generated and secure from alteration

In addition to audit trails on GxP-relevant activities, systems may also be deficient in capturing technical system logs of important system activities, such as:

- Configuration changes
- Backup activities
- User logons

These types of technical system logs need to be reviewed for any indicators of potential data integrity risks as part of a periodic review, data integrity audit, or an investigation.

The WHO guidance [3] gives clear instruction that the content, independence, and review capability of the audit trail is more important than the particular term used to describe it in a software application.

Validation

Some systems originally implemented as supporting a non-GxP process may have been put into use without adequate computerized system validation and data integrity controls.

Failure to validate a system for its intended use leaves potential data integrity risks across the business process, as there is no documented evidence that the technical controls function correctly and that, as a whole, the system is fit for purpose.

Archival

Data is often left to build up in the system without a clear archiving process. Storing files indefinitely in the original system for the retention period brings two main concerns:

- **Backup time and cost:** The time taken to backup the system increases over time with the increase in quantity of records in the system.

Backup storage has a cost associated with it, and backup files become very large as they contain copies of inactive data that may not have been modified or viewed in years. So long as this inactive data continues to reside only in the original system, it must be included in the backup. (The regulators are clear that a temporary backup is not an acceptable archiving solution [4].)

Additionally, there is an increased risk of data loss if the time to complete a single backup operation exceeds 24 hours (due to the sheer quantity of data), since it is then no longer possible to complete a daily backup of data within the system.

- **Original System Performance:** As the quantity of files stored within the original system increases, system performance may decrease. Also, searching for data may take longer or may not work as the quantity of files to be searched increases.

Generic Issues across the Data Life Cycle

The data life cycle is discussed in Section 3.2. Each life cycle stage may present potential record and data integrity issues for systems.

Table 2.4: Generic Issues with Systems

Life Cycle Stage	Possible Issues and Deficiencies
Data Creation (Generation, Capture and transmission)	• Lack of secure, individual access control • Inability to support granular accounts enabling segregation of duties for system user, administrator, QA, etc. • Lack of control over date and time stamp settings and format • Inability to deal with multiple time zones • Inadequate audit trail capability on data creation, modification, or deletion • Manual transcription between systems due to lack of interface capability • Interfaces that lack data encryption and fail-safe mechanisms • Unvalidated data transfer between systems • Inability to preserve content and meaning of records during transmission
Data Processing	• Processing methods not version controlled, and earlier versions not retained
Data Review, Reporting, and Use	• Lack of data review capability (including searching and filtering of audit trails) • Inability to support exception reporting processes • Lack of electronic signature capability • Lack of traceability between original data, processed data, and reports • Lack of flagging or reporting invalid and/or atypical data • Lack of ability to trend or perform analytics
Data Retention and Retrieval	• Lack of backwards/forwards compatibility for records from earlier software versions • Lack of automated archiving ability • Lack of checksums or other integrity checks on archived records • Reliance on a manual backup process
Data Destruction	• Reliance on manual deletion at the end of the retention period • Inability to deal with data inadmissible for destruction due to legal hold requirements

These technical failings can be exacerbated when combined with inadequate procedural controls, such as:

- Failure to apply segregation of duties to user roles
- Procedural definition of printouts to be the original data
- User-defined equations created without version control or verification
- No defined record retention period
- Data disposal process inadequate, not procedurally described, and/or not documented

Paper Records

Paper records are unsuitable for complex data as the dynamic nature of the data is lost in the static format (see Section 3.1 for a more detailed discussion), so relying on a printed record in place of secure electronic archiving is unacceptable for chromatography, spectroscopy, video images, electronic Case Report Form (eCRF) with clinical data, and other dynamic data.

Report data that is recorded by hand or transcribed by typing has a high dependency upon the individuals performing the task, which increases the data integrity risk.

System Support

Service and technical support for operational computerized systems is crucial to maintain efficient and effective operations. Service support is often provided by the external supplier that has developed the system or has been involved with the implementation.

In the case of older or outdated systems, ongoing service support from the supplier can be problematic due to a number of factors including:

- Limited support for previous versions
- Discontinued support
- Loss of knowledge, skills, and experience due to personnel change
- Supplier no longer exists

Obsolescence

At some point in its operational life, the hardware within an older system may become obsolete such that it can no longer be maintained in a functional state or cannot support the installation and operation of the latest software version or operating system. This eliminates upgrade as a mitigation option and introduces additional risks around hardware failure with no possibility of like-for-like replacement.

Personnel

Change of company management and staff can have an impact (positive or negative) on the potential risks to patient safety and product quality within a system that relies upon the right behaviors and procedural practices used to mitigate a lack of safeguards and technical controls.

Evaluation of working techniques and associated levels of competency supporting the process can identify risk factors that can compromise data integrity, such as:

- Inappropriate or overly complex workflows
- Outdated customs and practices
- Lack of records and data integrity or ALCOA+ understanding
- Excessive focus on productivity and yield

3 Data Life Cycle

3.1 Data Definitions and Requirements

3.1.1 Introduction

The objective of this Chapter is to appreciate the variety of terms used to describe regulated data, and to explain the terminology that will be used in this Guide. It is important to understand that there is both commonality and variation among global GxP regulations and regulatory guidance.

The terminology defined here is an essential precursor to the detailed examples around applying the data life cycle contained in Section 3.2.

3.1.2 Data Definitions in Regulations

Table 3.1 summarizes which regulations define and use which data term. A full listing of the comparative regulations and terminology is contained in Appendix 5.

Table 3.1: Data Definitions in Regulations

Regulatory Reference/Term	Raw Data	Complete Data	Source Data	Critical Data
US FDA GLP (21 CFR Part 58.3k) [21]	Yes	No	No	No
OECD GLP (GLP No. 1) [22]	Yes	No	No	No
OECD GLP (GLP No. 17) [23]	Yes ¹	No	No	No
EMA/CHMP/ICH Tripartite GCP ² (Guideline for good clinical practice E6(R2)) [24]	No	No	Yes	Yes, with partial definition ³
US FDA Guidance for Industry (Electronic Source Data for Clinical Investigations) [25]	No	No	Yes	No
US FDA GMP (21 CFR Part 211.194(a)/211.188 ⁴) [26]	No	Yes, with list of requirements	No	No
EU GMP (EudraLEX Volume 4 Chapter 4) [27]	Yes, no definition ⁵	No	No	Yes, no definition ⁶
UK MHRA 'GXP' Data Integrity Guidance and Definitions [5]	Yes	No	Yes, synonymous with raw data	Yes, no definition ⁷

Notes:

1. Also defines Derived Data.
2. Also uses the term Source Documents.
3. Used in context: “The sponsor should develop a monitoring plan that is tailored to the specific human subject protection and data integrity risks of the trial.”; “monitoring of critical data and processes”
4. Uses the term complete information in place of complete data
5. Used in context: “For electronic records regulated users should define which data are to be used as raw data”
6. Used in context: “A system should be in place to indicate special observations and any changes to critical data.”
7. Used in context: “The approach to reviewing specific record content, such as critical data and metadata...”

3.1.3 Original Data

For paper and electronic processes, original records of an activity must be captured and retained for the applicable retention period, with the retained format of the records preserving both the content and meaning.

Within this Guide, and based upon the precedent set in the *ISPE GAMP® Guide: Records and Data Integrity* [1]:

- A regulated record is a collection of regulated data (and any metadata necessary to provide meaning and context) with a specific GxP purpose, content, and meaning, and required by GxP regulations. Records include instructions as well as data and reports.
- The terms “original data” and “original record” will be used interchangeably to describe the first recording of original observations and/or acquired data before any processing.

Original data includes the metadata that describes the attributes of other data and is required to maintain GxP content and meaning [1] and that forms an integral part of the original record [5].

The term “original record” is not specific to any one GxP (see Table 3.1 above for the variety of terms used) and is consistent with the definition of “Original” within ALCOA+, as documented in WHO TRS No. 996 Annex 5 §9.5 [3]:

“Original data include the first or source capture of data or information and all subsequent data required to fully reconstruct the conduct of the GXP activity.”

For electronic records, the concept of first capture creating an original record that is then preserved has less relevance since the electronic record lacks the physical presence of a paper record.

A true copy of an electronic record that preserves content and meaning is identical to the original record, whereas a true copy of a paper record is typically distinguishable from the original while still preserving the content and meaning of the original (e.g., a photocopy of a filled page in a laboratory notebook can be a true copy of the original page, but the laboratory notebook page will always be the original, contained within the notebook and bearing the original handwritten recording of the data).

The requirements for true copies are discussed in Section 3.1.4.1, where the key import is that all of the content and meaning must be preserved in the true copy.

3.1.4 Data Requirements

3.1.4.1 Original Record and True Copies

The original record is the first capture of data into an authorized recording medium (paper, photograph, or electronic records including image or spectral files). Maintaining the original record for the retention period may not always be practical, with records such as:

- Thermal printouts, which degrade over time
- Electronic data where the original record alone is not sufficient; there must also be a backup copy of the electronic data
- Records in systems where large volumes of electronic data are created, the ability to archive true copies of the data into offline storage becomes increasingly important

In such situations, a true copy can be retained in lieu of the original records if the copy is equivalent to the original in that it preserves the GxP content and meaning of the original record. Such a true copy ensures that predicate rule requirements are fully met, and the ability to fulfill regulatory obligations preserved.

The various GxP regulations again use a variety of terms, this time for copies of original records: true copies, certified copies, verified copies, or exact copies as documented in Appendix 5. The regulations also discuss static and dynamic data where the static data can be used and understood “as recorded” directly to make a decision (on patient safety, product quality, or regulatory approval), but dynamic data requires additional processing to derive a meaningful value (reportable result) that can be used for decision making.

Regardless of the terms, the content and meaning of the copy must be the same as the original record, including preserving the ability to interpret and process the record for dynamic data.

Table 3.2 lists a variety of original records with examples of permitted true copies for each. Note: in all cases, the true copies must:

- Have been through either a manual verification (e.g., signing and dating the photocopy) or validation (verifying the consistent creation of a true copy via a controlled and automated process)
- Be secured against unauthorized changes
- Be protected throughout the record retention period

Table 3.2: Examples of True Copies of Original Records

Original Record	Nature	Examples of True Copies
Paper notebook or paper form with handwritten values, observations or information	Static format	Photocopies, microfilm, scanned to secure/protected PDF
Electronic file containing a list of values from a data logger – no further processing required	Static format	Paper printout, conversion to secure/protected PDF
Electronic file containing a list of values from a data logger – statistical analysis and trending required	Dynamic format	Electronic record containing all logged values, with the ability to interpret or export the data for analysis and trending
Electronic file containing 3D or complex data requiring processing or trending	Dynamic format	Electronic record containing all raw data, metadata, and supporting data, and giving the ability to further interpret the data in the same way as with the original record

3.1.4.2 Good Clinical Practice Special Requirements

Good Clinical Practice (GCP) requirements include additional provisos unique to GCP: that the investigator must have access to the source data at all times, and that source data cannot reside only in a storage location that is under sole control of the trial sponsor.

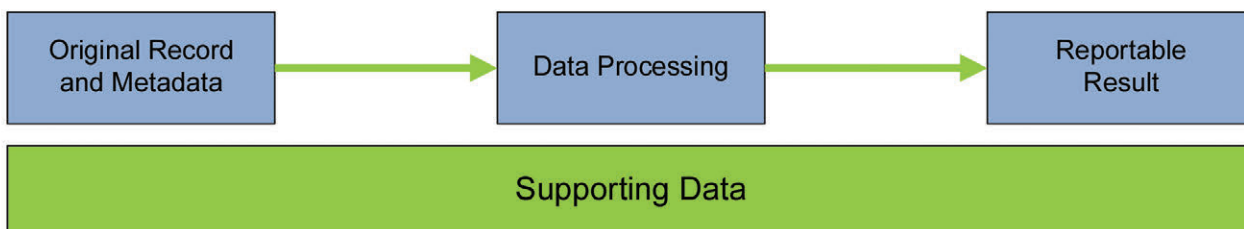
3.1.4.3 Supporting Data

An original record contains metadata, which provides the context and meaning to the data and forms an integral part of the record [5]. When reviewing the original record, the regulatory expectation is that the original data, including metadata, be included in the review when needed to establish the integrity of the data [3, 4]. This is reflected in MHRA ‘GXP’ Data Integrity Guidance and Definitions §6.15 [5]:

“Summary reports of data are often supplied between organisations (contract givers and acceptors). It must be acknowledged that summary reports are limited and critical supporting data and metadata may not be included.”

Figure 3.1 shows how some original records may need to be processed to gain a reportable result.

Figure 3.1: Components of Data



Underneath this process is the supporting data comprising records that could impact the data processing and/or are used to ensure that the system is fit for its intended use and is under control (e.g., effective change control), maintained, and calibrated, as appropriate. Table 3.3 contains examples of reportable results and supporting data for a process control system and a Chromatography Data System (CDS). In a manufacturing environment, supporting data may apply across multiple batches or even be batch independent (see Appendix 18 for examples).

Table 3.3: Original Records, Metadata, Reportable Results, and Supporting Data

System Type	Data			
	Original Record	Metadata	Reportable Result*	Supporting Data
Process Control System	- Process values from sensors - Operator-entered parameters	- Audit trails - Date/time stamp - User ID - System ID	- Critical Process - Parameters (feeding into batch record)	- Sensor location - Warnings and alarms - Sensor calibration data - Cleaning records - Alarm settings - Batch run paper checklist
Chromatography Data System	- Acquired channels from detectors - Operator-entered sample information and method selection	- Audit trails - Date/time stamp - User ID - Method ID - Instrument ID	Critical Quality Attributes (amount, purity, impurities, feeding into Certificate of Analysis)	- Instrument method - Acquisition method - Processing method - Method history - System configuration settings - Calibration and qualification data - Method validation data - Invalidated out of specification results - Laboratory notebooks
*Referred to as Derived Data in OECD GLP No. 17 [23].				

3.1.5 GxP Impact/Data Criticality

Throughout this Guide, the GxP impact of a record or data is a measure of the extent to which it can affect product quality, patient safety, or a regulatory decision. Data criticality expands on this to combine the importance of the decision to be made based on the data [5, 7]. The risk to the data is related to its vulnerability to unauthorized manipulation and the ability to detect any such manipulation. The rigor of the controls established around those risks should be commensurate with the criticality of the data.

Whether data is recorded electronically or manually, it is important that controls are established to prevent and detect manipulation. The use of validated exception reporting processes [5] are fundamental to such detection without adopting a forensic approach to routine data review (which is clearly stated as not a regulatory expectation [5]). Identifying, understanding, and preserving data with GxP impact is further discussed in Section 3.2 Data and System Life Cycle Interrelationships.

3.1.6 Data within the Data Life Cycle

The data life cycle presented within the *ISPE GAMP® Guide: Records and Data Integrity* [1] consists of five phases:

1. Creation
2. Processing
3. Review, Reporting, and Use
4. Retention and Retrieval
5. Destruction

If one looks at a complex business process, including manual and computerized processes, some of these phases may become more complex:

- During the data creation phase, data can be collected electronically from a computerized system or via data recording from a manual process.
- The approach to data review, reporting, and use must be risk-based and consider all relevant supporting data including original paper and electronic records or true copies thereof.
- During the data retention and retrieval phase there may be a need to transfer or migrate data from one system to another, or from one format to another. It is important that any such data movement be performed so as to preserve GxP content and meaning.

For further information around these topics, see the following in this Guide:

- Chapter 2 Data Governance
- Section 3.2 Data and System Life Cycle Interrelationships
- Section 4.3.6 Preserving Existing Data

3.2 Data and System Life Cycle Interrelationships

3.2.1 Introduction

Regulated companies using computerized systems to create, process or store data with GxP impact have three life cycles to consider: the data life cycle, the records life cycle, and the computerized system life cycle.

The scope of the records life cycle extends beyond data with GxP impact and will, therefore, not be covered within this Guide; it is discussed in the *ISPE GAMP® Guide: Records and Data Integrity, Appendix M7* [1].

The data life cycle and the computerized system life cycle form two of the key concepts in the *ISPE GAMP® Guide: Records and Data Integrity, Section 1.5* [1].

This Section will build on these two life cycles, and then look at the relationships and synergies between them.

The objective of this Section is to explain the relationships between the life cycles, to allow the reader to understand how to map the life cycles in their organization to identify their data integrity risks, and then use that knowledge to drive the definition of system requirements and the scope of verification activities.

Understanding and leveraging the life cycles provides a systematic process for the assessment, control, communication, and review of risks to data integrity. This should be a continuous process throughout the computerized system life cycle from concept to retirement, and ongoing for those GxP records within their retention period.

This approach should be applied to any business process creating, processing, using, or storing GxP-regulated data.

3.2.1.1 Data Life Cycle

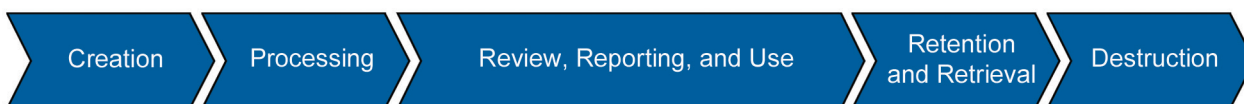
The *ISPE GAMP® Guide: Records and Data Integrity* [1] introduces the data life cycle as follows:

“All phases in data life cycle from initial data creation, capture, and recording through processing (including transformation or migration), review, reporting, retention, retrieval, and destruction should be controlled and managed in order to ensure accurate, reliable, and compliant records and data.”

Data, in this context, is data that directly supports a GxP process or decision; hence, the data life cycle is managed at the business process level.

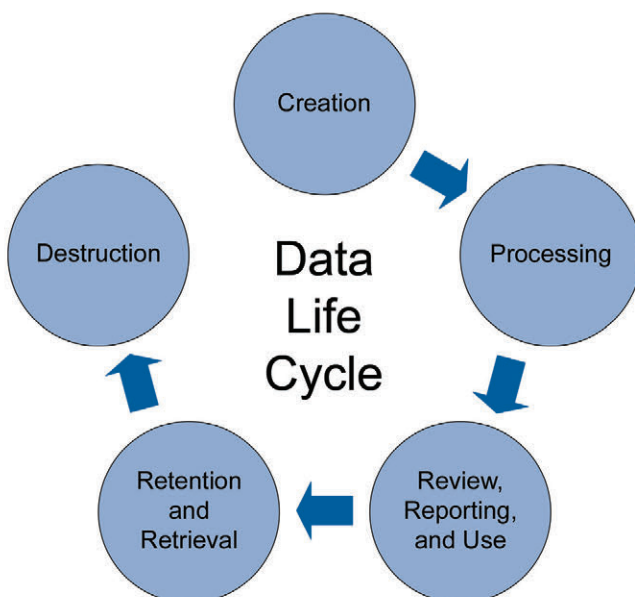
There are typically two representations of the data life cycle: a linear format, which reflects the similarity to the records life cycle, and a cyclic portrayal.

Figure 3.2: Data Life Cycle in Linear Format [1]



The cyclic representation is the main one used in the *ISPE GAMP® Guide: Records and Data Integrity* [1].

Figure 3.3: Data Life Cycle in Cyclic Format [1]



Irrespective of the diagrammatic representation used, the key points in the life cycle are [1]:

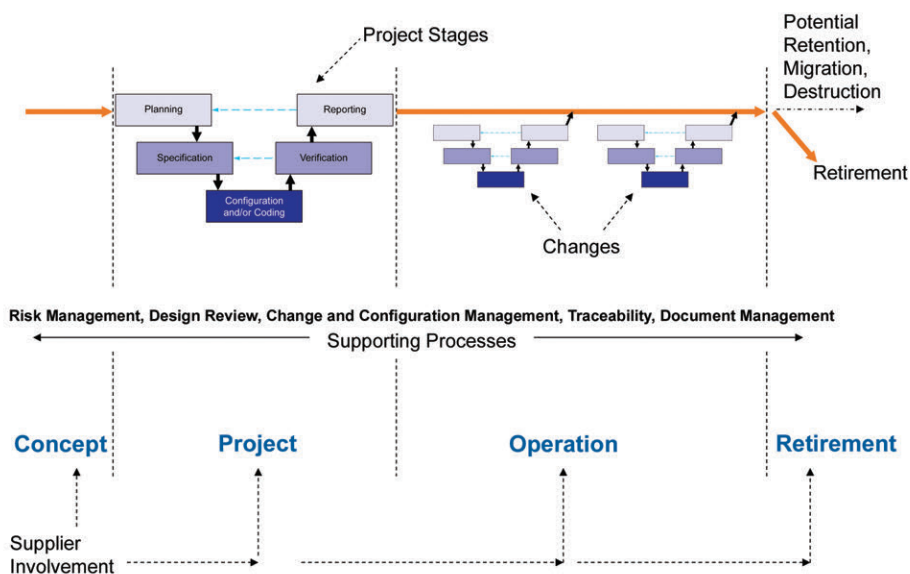
- **Creation:** Data capture or recording should ensure that data of appropriate accuracy, completeness, content, and meaning is collected and retained for its intended use.
- **Processing:** Data is processed to obtain and present information in the required format. Processing should occur in accordance with defined and verified processes (e.g., specified and tested calculations and algorithms), and approved procedures.
- **Review, Reporting, and Use:** Data is used for informed decision making. Data review, reporting, and use should be performed in accordance with defined and verified processes and approved procedures. Data review and reporting is typically concerned with record/report type documents. Second person reviews should focus on the overall process from data creation to calculation of reportable results. Such reviews may cross system boundaries and include the associated external records and may include verification of any calculations used. The data reporting procedures should contain the complete data set and define the data handling procedures, and ensure the consistency and integrity of the results.
- **Retention and Retrieval:** Data should be retained securely. Data should be readily available through the defined retention period in accordance with defined and verified processes and approved procedures. Retention periods vary by record type and applicable regulation, and some records e.g., validation and qualification records, need to be retained for the life of the system or process.
- **Destruction:** The data destruction phase involves ensuring that the correct original data is disposed of after the required retention period in accordance with a defined process and approved procedures. Data should only be retained beyond its retention period if there are legal requirements.

A more detailed discussion of the data life cycle is contained in the *ISPE GAMP® Guide: Records and Data Integrity, Section 4* for Data Life Cycle and *Appendix M4* of the same Guide for Audit Trail Review [1].

3.2.1.2 Computerized System Life Cycle

Computerized system life cycles have been thoroughly documented in *ISPE GAMP® 5* [2], and therefore this Section only contains a brief outline.

Figure 3.4: Project Stages and Supporting Processes within the {Computerized System) Life Cycle (according to ISPE GAMP® 5) [2]

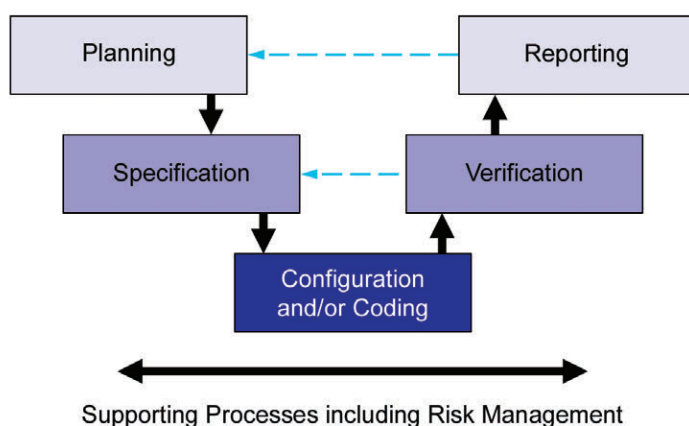


A computerized system life cycle provides a systematic approach through four phases [2]:

- **Concept:** During the concept phase, the regulated company considers opportunities to automate one or more business processes based upon business need and benefits. Typically, at this phase, initial requirements will be developed and potential solutions considered. From an initial understanding of scope, costs, and benefits, a decision is made on whether to proceed to the project phase. The concept phase may be initiated as part of a remediation plan for an existing system. (See Section 4.3 for further details.)
- **Project:** The project phase involves planning, supplier assessment and selection, various levels of specification development, configuration (or coding for custom applications), and verification leading to acceptance and release for operation. Risk Management is applied to identify risks and to remove or reduce them to an acceptable level.
- **Operation:** System operation, typically, is the longest phase and is managed by the use of defined, up to date, operational procedures applied by personnel who have appropriate training, education, and experience. Maintaining control (including security), fitness for intended use, and compliance are key aspects. The management of changes of different impact, scope, and complexity is an important activity during this phase. Another aspect is the completion of a periodic review to determine if the system is still fit for purpose.
- **Retirement:** The final phase is the ultimate retirement of the system. It involves decisions about data retention, migration, or destruction, and the management of these processes. Retiring the system has major implications about how the data created in the system during its operational life remains available, enduring, and readable throughout the data retention period.

This Chapter focuses primarily on the GAMP Specification and Verification approach in the project phase of the computerized system life cycle as an important mitigation of risks related to data integrity within the data life cycle. Applying this approach is comprehensively discussed in *ISPE GAMP® 5* and is not duplicated here [2].

Figure 3.5: A General Approach for Achieving Compliance and Fitness for Intended Use (Specification and Verification Approach from ISPE GAMP® 5) [2]



3.2.2 Business Process Definition

The starting point for all activities within a regulated company must be to define and understand the business process. A business process is a group of related activities or tasks. A data flow diagram shows how data flows through a system or process.