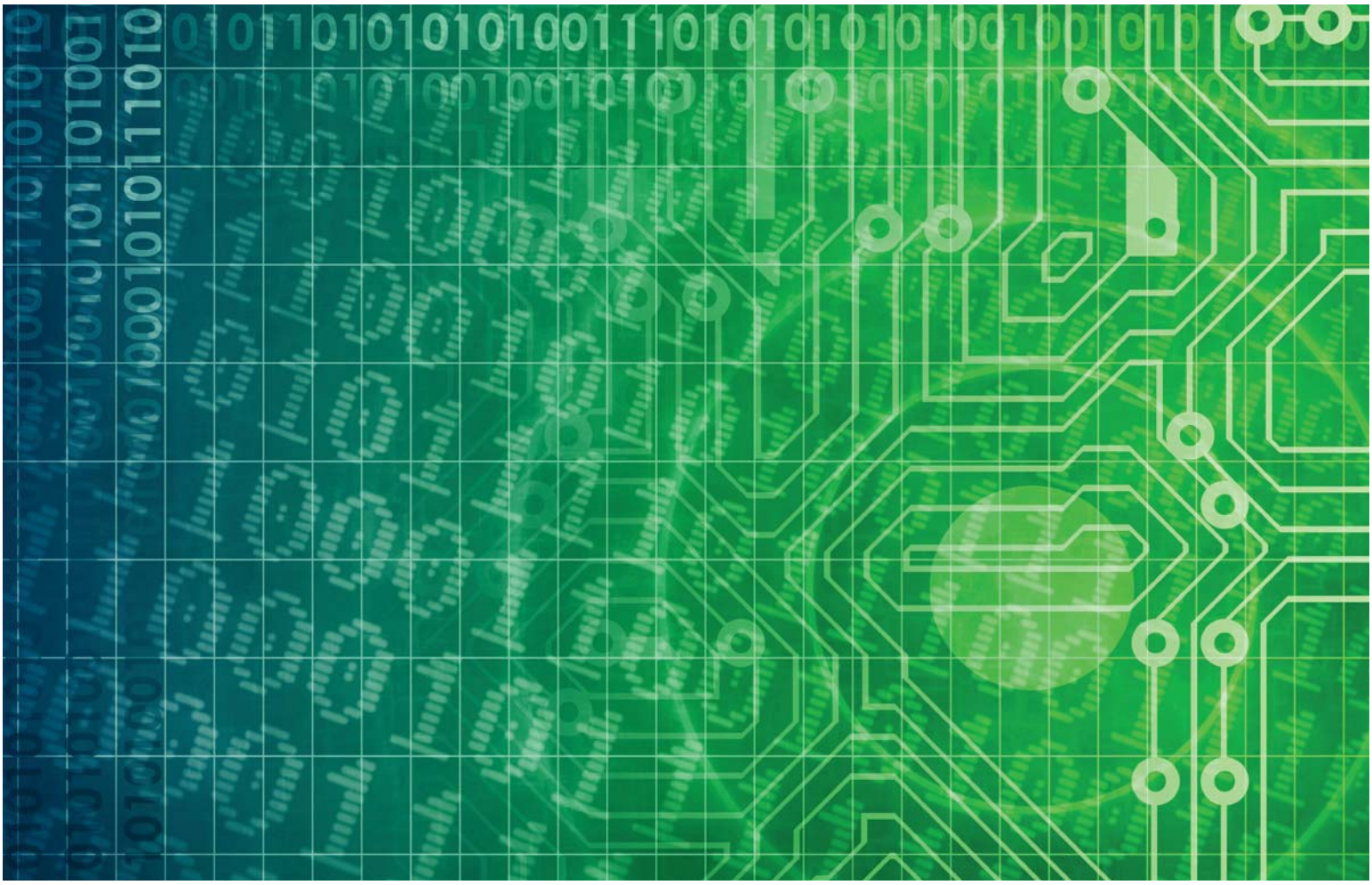
CPA CHARTERED PROFESSIONAL ACCOUNTANTS CANADA

A Framework for Information Integrity Controls

Principal Author: Dr. J. Efrim Boritz, FCPA, FCA, CISA
Co-author: Malik Datardina, MAcc, CPA, CA, CISA



DISCLAIMER

This paper was prepared by the Chartered Professional Accountants of Canada (CPA Canada) as non-authoritative guidance.

CPA Canada and the authors do not accept any responsibility or liability that might occur directly or indirectly as a consequence of the use, application or reliance on this material.

Foreword

With the exponential growth in the importance of data analytics and emerging technologies such as machine learning that depend on data, organizations are looking for simple and applicable guidance on risks and controls that will help them ensure the information they rely on or disseminate possesses integrity. This guidance can help organizations assess and document their information-integrity controls and assist with compliance and audit engagements.

Information quality and integrity are ultimate objectives of many control frameworks, but these frameworks do not provide a clear and direct link between information integrity and the processes, enablers and controls required to ensure such integrity. The absence of such a link may lead to inadequate balancing of implementation and the audit effort directed at indirect (general) controls and direct (application) controls by managers and auditors. Also, it may lead to measures insufficient to achieve the required level of comfort about information integrity. In addition, organizations' sufficiency of data integrity controls often varies across datasets with regulatory and financial data being subject to more robust controls, as compared to operational data.

This publication aims at providing managers, auditors, compliance specialists, financial administrators and information management professionals with a framework for thinking about and identifying key business risks related to information integrity and on designing and implementing appropriate enablers and controls focused on the objective of information integrity. The framework can also assist in planning procedures aimed at assessing the suitability of the design and the operating effectiveness of controls aimed at achieving information integrity. Other companion publications expanding on this one provide more detailed analyses of information integrity risks, specific enablers and controls that can be used to respond to those risks and assurance services that can be used to assess the suitability of the design and operating effectiveness of those enablers and controls.

CPA Canada expresses its appreciation to the principal author of the study Dr. J. Efrim Boritz, FCPA, FCA, CISA and to Malik Datardina, MAcc, CPA, CA, CISA, co-author of the study. Thanks are also expressed to the advisory group for the time and effort involved in undertaking the research study. Thanks also go to Andrée Lavigne, CPA, CA who directed the project.

Advisory Group

Principal Author and Chair

Dr. J. Efrim Boritz FCPA, FCA, CISA
University of Waterloo
Waterloo, Ontario

Co-author

Malik Datardina, MAcc, CPA, CA, CISA
Auvenir
Toronto, Ontario

Chris Anderson, CA (NZ), CISA, CMC, CISSP
Toronto, Ontario

Usuff Curim, FCCA, CISA, CPA
PricewaterhouseCoopers
Toronto, Ontario

Ray Henrickson, CPA, CA, CISA
Scotiabank (Retired)
Toronto, Ontario

Darren James ACA, CISA
Deloitte
Toronto, Ontario

Richard Livesley
BMO Financial Group (Retired)
Toronto, Ontario

Madhavan Nayar
Infogix, Inc.
Naperville, Illinois, USA

Sheryl A. Teed, FCPA, FCA, CISA, CFE
Ernst & Young LLP (Retired)
Toronto, Ontario

Project Director

Andrée Lavigne, CPA, CA
Former CPA Canada (Montréal) staff

Table of Contents

A Framework for Information Integrity Controls

Objective of a Framework for Information Integrity Controls

The value of information comes from its relevance, usefulness / usability and integrity, which can be assessed against the purpose for which the information is produced. Although relevance and usefulness / usability are important contributors to the value of information, the focus of this publication is on information integrity. Information integrity is essential for effective planning, decision-making, monitoring, and control. Senior executives' accountability for the integrity of entity information and internal controls is now well understood in the business community and public sector. While concerns about the risks of information integrity have sometimes been limited to the financial reporting area, in actuality they pertain to all information obtained, created, stored, used, and distributed by businesses and other entities. Such concerns should lead entities to monitor their operations and ensure compliance with relevant laws, regulations and standards related to information integrity.

The increased attention being given to data analytics—both in the "big data" and "small data" varieties—and other emerging technologies, such as machine learning, that depend on data reflects the desire of organizations to extract value from data. However, to draw meaningful insights from data, management must ensure the underlying data has integrity. Poor data integrity costs the economy billions of dollars annually, detracts from the trust business leaders place in the information they rely on to make decisions and creates uncertainty on the part of users about the accuracy of their data. Thus, IBM includes "veracity" as one of the "4 Vs"[1] used to describe big data to highlight the importance of the link between information integrity and the effective use of analytics to extract "actionable insights" from information.[2]

The purpose of this publication is to define information integrity and to provide a context for users and preparers of information who need to understand how information integrity can be achieved and maintained. There is an emphasis on the risks that can lead to impairment of information integrity and the countermeasures to those risks provided by enablers of information integrity and information integrity controls. Those involved in providing assurance on information integrity will also naturally benefit from this publication.

1 The other three Vs include: Volume, Variety and Velocity.

2 https://www.ibmbigdatahub.com/infographic/four-vs-big-data

Addressing information integrity impairment risks in an organized and rigorous manner requires a comprehensive framework both to guide management's risk assessments and its selection of information-system features and internal controls to address the identified risks. This framework will, in turn, guide assurance providers through the criteria they need to consider when providing information-integrity-oriented assurance services. This publication provides such a framework. It is organized around the following key elements:

- information and the information lifecycle
- information integrity characteristics
- domains of information processing
- risks by domain and lifecycle phase
- enablers and controls by domain and lifecycle phase

Other related publications arising from this one provide more detailed analyses of:

- information integrity risks
- specific enablers and controls that can be used to respond to those risks
- assurance services that can be used to assess the suitability of the design and operating effectiveness of those enablers and controls

Information and the Information Processing Lifecycle

Information is created from content (i.e., raw data) through the use of processes within an information system (IS) environment that gather and transform content into information that can be used for planning, decision-making, monitoring and control. Content can range from various types of sensory data to semi-processed structured and unstructured information, metadata, and parameters used to produce information. At a high level, the information lifecycle consists of several key phases:

- creation
- operation
- use
- change
- retirement

Gathering and transforming data into information involves:

1. defining the data to be collected
2. collecting the data based on the definition
3. recording the data in a repository such as a file or database
4. transforming the data into information for use in:
    a. planning
    b. decision-making
    c. monitoring
    d. control

The phases and sub-phases of the information and information processing lifecycle are subject to risks that need to be managed to ensure the information has integrity.

Information Integrity = Representational Faithfulness

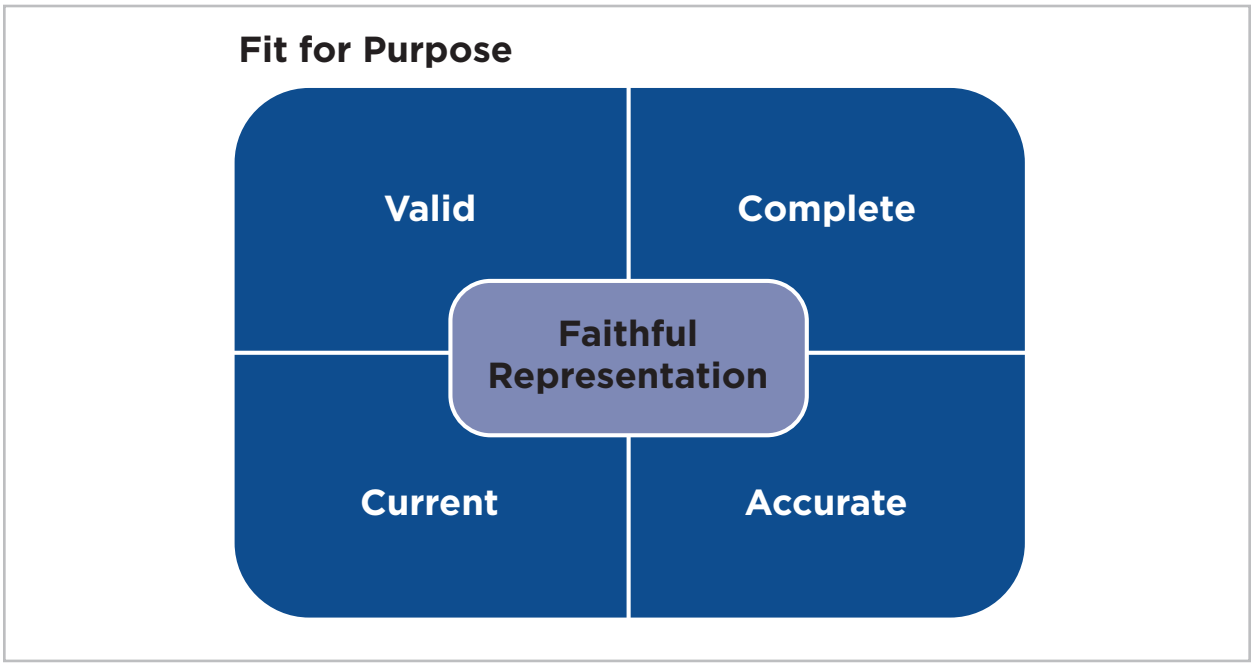
Information integrity is the consistency of the information (i.e., representational fidelity) with the subject matter it purports to portray or represent. For example, during the financial crisis of 2008, the sub-prime bonds previously rated as "AAA" had to be "downgraded 16 notches, all the way to B"—suggesting that the original rating lacked representational faithfulness.[3] The AAA rating was inconsistent (i.e., lacked representational fidelity) with the actual financial risk borne by the bonds at the time of the rating.

Representational faithfulness and its underlying elements

Representational faithfulness can be described by many attributes. However, as portrayed in Figure 1, we believe that the core attributes of representational faithfulness are validity, completeness, currency and accuracy. By "core" we mean that these attributes are the minimum criteria by which the degree of information integrity can be assessed. This assessment must be made while keeping in mind the purpose for which the information is intended. The core attributes can be described as follows:

- Validity: The information portrays what it purports to portray.
- Completeness: The information is complete over time and across items.
- Currency: The information is the most up-to-date version.
- Accuracy: The information is free of error and sufficiently precise for its intended purpose.

FIGURE 1: REPRESENTATIONAL FAITHFULNESS AND ITS UNDERLYING ELEMENTS



3 Lowenstein, Roger. "Triple-A Failure," *New York Times* (www.nytimes.com/2008/04/27/magazine/27Credit-t.html?pagewanted=all&_r=0, April 27, 2008)

The required degree of information integrity depends on the information's intended use. Uses related to health and safety (e.g., information on the effectiveness of a proposed drug) may require very high levels of representational faithfulness, whereas uses related to entertainment (e.g., movie ratings) may require lower levels.

Meta-information

Meta-information is information about information. Meta-information can provide users of information with the context of that information to help reduce the risk that the information will be used for an unintended purpose or used inappropriately for an intended purpose. Thus, information must be accompanied by meta-information or linked to meta-information that describes the information, including its:

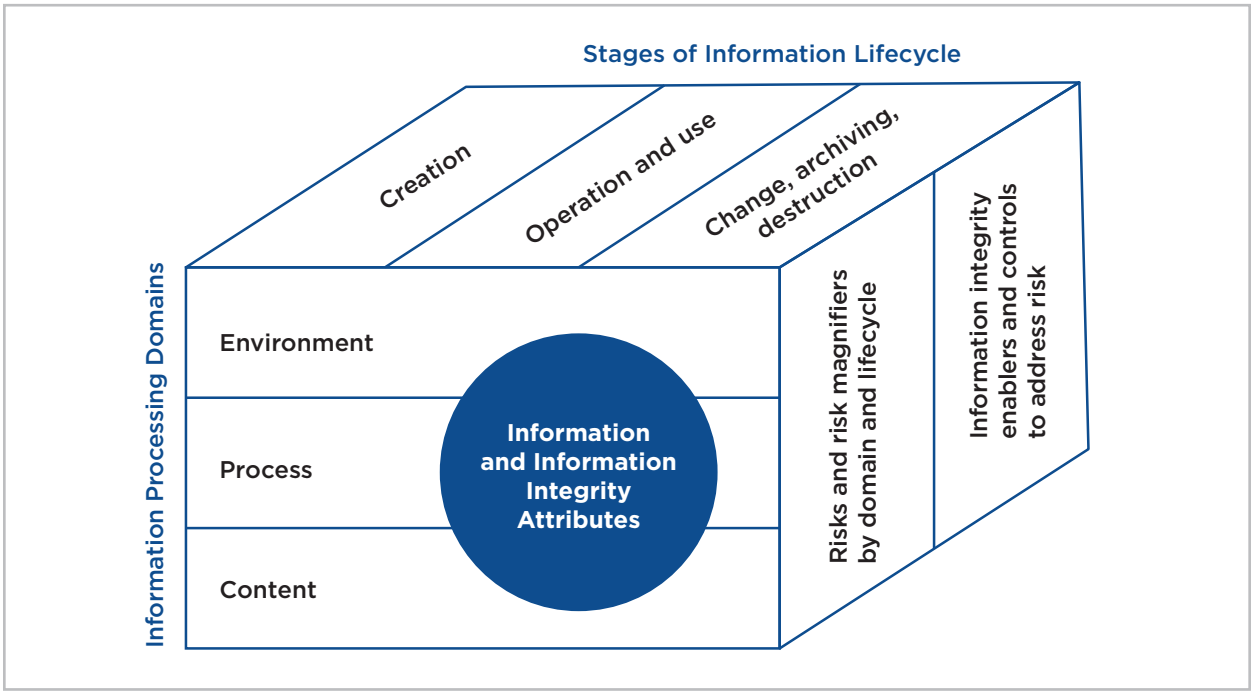
- intended use
- source(s)
- method of compilation
- components and their shared characteristics and relationships
- limitations such as omissions; time period(s) excluded
- measurement uncertainties
- other factors that could affect the appropriate use of the information

Information Integrity Framework

The information integrity framework illustrated in Figure 2 has several components:

- information and the attributes of its integrity
- the information lifecycle
- information processing domains (environment, process and content)
- information integrity risks and risk magnifiers
- information integrity enablers
- controls designed to address risks

FIGURE 2: INFORMATION INTEGRITY FRAMEWORK



Addressing users' information needs must include all the stages of the information lifecycle and their sub-phases. In this framework the information lifecycle is summarized for the sake of conciseness under the three headings of:

- creation
- operation and use
- change, archiving or destruction

The information lifecycle begins with the recognition of the need for particular information. Once this need has been identified, the requirements of the intended users and uses of the information and any ancillary operational and managerial requirements are identified. Then, the conditions, events or instances of interest are identified or defined, together with the attributes of the conditions, events or instances that will be observed, evaluated, measured, recorded and reported. Identifying the intended users and uses of the information is crucial for enabling the information to fit its purpose.

Domains

As previously noted, information is created from content (i.e., "raw data") through the use of processes within an information system (IS) environment that transform content into information. Content can range from various types of raw or sensory data to semi-processed structured and unstructured information, metadata / meta-information, and parameters used to produce information. One or more information processes can transform a collection of inputs into outputs and store them for subsequent use in processing or reporting. Processes operate

within one or more IS environments on which they depend for their continued operational effectiveness. Thus, this framework identifies three information processing domains that affect information integrity and can be used for organizing information integrity risks:

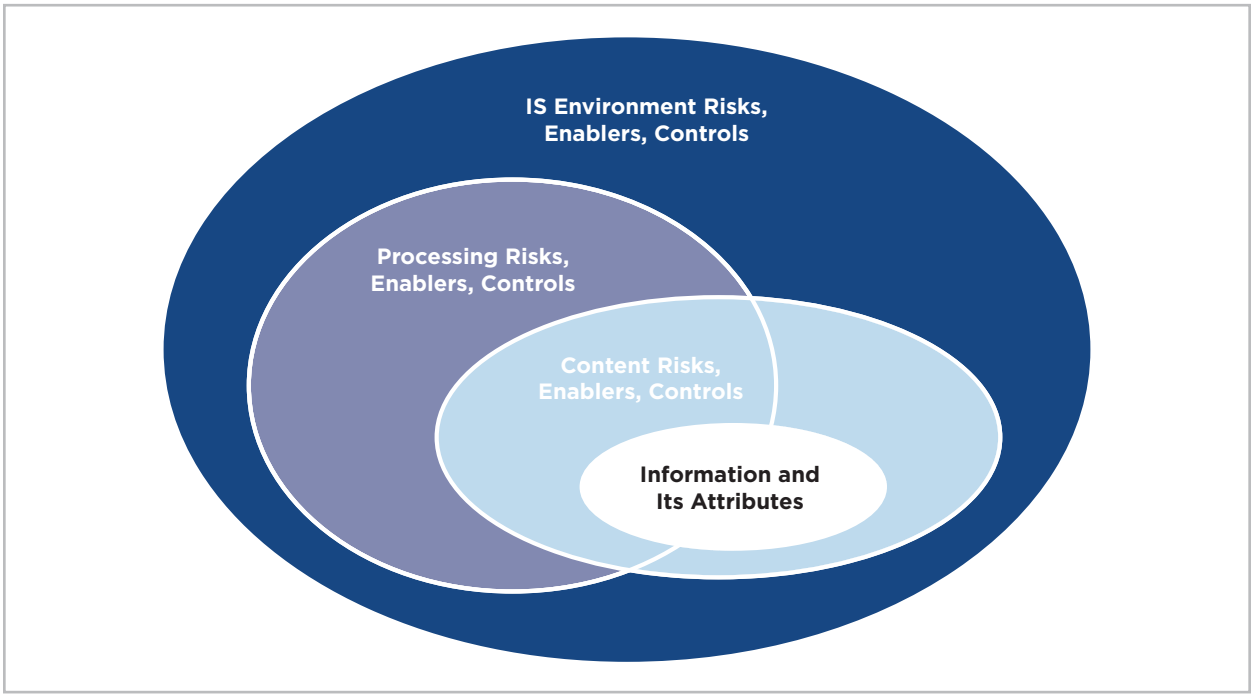
1. content
2. processing
3. IS environment

The three domains have different as well as overlapping information integrity risks which must be mitigated by information integrity enablers and controls tailored to those risks. The risks in all three domains occur throughout the stages of the information lifecycle. The sizes of the risks in these lifecycle stages are determined by the presence or absence of key magnifiers of risk such as the nature of the information system, the complexity of the processes employed to collect content or to transform it into information, and the presence and degree of malicious intent to impair information integrity. For example, most online systems are subject to a high degree of malicious intent aimed at stealing, tampering with, misusing or destroying information.

Enablers and controls are components, features and practices associated with content, processing and IS environment domains that contribute to information integrity. Some enablers are often classified as controls, but many enablers are not controls at all but are, rather, features aimed at enhancing information integrity (e.g., training of personnel enables them to perform their jobs effectively; using servers with excess capacity enables them to operate with fewer malfunctions).

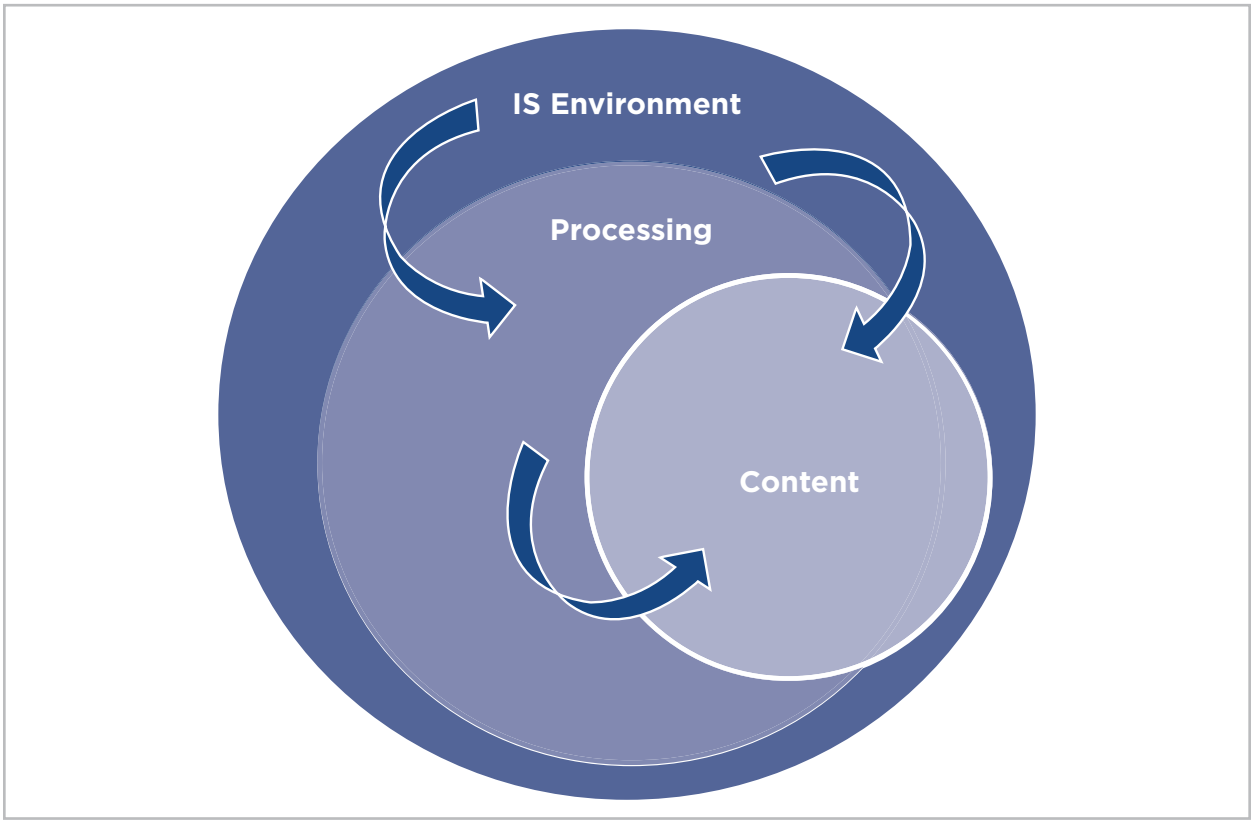
Controls can be thought of as a subset of enablers whose role is to monitor and check whether other enablers are properly designed and implemented, are operating effectively and are updated as required. Thus, controls monitor information-integrity attributes and aspects of the content, processing and IS environment to prevent, detect and correct information integrity impairments, recover from them and mitigate their consequences. For example, a control may monitor system access granting activities and bring instances of non-compliance against policy to management's attention so that corrective action can be taken. In addition to correcting the faulty procedures, steps would be taken to investigate and mitigate the effects of the non-compliance prior to the time that it was discovered.

FIGURE 3: DOMAINS OF INFORMATION INTEGRITY



The diagrams in Figures 3 and 4 further elaborate the relationship between these domains. The IS environment envelops processing and content; processing, in turn, partly envelops the content domain. Weaknesses in the environment domain may detrimentally affect the design of the content and the design, operation and use of the process(es) in the processing domain. Weaknesses in the design of content and the design and operation of the process(es) in the processing domain may permit the integrity of the content to be impaired. For example, cutbacks in IT budgets / services or increases in charge-out rates for IT services may lead business units to use spreadsheets to build applications to meet their information processing needs instead of using the central IT department. This, in turn, may result in weaknesses in the system definition, design, development and deployment processes that ultimately may result in impaired information integrity of the content.

FIGURE 4: RELATIONSHIP BETWEEN IS ENVIRONMENT, PROCESSING AND CONTENT DOMAINS



Content

The **content domain** includes the various types of data, metadata, information and meta-information that are the subject matter whose information integrity is of interest to the organization and decision-makers. Content can include raw data (e.g., scanned codes) as well as semi-processed information (e.g., alphanumeric records of data the codes represent) and parameters used to control processing and information flows (e.g., a table entry that sends the records to a particular device at a particular time). The resulting information in the form of displays, reports, messages and other outputs may be used for planning, decision-making, monitoring and control. Content can be recorded, stored and transferred to and from a variety of media.

The attributes of each condition, event or instance included in the content are affected by the characteristics of the IS environment, and the processing that acts on the content, which may change during the period of interest. Therefore, understanding the attributes involves considering how the IS environment and the processes change during the lifecycle of the information, including related enablers and controls. For example, the sales of ice cream at a particular store location will be affected by temperature, activities in the immediate vicinity,

the people in the area, etc. In designing the information for reporting ice cream sales to management and others, all these characteristics of the environment need to be considered for inclusion in the information.[4]

Some attributes of conditions, events and instances (and environmental characteristics) may be difficult to observe, evaluate or portray (e.g., what the audience of a particular television show finds appealing), or may be undeterminable (e.g., the intent of a borrower to perform maintenance on collateral). In other words, there may be unobservable factors that affect the nature and interpretation of a portion or aspect of the information. For example, the likability of a particular actor in the cast of a television show may impact the way people perceive a show and thus directly affect viewership. If the likability of the actor changes, there may be a direct impact on the viewership; however, this change may not be easily measured or described. The design of information about the television show needs to take into account the impact of such attributes on the fitness for purpose of the information and consider whether the omission of such attributes would make the information misleading.

Attributes may be quantifiable or qualitative and may be measurable to varying degrees at various times in the past, present or future. If an attribute will be measurable in the future, some attributes of the current event or instance are probably contingent on the occurrence of one or more future events. If the measurability of an item depends on the occurrence of a future event, it can become measurable at a date that is certain or one that is uncertain. For example, the number of future sales returns within a 30-day return period relates to a certain period, while the date of collection of an account receivable subject to bankruptcy proceedings is most likely to be uncertain.

The objectivity or subjectivity of an item can have an impact on the ease with which it can be measured. The more subjective an item is, the more difficult it is to measure; in extreme instances, measurement may even be impossible. Nevertheless, subjective assessments made by shoppers on many e-commerce sites are relied upon by other shoppers to determine whether or not they should buy an item. Thus, subjectivity of an item may not be a sound reason for excluding it from decision information. Instead, descriptive information about the nature and extent of the subjectivity might be provided with the subjective item.

Every item of information has **meta-information** associated with it, such as the environmental characteristics noted above, which permits the user to understand and interpret the information. Meta-information is defined as information about information; it describes what the information is and contributes to understanding the information and its attributes by placing it in context, making it fit for purpose. For example, an amount of 35,300 is meaningless because we do not know what the number represents. It could be dollars or miles or numbers of automobiles. If we add a dollar sign, we know that it is a monetary measure, but we still do not know what it represents. If a label “Inventory” is added, we have more information,

4 This example and several others in this publication are drawn from the AICPA whitepaper Information Integrity (January 2013). The principal author of this guide was a member of the task force that produced that whitepaper.

but that information is still not enough to be very useful. However, adding a description such as: "Inventory of Finished Goods for Jones Corporation as at December 31, 20XY, valued under IFRS at the lower of cost and net realizable value" provides a reasonable amount of information, including ownership, date and valuation.

Organizations and individuals may be primarily concerned about integrity of information content. However, as Figure 3 emphasizes, integrity of information content depends on the effectiveness of enablers and controls within the processing and IS environment domains as well as how the two interrelate (i.e., enablers and controls in the content, processing and IS environment domains should complement and reinforce one another and not conflict).

Processing

The **processing domain** includes the content-related activities that identify, collect, record and transform raw data, semi-processed information and parameters into information used for planning, decision-making, monitoring and control. It also includes storage of information for subsequent use in additional processing or reporting.

The processing domain is usually divided into a number of phases—input, process, output and storage (including archiving or destruction)—and sub-phases that contribute to information integrity.[5] Table 1 summarizes key processing activities, enablers and controls by phase and sub-phase.

TABLE 1: PROCESSING DOMAIN PHASES, ACTIVITIES, ENABLERS AND CONTROLS

Phase	Key Activities, Enablers and Controls
Input	• identification or recognition of relevant events or instances triggering other actions • data capture, observation or measurement • data preparation and recording • other activities (see note)
Processing	• transformation of input by aggregating information • performing calculations, logic functions and analyses • performing updates to temporary files (e.g., suspense files) • performing updates to permanent or semi-permanent files, tables and databases • other activities (see note)
Output	• output display • output transmission and distribution to users and other processes • other activities (see note)

5 A distinction is made between the information lifecycle with its three key phases of information creation, operation and use and change and the information processing lifecycle which refers to the path of a particular item of information from the time it becomes identifiable until it is archived or destroyed.

Phase	Key Activities, Enablers and Controls
Storage	• onsite and offsite storage • periodic updating • archiving, anonymization or disposal / destruction of content that is not to be retained • other activities (see note)

Note: All the above phases also include the following generic activities.

Initiation of the phase or process:

- Receipt of data / information from other phases or processes
  - registration / recording / logging of records (including their origins and destinations) and activities performed during the phase or process
  - matching data classifications against access privileges and requested actions against permitted access and functions
    - input, processing, output and storage phase activities as described above
    - error prevention, detection and correction; recovery from failures and mitigation of related consequences
  - assignment / update of metadata
- Transmission / distribution of data / information to other phases or processes
- Back-up and recovery
- Maintenance and change management of the phase or process

Termination of the phase or process.

IS environment

Processes operate within IS environments on which they depend for their continued operational effectiveness. The **IS environment domain** includes the practices used to:

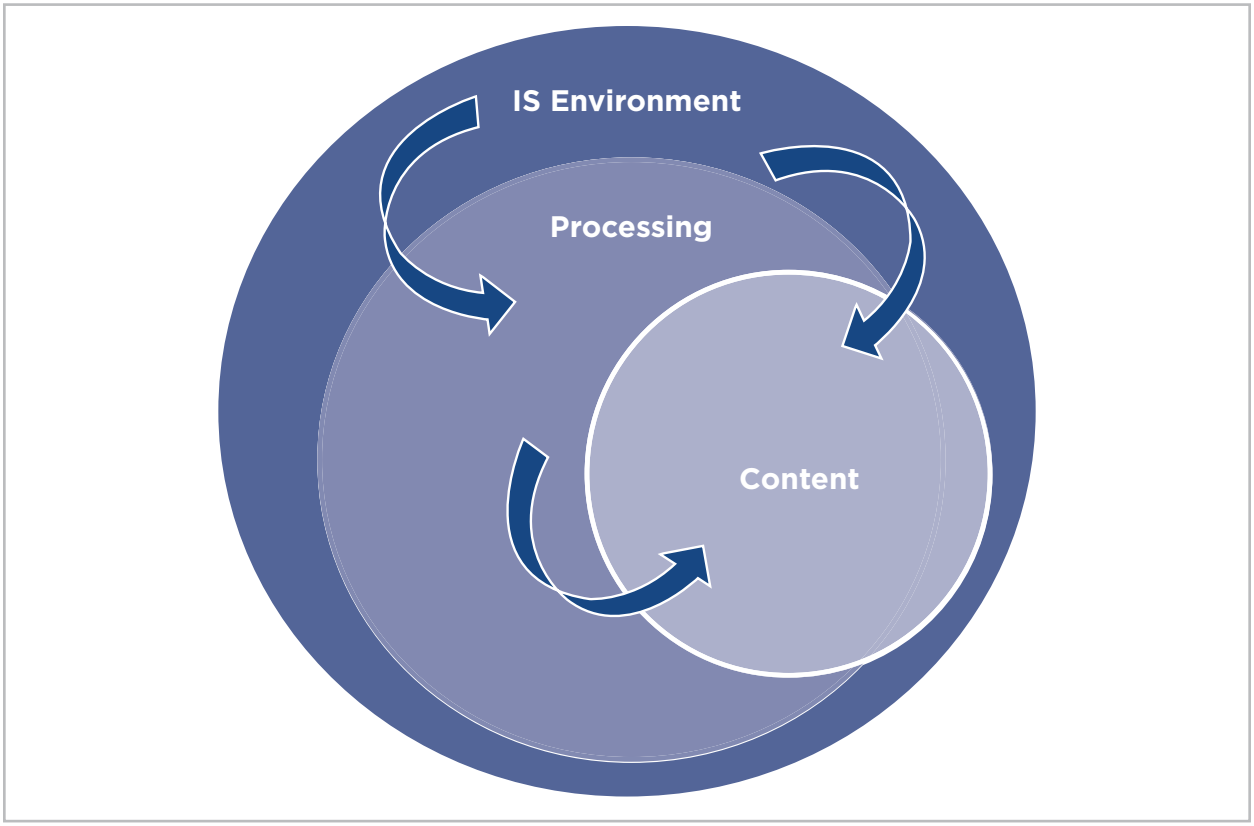
- Manage information to extract and protect the strategic value that high-quality information brings to the organization
- Define, design, develop and deploy processes that ensure that information is fit for its intended users and uses and possesses information integrity
- Operate those processes dependably and consistently
- Ensure the information is:
  - protected against theft, tampering, misuse and destruction
  - available and accessible to authorized users when required
  - verifiable and assured.

It is important to recognize that there may be more than one IS environment affecting processing. Many entities interact with customers, vendors, business partners and others who may have access to or perform processes on their behalf. For example, an entity may outsource some of its processes to one or more vendors. Each vendor has one or more IS environments in which those processes operate and may, in turn, outsource some processing to other vendors. Many entities that rely on cloud computing could be affected by enablers and controls operating in a chain of outsourced processes and their respective IS environments as well as their own in-house IS environment.

Risks and Consequences

This publication takes a risk-based approach that identifies key business risks and the consequences that can flow from those risks introduced in the content, processing and IS environment domains during the information lifecycle. As Figure 5 illustrates, IS-environment risks consist of risks that enablers of and controls over system and information creation, operation and use, and change in the IS environment domain will fail to protect processing integrity and information integrity. Processing risks consist of risks that enablers and controls over information processing in the IS environment and processing domains will fail to protect information integrity. Content risks consist of risks that enablers and controls over information and meta-information in the IS environment, processing and content domains will fail to protect information integrity.

FIGURE 5: RELATIONSHIP BETWEEN IS ENVIRONMENT RISKS, PROCESSING RISKS AND CONTENT RISKS



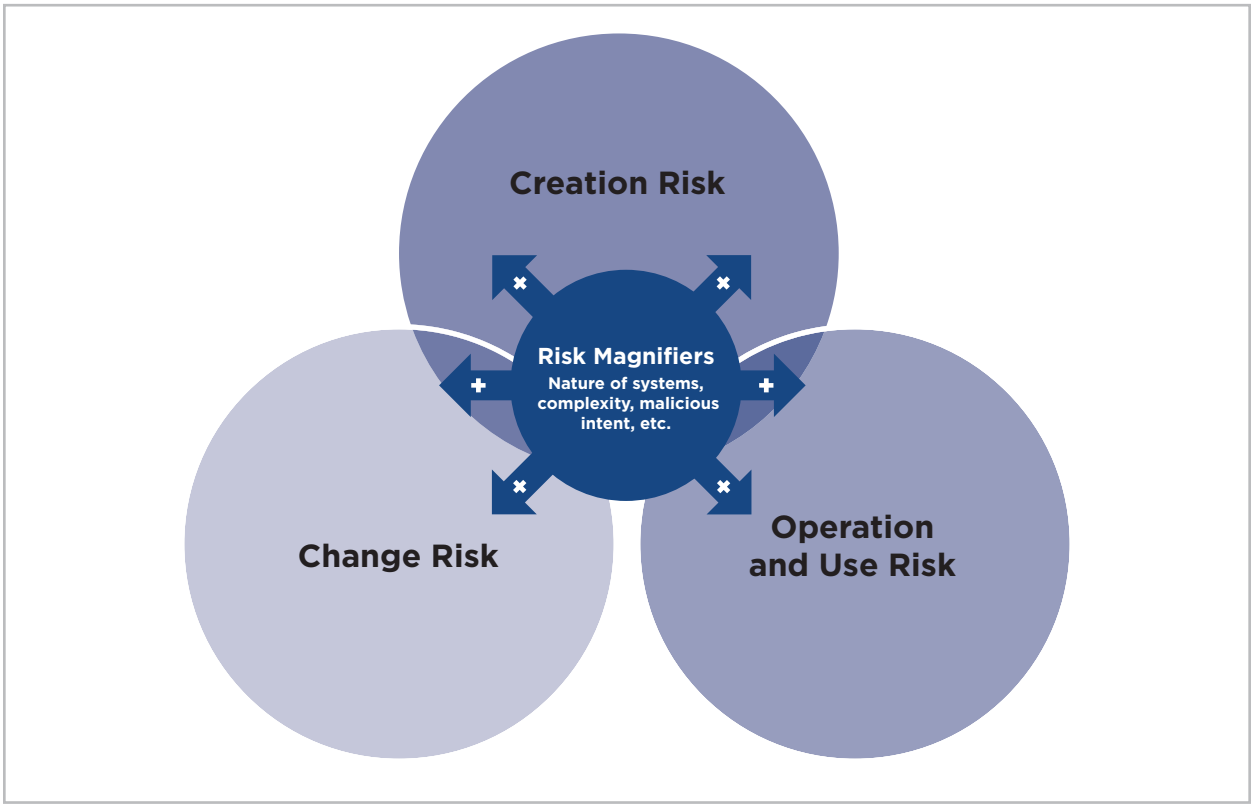
The three domains have different risks as well as overlapping risks which must be addressed by enablers and controls tailored to each domain. In other words, risks in the content domain need to be mitigated by enablers and controls in the content domain, the processing domain and the IS environment domain. Risks in the processing domain need to be mitigated by enablers and controls in the processing domain and the IS environment domain. Risks in the IS environment domain need to be mitigated by enablers and controls in the IS environment domain. For example, the risk of someone tampering with sensitive information should be addressed:

- In the IS environment domain by having a policy covering the granting of access privileges, classifying the information as "sensitive", and defining access privileges and processing rights for the information
- In the processing domain by having a process in place to authenticate a user's identification and access privileges to prevent unauthorized parties from accessing the information or performing unauthorized functions on the information
- In the content domain by having meta-information associated with the information that specifies the sensitivity of the information, the parties that have been granted access privileges to the information and the processing rights that they have been granted

Causes of Information Integrity Impairment Risks

Various risks that can impact information integrity exist throughout the information lifecycle and increase the possibility of material errors and omissions in information leading to erroneous or sub-optimal decisions arising from the use of the information. As Figure 6 illustrates, the risks fall under three headings that correspond to the lifecycle phases of creation, operation and use, and change. Certain factors such as the inherent nature of the system, complexity, and malicious intentions can magnify those risks. These factors must be given special consideration.

FIGURE 6: CAUSES OF INFORMATION INTEGRITY RISKS



Creation risks

Information creation risks include risks of information integrity impairment that arise from the failure of the information design to adequately address users' information needs, information system and information operation and use, and information system and information change, as well as the risks inherent in the activities that occur throughout sub-phases of the information lifecycle. They include the risks that the information attribute / characteristic to be reported is:

- an invalid representation of the desired information (it does not represent what it is purported to represent)
- out of date (measured too early or too late)
- biased, incorrect or insufficiently precise for the intended use
- at an inadequate level of aggregation / disaggregation
- inconsistent / not replicable (between measurers or between measurements) because of qualitative factors and uncertainty
- inconsistent with norms or other sources

Inadequate definition of requirements for information content, processing and the IS environment can create insurmountable barriers to information integrity:

- Definition risks may stem from incomplete or inaccurate understanding of users' and other stakeholders' needs and requirements as well as failure to involve the right participants in the requirements determination process.

- Risks related to the design of information content, processing and the IS environment may stem from failure to follow appropriate design methods, failure to involve the right participants in the design process, and limitations of human judgment that can limit the suitability of a design for a given purpose and thereby permit impairments to occur.
- Risks related to the development and deployment of information content, processing and the IS environment may stem from failure to follow appropriate methods to acquire, develop and deploy systems and information, failure to involve the right participants in the development and deployment process, insufficient testing, and organizational issues that can affect the quality of the outcomes and thereby permit impairments.

Operation and use risks

Operation of systems and production and use of information expose IS environments, processing and content to error, malfunction, malicious attacks and exploitation of known vulnerabilities and unforeseen flaws in the design of information content, processes and systems, and entropy (i.e., the natural tendency of all things to deteriorate over time) resulting in information integrity impairments.

Use risk is the risk that the information or meta-information will be used for other than its intended purpose, used incorrectly, or not used when it should be. The result will be an ill-informed or erroneous judgment or decision. Inappropriate use includes:

- selection of inappropriate information or omission of appropriate information for use in decision-making
- inappropriate substitution of available information for unavailable information
- inappropriate projection of information to other events / instances
- inappropriate combination / transformation / synthesis of information
- inconsistencies in the decision-making process both internal to the user and between users
- inconsistency / misunderstanding between the intent of the information supplier and that of the information user.

All these use risks may result from misinterpretation or misapplication on the part of the intended user or someone other than the intended user of the information or meta-information or due to a lack of information integrity. Misinterpretation or misapplication of information could occur if the information or the meta-information supplied are not appropriate for the intended purpose, are not current, are incomplete, contain errors or are otherwise misleading. Inappropriate application of meta-information would occur, for example, when the information supplied is given excessive weight in the decision-making process or the information does not contain all the meta-information required for the intended use or is not well understood by the user (e.g., use of the information and disclosures written in German by someone with limited knowledge of German).

Risks of misinterpretation or misapplication of information may be reduced by providing information in a format that can be used by the intended users and attaching meta-information (a description) to the information describing the intended user and the intended use of the information, how the information was compiled (what it includes and excludes) and its limitations.

Change risks

As an entity experiences changes in its organization, business practices, personnel, infrastructure and software, it faces increased risks that:

- IS environments and processing will deteriorate
- Changes will defeat or circumvent existing controls
- Unauthorized or untested additions or modifications will be made
- Current content will become irrelevant and will need to be updated, archived or destroyed.

Any of these risks can impair information integrity. Co-ordinated use of information integrity enablers and controls is required to mitigate these risks.

Enablers of Information Integrity

Some risks can be addressed by effective enablers and controls whereas others may need to be addressed by other risk mitigation strategies such as risk avoidance. The enablers of information integrity include content domain, processing domain and IS environment domain enablers and controls. Enablers include policies, procedures and techniques oriented to enhancing information integrity.

Content domain enablers include type of content, type of media, metadata content, metadata creation, use and change management processes. An example of a content domain enabler is the incorporation of features in information (e.g., source information, unique transaction identification codes, time stamps, and other data) that enable users, management, internal auditors and external auditors to verify the integrity of information in an organization's database.

Processing-domain enablers include the activities that form the information processing lifecycle (input, processing, output and storage) and the characteristics of the input, processing, output and storage phases that contribute to information integrity. An example of a processing domain enabler is a table with the identifications of all authorized users of a database that can be checked before permitting a user to access or modify information in the database.

IS environment domain enablers include information governance practices, information design practices aimed at achieving fit for purpose, security practices to protect the information against unauthorized creation, change, misuse and destruction, availability practices to ensure the information is available to and accessible by authorized users, and operations practices to ensure dependability of operations and consistency of information production. An example of an IS environment domain enabler is a disaster recovery plan for the entire organization that ensures the entity is protected against loss of information as a result of intentional and unintentional threats.

Controls

Because information is processed content, the reliability of processes that transform content into information must be part of the information integrity framework. Similarly, since the processing occurs in an IS environment, the reliability of the IS environment surrounding the processing must be part of the framework. Controls can be thought of as a tactical subset of enablers. Their role is to monitor and check whether other enablers are properly designed and

implemented, operate effectively and get updated as required to prevent, detect and correct information integrity impairments and recover from and mitigate the consequences of any impairments that do occur. In other words, enablers are features built into content, process and the IS environment that help achieve information integrity; controls are verification processes and assurance services that monitor those features to ensure they are effective.

An example of a content-domain control is a control that verifies that content attributes are reviewed and approved by an authorized employee before recording data / information in an organization's database. An example of a processing-domain control is a control that verifies that the identity of a user has been checked before the user is permitted to access or modify information in a database. An example of an IS environment-domain control is a control that periodically tests and reviews the effectiveness of the disaster recovery plan for the entire organization that ensures the entity is protected against loss of information as a result of intentional and unintentional threats.

Relationship Between Information Integrity Attributes, Risks, Enablers and Controls

Core attributes of representational faithfulness are the minimum criteria that must be satisfied to an acceptable level for a given information item or information set to be judged as having integrity. In other words, all are necessary, but none is sufficient by itself to warrant the label. A monthly sales figure that omits one day of sales (i.e., is incomplete) but is otherwise accurate is not representationally faithful; a supermarket price table that has not been updated for this week's advertised sale prices (i.e., is not current) is not representationally faithful; an accounts receivable aging with dating errors (i.e., is not accurate) is not representationally faithful; a supplier list that contains fictitious suppliers (i.e., is not valid) is not representationally faithful; and so forth.

Because of the inherent limitations of information processing systems and the people creating and operating those systems, perfect completeness, currency, accuracy and validity are not achievable. The limitations of enablers and controls mean that representational faithfulness is subject to some degree of imperfection; the tolerable degree of imperfection (materiality) is defined differently in different domains and contexts.

An assessment of the effectiveness of enablers and controls can help decision-makers assess the degree of representational faithfulness possessed by an information item or information set so that, if necessary, remedial action can be taken to achieve an acceptable level of information integrity or to discount the amount of reliance placed on the information.

The framework discussed in this publication should be a valuable resource to organizations looking for guidance on achieving information integrity. Other related publications arising from this one provide more detailed analyses of information integrity risks, specific enablers and controls that can be used to respond to those risks, and assurance services that can be used to assess the suitability of the design and operating effectiveness of those enablers and controls.

Definitions

In this publication we use a number of key terms in a particular way. These terms are identified and defined below.

Accuracy: includes concepts such as correctness and precision of calculation, measurement or estimation as well as consistency of processing over time and across items

Change: one of three stages of the system and **Information Lifecycle** (the others being **Creation, Operation and Use**) used to organize risks in this publication; includes replacement of a pre-existing organizational element, business practice, infrastructure, or software with a revised version; also includes departure or replacement of personnel, archiving and destruction of information

Completeness: the completeness of processing, including all the time periods, data items and attributes of the data items required for the intended purpose as well as the **Metadata** with the **Contextual Information** required to understand the **Information**

Complexity: presence of a large number and/or variety of interacting components

Content: all types of data used to generate **Information**, including **Raw Data**, sensor data, semi-processed **Information**, **Metadata**, and parameters

Contextual Information: see **Meta-Information**

Control: feature or activity that monitors and checks elements of content, processing or the IS environment against criteria to prevent, detect or correct **Information Integrity** impairments; can be thought of as the tactical subset of enablers whose role is to monitor and check whether other enablers are properly designed and implemented, are operating effectively and are updated as required

Creation: one of three key stages of the system lifecycle (the others are **Operation and Use** and **Change**) used to organize risks in this publication; consists of activities such as definition, design, acquisition, development and deployment

Currency: evaluated relative to the time period or cut-off date of the **Information** relative to its purpose and the timing of its use

Data: a recorded set of qualitative and quantitative measurements of the characteristics or attributes of events and instances; may be presented in various formats, ranging from structured alphanumeric data to unstructured text to audio to images (see **Raw Data**)

Data Quality: a label for a variety of concepts describing desirable attributes of data ranging from relevance and usefulness to integrity; at a minimum, the level of completeness and accuracy of valid data captured and processed for a specific purpose

Downstream: the subsequent use of current **Information** (see **Upstream**)

Enabler: component, feature or practice associated with the content, process or IS environment domain that contributes to **Information Integrity**

Event: category of occurrences to be captured by a system of business rules

Event Instance: actual and particular occurrence of the event type to be captured

Fit for Purpose: relevant for its actual or intended use; applicable, clear, understandable, and at an appropriate level of granularity or aggregation

Information Activities: individual components or tasks that aggregate into a process (see **Process**).

Information: data presented to a user in a meaningful context for a given purpose (see **Data** and **Raw Data**)

Information Assurance: incremental **Information** or **Meta-Information** attached to subject matter that serves to increase the confidence of a user in the integrity of that subject matter

Information Governance: policies, standards, procedures and other mechanisms established by the board of directors and executive management to make **Information Integrity** a high priority within the organization

Information Integrity Impairment Risk: see **Risk**

Information Integrity: **Representational Faithfulness** of the **Information** to the condition or subject being represented by the **Information**

Information Lifecycle: the process running from the **Specification** of **Information** to its retirement (archiving or destruction); in this publication: creation, operation and use, and change (including retirement, permanent archiving, anonymization or destruction)

Information Processing Lifecycle: a part of the overall **Information Lifecycle**, including the following phases:

a. input: creation or identification of **Data**, observation or measurement, documentation or recording
b. processing: analysis, calculation, transformation or aggregation (to transform data into information)
c. storage or archiving
d. periodic updating

e. output display, transmission and distribution
f. use
g. archiving, anonymization or destruction

Information Quality: a label for desirable attributes of **Information**, including relevance,[6] usefulness and **Representational Faithfulness**[7]

IS Environment: all elements of the supporting organizational infrastructure relied upon by the processing domain, including policies, standards, procedures and IT services

Metadata: describes the content, context and structure of **Raw Data** before it is turned into **Information** (e.g., description, purpose, origin, used by, owned by, custodian / steward, standard, classification for security / privacy, access privileges, location, version, date / timestamp, retention / disposal requirement, lineage / audit trail, assurance) (see **Meta-Information** and **Raw Data**)

Meta-Information: enables information processing systems to maintain information integrity during processing and for users to understand and use information appropriately; the context for understanding "processed data" (see **Metadata**)

Operation and Use: one of the three stages of the system and **Information Life Cycle** (the others being **creation and change**); business activity that involves the use of content, **Information**, or systems

Process: all activities that transform a collection of inputs (e.g., **Raw Data** or other items from the content domain) into outputs and store them for subsequent use in processing or reporting

Processing Integrity: completeness, timeliness, accuracy and validity of system processing in the context of the aim or purpose of the system and its intended users (see **Information Integrity**)

Raw Data: data requiring further processing to be useful (see **Information**)

Representational Faithfulness (of Information): a depiction connected to the actual phenomena (or the conformity of **Information** to the item to which it corresponds). According to Financial Accounting Standards Board® (FASB), to be a perfectly faithful representation, a depiction must be complete, neutral, and free from error.[8] In this publication, representational

6 According to FASB (2010), relevant financial information is capable of making a difference in the decisions made by users. Information may be capable of making a difference in a decision even if some users choose not to take advantage of it or already are aware of it from other sources. Financial information is capable of making a difference in decisions if it has predictive value, confirmatory value, or both.

7 Committee of Sponsoring Organizations of the Treadway Commission (COSO) (2013) identifies the following determinants of information quality: timely, current, accurate, complete, accessible, protected, verifiable, retained.

8 FASB, *Conceptual Framework for Financial Reporting: Objective of Financial Reporting and Qualitative Characteristics of Decision-Useful Financial Reporting Information* September 2010.

faithfulness of **Information** is determined with reference to whether it is complete, current, accurate and valid. These characteristics must be assessed in the context of the intended or actual use of the **Information** (see **Processing Integrity**)

Relevance (of **Information**): applicability to the purpose for which the **Information** is created; has the capacity to make a difference to users' decisions based on that **Information**

Risk (of **Information Integrity Impairment**): may undermine or threaten one or more of the core attributes of **Information Integrity**; can arise from intentional malicious acts or unintentional errors; organized into the three information system and **Information Lifecycle** phases: creation, operation and use, and change

Risk Magnifier: a factor that magnifies a risk (e.g., complexity, nature, malicious intent, etc.)

Subject / Subject Matter: set of phenomena (i.e., conditions, events or instances) about which **Information** and accompanying **Meta-Information** are provided

Subject Matter Information: Information and **Meta-Information** that portray a subject / subject matter based on the observation, evaluation, measurement and representation of the subject matter (against criteria)[9]

Threat (to **Information Integrity**): arises from internal and external sources; may come from people, technology and the environment; may stem from intentional and unintentional actions (see **Risk**)

Timeliness: **Information** available in time to be used for its intended purpose

Understandability (of **Information**): appropriate level of detail or aggregation, labelling and contextual **Information** for the intended use

Upstream: the **Creation** and/or previous processing of current information (see **Downstream**)

Usefulness / **Usability** (of **Information**): sufficient understandability, relevance and integrity for the purpose for which the **Information** is intended

Validity (of **Information**): represents what it purports to represent; results from authorized processes, complies with policies, laws and regulations, is properly formatted, authentic, traceable to its source(s) and its ultimate destination, verifiable and free from bias

9 Adapted from *International Auditing and Assurance Standards Board Handbook of International Quality Control, Auditing Review, Other Assurance, and Related Services Pronouncements* Volume 1, 2012.

References

AICPA (2013). *Information Integrity*. Whitepaper. (January). AICPA.

Boritz, J.E. (2004). *Managing Enterprise Information Integrity: Security, Control and Audit Issues*. IT Governance Institute.

Boritz, J.E. (2005). "IS Practitioners' Views on Core Concepts of Information Integrity." *International Journal of Accounting Information Systems 6* (2005), 20.

CICA (1998). *Information Technology Control Guidelines*. Third Edition. CICA.

The Committee of Sponsoring Organizations of the Treadway Commission (COSO) (2013). *Internal Control –Integrated Framework*.

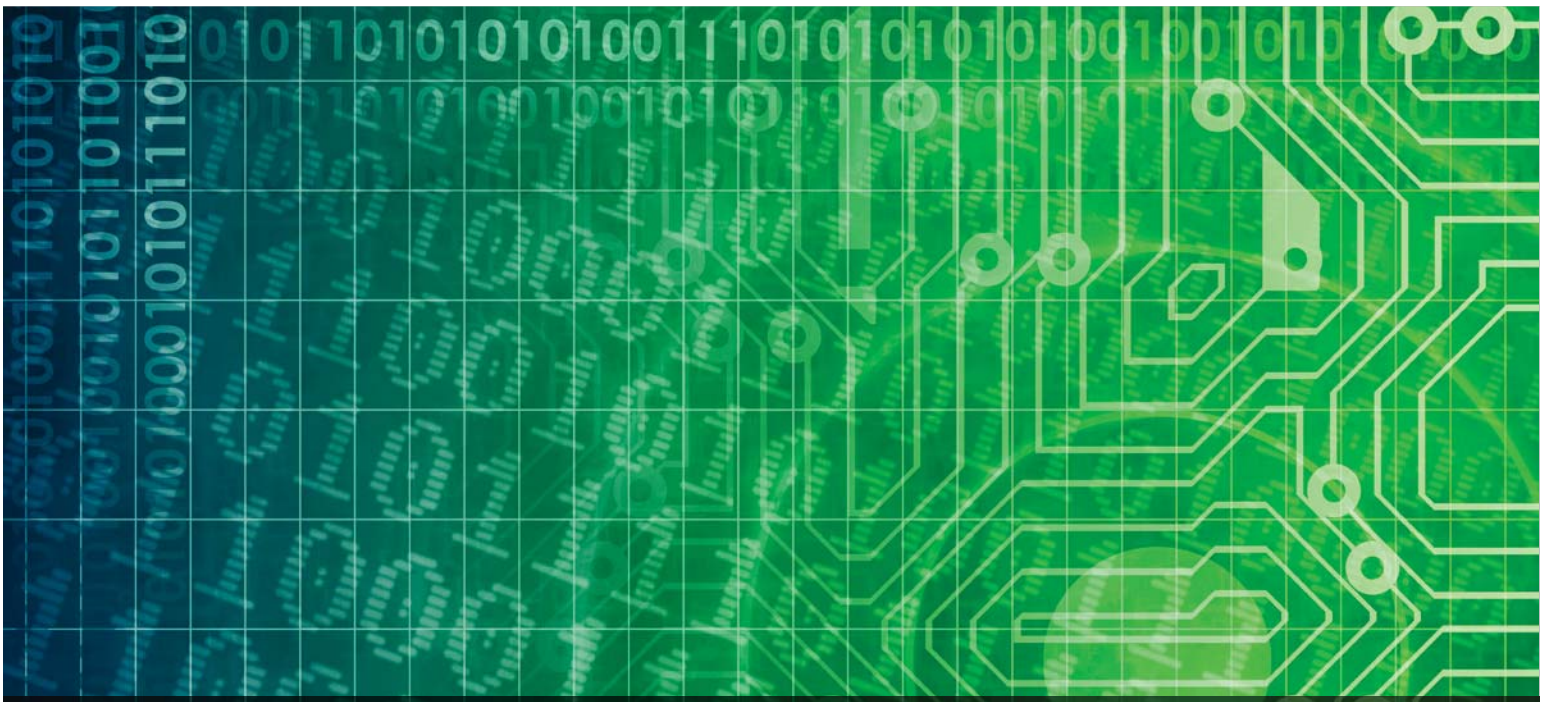
Financial Accounting Standards Board® (FASB) (2010). Concepts Statement No. 8, Conceptual Framework for Financial Reporting, Chapter 3, "Qualitative Characteristics of Useful Financial Information." FASB.

International Organization for Standardization (ISO) (2001). 15489-1. Information and documentation – Records management – Part 1: Concepts and principles. ISO.

ISACA (2012). COBIT (Control Objectives for Information and Related Technologies) 5 Framework. ISACA.

Neely, M. P. and Cook, J. S. (2011). "Fifteen Years of Data and Information Quality Literature: Developing a Research Agenda for Accounting." *Journal of Information Systems* 25 (1) (Spring), 79 – 108.

Wang, R. Y. and Strong, D. M. (1996). "Beyond accuracy: What data quality means to data consumers." *Journal of Management Information Systems* 12 (4), 5 – 34.



CHARTERED
PROFESSIONAL
ACCOUNTANTS
CANADA

277 WELLINGTON STREET WEST
TORONTO, ON CANADA M5V 3H2
T. 416 977.3222 F. 416 977.8585
WWW.CPACANADA.CA