Chapter 1 Information Integrity

Stakeholders use various types of information in making decisions, understanding, interpreting or using other information and generally increasing their body of knowledge concerning historical events or the current state of affairs. To make the best decisions, they need to have confidence in the integrity of the information. This ranges from the traditional financial information based on generally accepted accounting principles (GAAP) to sustainability reporting of information such as baseline-year emissions data, energy produced/consumed and resource reserves (bbl, tons, etc.). Service organizations routinely produce reports on performance measured against metrics defined in service level agreements and commitments and other criteria.

The value of information comes from its relevance,[1] usefulness/usability and integrity that collectively reflect its quality.[2] Information integrity is a key aspect of information quality. Integrity means an unimpaired or unmarred condition — entire correspondence of a representation with an original condition (Webster's Third New International Dictionary). The word integrity is used in many contexts. Applied to information, integrity is the representational faithfulness of the information to the subject matter (e.g., the events or instances) being represented by the information. The FASB's and IASB's Conceptual Framework (FASB, 2010; IASB 2010) discusses representational faithfulness in the context of financial information and notes that to be a perfectly faithful representation, a depiction must be complete, neutral, and free from error.[3] The definition of information integrity provided in ITGI's COBIT 5 (ITGI, 2012) defines it by the attributes of completeness and accuracy.[4] The CICA's ITCG (1998) includes additional attributes such as authorization, timeliness, consistency and segregation of incompatible functions. An extensive series of studies of data quality at MIT by Wang et al. (1993, 1995, 1996) has provided

[1] Relevance includes attributes such as feedback value and predictive value that are not included in information integrity. Usefulness/usability include attributes such as perceived usefulness and ease of use.

[2] COSO (2013) identifies the following determinants of information quality: timely, current, accurate, complete, accessible, protected, verifiable, retained.

[3] FASB (2010) also identifies several characteristics of information that are required to make the information useful: relevant, comparable, verifiable, timely and understandable. COSO (2011) identifies the following attributes that contribute to the quality of information: sufficient; timely; current; correct; accessible; protected; verifiable; and retained.

[4] IT Governance Institute, *COBIT (Control Objectives for Information Technology) 5* Rolling Meadows, Il: ITGI, 2012.

valuable insights into information users' views about data quality, which also have a bearing on data integrity and information integrity. In summary, information integrity focuses on a narrower set of information attributes than information quality but information integrity is the sine qua non of information quality as it would be hard to imagine information having quality in the absence of integrity (ITGI, 2004).

Table 1.1 summarizes the key attributes of information integrity identified in a number of key sources. The main purpose of Table 1.1 is to document the authoritative sources of the key attributes that are used to define the term information integrity in this publication and the content, process and IS environment domain enablers of information integrity. The reason for doing this is that the attributes of information integrity are the minimum criteria by which the degree of information integrity can be assessed. In other words, to come to a conclusion on the integrity of information about a subject matter, it is necessary to assess the completeness, currency, accuracy and validity of the information. This assessment must be made while keeping in mind the purpose for which the information is to be used and the minimum level of information integrity that can be tolerated given that purpose. Note that detailed enablers and controls related to the attributes of information integrity are not the same in each framework listed in Table 1.1; that is, completeness enablers and controls in one framework are not the same as completeness enablers and controls in another, although there are many overlaps.

The relevance, usefulness/usability and integrity of the information must be assessed relative to the purpose for which the information is produced as well as its actual use. Information can be structured (e.g., accounting transactions), partly structured (e.g., object-oriented data bases) or unstructured (e.g., raw data such as a string of digits). For purposes of this publication, information consists of representations regarding one or more events and/or instances that have been created for a specified use. Such events or instances can have numerous attributes and characteristics that may or may not be included in a set of information, depending on the intended use of the information. Some uses may require a small number of attributes to be recorded about a given set of events or instances whereas other uses may require a large number of attributes to be recorded about those same events or instances.[5]

For the information to be useful, it is important to provide meta-information that describes the purpose of the information and other contextual information necessary to make use of the information.

In summary, information is prepared for a specified purpose and includes: (1) the data about the characteristics of the specific events or instances which have been included within the purview of the information by virtue of the information design aimed at the specified purpose, (2) information about the environment in which the events occurred or the instances existed, and (3) other information necessary for the observations to be used for their intended purpose. Information integrity is determined based on both the information's consistency with its meta-information and its representational faithfulness.

Information integrity is not a binary quality. It can vary from 0 to 100%. In using information, users need to assess their level of confidence in the integrity of the information. Otherwise, they may place unwarranted reliance on the information. Confidence in information integrity can come from many sources, including:

1. Additional information supplied by the party responsible for the information such as, a description of the process that produced the information.

2. The reputation of the responsible party.

3. Knowledge possessed by the user, whether pre-existing or specifically obtained for the purpose of evaluating the integrity of the information.

[5] For example, a log of accounting transactions used to assess the completeness of information transmitted from a branch to headquarters may only require an information identifier and a message digest that can be checked for completeness of transmissions for each item. In contrast, an audit trail used to trace transactions from cradle to grave and vice versa, may need an information identifier, message digest, date stamp, source, destination and intermediate processing steps that were performed on the information.

4. Validation of the information by a third party with knowledge sufficient to evaluate the integrity of the information, which may or may not be in the context of a professional engagement.

5. Assurance provided by an independent third-party based on procedures performed to evaluate the integrity of the information provided by the responsible party. This assurance would normally be provided as an assurance engagement carried out in accordance with generally accepted standards.

This publication is comprehensive, in that it has identified and consolidated all information integrity controls into a single framework, organized by integrity attribute, identifying threats to information integrity and enablers and controls within the three domains of content, process and the IS environment.

Table 1.1
Comparison of Core Attributes of Information Integrity and their Enablers with Key Frameworks

ISACA COBIT 5 (2012)	CICA ITCG	FASB/IASB Conceptual Framework (2010)	COSO (2013)	MIT Research Group	ISO 15489-1:2001	Core attributes and enablers in this publication
Integrity						
N/A	N/A	Relevant = Decision Usefulness: Predictive value, confirmatory value; Timely	N/A	Relevant; Value added	N/A	Fit for Use – Relevant; Useful; Usable
Complete	Complete	Complete; Part of Representational Faithfulness	Complete	Complete	Complete	Complete
N/A	Timely	N/A	Current/ Timely	N/A	N/A	Current; Timely
Accurate	Accurate	Free from Error; Part of Representational Faithfulness	Correct	Accurate	N/A	Accurate - Correct; Free from Error; Sufficiently precise
N/A	Valid	N/A	N/A	N/A	Unaltered Authentic	Valid - Authorized (in accordance with laws, policies, etc.); Traceable; Authentic; Non-Repudiable; Believable; Credible; Assured

ISACA COBIT 5 (2012)	CICA ITCG	FASB/IASB Conceptual Framework (2010)	COSO (2013)	MIT Research Group	ISO 15489-1:2001	Core attributes and enablers in this publication
N/A	Authorized	N/A	N/A	N/A	N/A	Part of Valid (see above)
N/A	Neutral; Part of Valid (see above)	Neutral; Part of Representational Faithfulness	N/A	Objective	N/A	Part of Complete, Current, Accurate and Valid
Intrinsic Quality						
Accurate	Accurate	Free from Error; Part of Representational Faithfulness	Accurate	Accurate	N/A	Accurate - Correct; Free from Error; Sufficiently precise; Consistent
Correct	N/A	Free from Error;	N/A	N/A	N/A	Part of Accurate (see above)
Reliable	N/A	N/A	N/A	N/A	Full; Accurate; Dependable	Part of Complete, Current, Accurate and Valid (see above)
Objective	Neutral; Part of Valid (see above)	Neutral; Part of Representational Faithfulness	N/A	Objective	N/A	Part of Complete, Current, Accurate and Valid (see above)
Believable	N/A	Credibility is part of verifiability	N/A	Believable	N/A	Part of Valid (see above)
Reputable	N/A	N/A	N/A	Reputable	N/A	Part of Valid (see above)
Contextual and Representational Quality						
Relevant	N/A	Relevant = Predictive value, confirmatory value; Timely	N/A	Relevant; Value added	N/A	Fit for Use – Relevant; Useful/Usable
Complete	Complete	Complete; Part of Representational Faithfulness	Complete	Complete	Complete	Complete
Current	Timely	Timely, part of Relevant above	Timely	N/A	N/A	Current; Timely
Appropriate Amount	N/A	Decision Usefulness	N/A	N/A	N/A	Fit for Use (see above)
Concise						Fit for Use (see above)
Comparable; Consistent	Consistent	Comparable; Consistency is a means to comparability	N/A	N/A	N/A	Part of Accurate (see above)
Understandable; Interpretable	N/A	Understandable – clear and concise;	N/A	N/A	Presentable; Interpretable	Fit for Use – Understandable; Transparent;

ISACA COBIT 5 (2012)	CICA ITCG	FASB/IASB Conceptual Framework (2010)	COSO (2013)	MIT Research Group	ISO 15489-1:2001	Core attributes and enablers in this publication
		transparent = faithfully represented (neutral, complete; understandable)				Appropriate granularity and aggregation
Ease of handling/ manipulation	N/A	N/A	N/A	N/A	N/A	Fit for use (see above)
Security/ Accessibility						
Available when required;	N/A	N/A	Accessible;;and retained	Accessible;	Locatable; Retrievable	Available; Accessible -.a means of achieving information integrity
Timeliness	Timeliness	Timeliness , part of Relevant above	Timely	N/A	N/A	Current; Timely
Restricted access	Segregation of Incompatible Functions	N/A	Protected	Secure	Protected against unauthorized alteration or disposition/ destruction	Secure – access control, segregation of incompatible functions; protected/ safeguarded (against tampering, loss, destruction) - a means of achieving information integrity
Believable	N/A	Credibility is part of verifiability	N/A	Believable;	N/A	Part of Valid (see above)
Reputable	N/A		N/A	Reputable	N/A	Credible; Assured.
Objective	N/A	Verifiable; Neutral	Verifiable;	Objective	Objective	Part of Complete, Current, Accurate and Valid (see above)
Compliance With laws, regulations and contractual arrangements	N/A	N/A	N/A	N/A	N/A	Part of Validity (see above)
Confidentiality Protection against unauthorized (read) access or disclosure	N/A	N/A	N/A	N/A	Protected against unauthorized access or removal	Outside the scope of this publication although confidentiality related enablers and controls overlap with integrity related enablers and controls

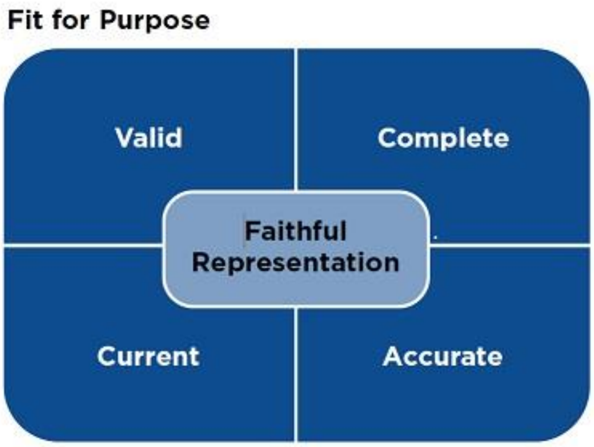
N/A = not applicable or not addressed

Core Attributes of Information Integrity

An underlying requirement of information is its fitness for its actual or intended use. To be fit for use, information needs to be relevant for its actual or intended use, useable (clear, understandable, and at an appropriate level of granularity or aggregation) and possess information integrity; that is, be free from material error, omission or distortion. An extensive review of literature in this area led to adopting the definition for information integrity as the representational faithfulness of information to the true state of the subject that the information represents or purports to represent, where representational faithfulness is composed of four essential qualities or core attributes:

- completeness,
- currency,
- accuracy, and
- validity.

Figure 1.1 Representational Faithfulness



The required degree of achievement of each of these four core attributes depends on the purpose for which the information is produced. In other words, the level of information integrity depends on the degree to which fitness for purpose is achieved by all of the core attributes individually and in combination. The tolerable amount of impairment in information integrity (materiality) is determined by the purpose for which the information is produced.

Completeness

Completeness is the starting point for information integrity. Information cannot be representationally faithful if it is materially incomplete.

In an information processing context, completeness relates the completeness of the population of events represented by the information, completeness of the attributes of the events included in the information, completeness of the values of those attributes and completeness of the related meta-information. To assess whether information is complete, it is necessary to understand the purpose for which the information is provided and the degree of information incompleteness that is tolerable before it materially affects fulfilment of that purpose.

Measurement and processing limitations of information processing systems will prevent 100% real-time completeness, especially for subject matter that changes frequently. This, in turn, prevents 100% accuracy. For example, if there are three cars on an auto dealer’s lot, two cars in the database, and one car in a receiving transaction that has yet to update the database, then a process that ensured processing completeness would contribute to database accuracy as well. In other words, the degree of information completeness that is achieved sets the upper bound on the degree of accuracy that is achievable. Thus, the degree of information completeness that is provided for in creating the information should be clearly communicated to the intended users of the information. In this regard, it is important to understand whether the information pertaining to the subject matter is complete as of a particular point in time or for a specified period of time.

Currency

Information currency is affected by real world changes over time (as well as by information processing delays) with a commensurate impact on information accuracy. Since time is continuous, completeness must be understood in a context that defines acceptable limits for information currency. For example, if certain information, such as cash receipts is only used to update accounts receivable on a weekly basis, then accounts receivable could be considered current if it was missing a day’s worth of transactions. However, in the world of high-frequency trading timeliness/currency is measured in milliseconds.

Many practitioners and academics often represent processing timeliness and information currency as aspects of information completeness; however, because of their unique relationship to the dimension of time and the change that time engenders, it is useful to identify currency as a separate attribute of information integrity.

Timeliness: Timeliness is different from currency, with a relationship to both information relevance and usefulness. Timeliness is a measure of how fast the content can be collected and transformed into information to be useful to a user. Currency is a measure of how closely the content reflects the present reality or condition of the subject matter being represented (or the condition of the subject matter at a specified cut-off date). Timeliness of processing is considered to be a necessary condition to achieve currency of information. However, the beneficial impact of information processing timeliness on decision-making also extends beyond the consideration of information integrity to other dimensions of information quality such as relevance and usefulness. As noted in connection with completeness, to assess whether information is current, it is necessary to understand the purpose for which the information is provided and degree of non-currency that is tolerable before it materially affects fulfilment of that purpose.

Date/Time Stamping: Given the foregoing discussion, it is important to recognize that completeness, currency and timeliness of processing are pre-requisites for a meaningful focus on information accuracy. There must be understood tolerances for omissions and delays in the timeliness of processing. Since the tolerances for information integrity may differ among stakeholders, it may be impractical to set standards for information currency or processing timeliness. Standards that meet the most demanding users’ currency/timeliness requirements may be stricter than those of many other users and impose unnecessary costs of achieving such stringent currency requirements on information processing systems. Instead, various forms of time stamping can be useful metadata to enable stakeholders to assess the temporal limitations of information integrity. When content is enhanced by time stamping, its currency is not improved; however, its degree of currency and accuracy are more understandable and more verifiable.

Historical Information: The attribute of currency does not mean that historical information is not valuable. Rather, it means that the information user or recipient is aware of how recent the content is. In a historical context, currency means that the most current events or transactions at a cut-off point in time in the past were used to record the information at that point in time.

Cut-off: In accounting, assertions such as measurement and cut-off convey that the information about a subject matter relates to the correct time period.

Accuracy

Accuracy (freedom from error) asserts that what is represented in the information corresponds to a real world object or event with some degree of precision. For example, if an auto dealer's database states there are two cars on the lot but there are actually zero cars on the lot, then the database is inaccurate. But if the auto dealer's database states there is one car, then it is less inaccurate than if it states there are two cars. If an insurer's database states that a policyholder is single and smokes, but she is married and doesn't smoke, then the database is inaccurate. If the insurer's database states that the policyholder is single and doesn't smoke, then whether it is more accurate than if it states that she is married and smokes would depend on the impact of the individual facts on the use of the information; for example, their respective impacts on the policy premium. In these examples, the databases may be inaccurate simply because certain transactions that are needed to update the databases have not been completely processed yet. Thus, as noted previously in the discussion of completeness, accuracy depends on completeness, in the sense that the degree of information completeness sets the upper limit on information accuracy. The materiality of any inaccuracy depends on the anticipated or actual use of the information.

Precision: The concept of precision of a measurement or estimate is also relevant to this concept, especially for information based on judgements. For example, a point estimate of the fair value of an asset or liability could be judged to be accurate even though it falls within a range that represents a comparatively large band of imprecision around the point estimate. As noted in connection with completeness and currency, to assess whether information is accurate, it is necessary to understand the purpose for which the information is provided and degree of imprecision that is tolerable before it materially affects fulfilment of that purpose. The degree of precision that is provided for in creating the information should be clearly communicated to the intended users of the information. In this regard, it is important to understand the units of measurement that were used and the measurement standards that were applied.

Correctness: In contrast with accuracy, which relates to measurement precision, correctness relates to conformity with an acknowledged standard. While there are subtle distinctions between accuracy, correctness and freedom from error these terms are considered as synonyms in this publication.

Valuation/Measurement: In accounting, accuracy is related to assertions such as valuation and measurement to convey that the information about a subject is valued or measured materially in accordance with a specified accounting framework such as International Financial Reporting Standards to the required degree of precision. Forecasts of future states and events which are widely used in accounting estimates generally result in a range of values (a reasonable range) rather than a single point estimate.

Consistency: When more than one item of information is produced about subjects with the purpose of comparing them on one or more dimensions, then this requires the use of consistent frameworks, measures and methods over time and across subjects, environments, populations and attributes.

Validity

Representational faithfulness of information implies that the information corresponds to real conditions or to characteristics of physical objects or, in the case of intangible objects, that the information represents the correct application of authoritative rules or relationships. In general, conditions, rules or relationships are valid if what they purport is true. In some branches of science, the concept of validity is used to describe constructs that measure what they purport to measure; for example, an IQ test is considered to be valid if it measures intelligence. In this publication, validity is used in a way that is consistent with this meaning but is a bit broader, incorporating ideas of authorization, non-repudiation/authentication, and non-duplication.

Authorization: In a business context, conditions, business rules or relationships are established or approved by parties with the delegated authority to do so. A business rule is a statement that defines or constrains some aspect of the business. It is intended to control or influence behaviour. For example a business rule might state that a credit check is to be performed on all new customers. Thus, transactions are valid if they were initiated and executed in accordance with one or more business rules by personnel or systems that have been granted the authority to do so and if approvals are authentic and within the scope of the authority granted to the approver(s). For example, if the credit limit assigned to a new customer reconciles to the company's rules and procedures used to set credit limits, the credit limit would be "valid." Thus, the concept of validity includes elements of accuracy and correctness, authorization and authenticity of the identity of the authorizer.

Compliance with policies, laws and regulations: Information integrity implies that information is produced in compliance with the policies, laws and regulations governing its creation, use, change, retention and destruction.

Data Formatting: In an information processing context, validity also refers to complying with data and identification code formatting rules, including data types, check digits and other such rules.

Non-Repudiation/Authentication: An important aspect of information validity in a transaction processing or information management context is non-repudiation/authentication, which requires a combination of information security, measures to clearly identify and confirm the parties to the transaction, audit trails and assurance practices that can convincingly demonstrate that a transaction occurred, that its timing is correctly stated, that it is what it purports to be, that the transactor/source is authentic (correctly identified) and is acting within the authorization framework (policy, regulation or law) governing the transaction, and that it was not and could not have been tampered with since it was created.

Non-Duplication: Duplicate records are often used in systems to enhance processing efficiency. However, in many processing contexts, duplicate records are considered to be information integrity impairments; for example, a duplicate payment to a supplier for the same invoice or a duplicate invoice to a customer for the same purchase. Some frameworks address duplicate records as part of completeness. Others treat duplication as a separate element (e.g., under the name uniqueness). In this framework, duplication is considered to be an impairment of validity.

Existence/Occurrence; Ownership, Rights and Obligations: In accounting, assertions such as existence, occurrence, ownership, rights and obligations convey that the information about a subject matter is non-fictitious; i.e., that reported assets exist, that transactions such as revenues and expenses actually occurred, and so forth. Accounting estimates based on forecasts of future events such as cash flows and interest rates may involve measurement uncertainties that may make it difficult to establish existence or occurrence with the same degree of confidence as is possible for current transactions.

Lineage: Validity implies that the information's "lineage" can be traced from the initiation of the information to its current status or ultimate destination and validated. An audit trail consists of the meta-information (i.e., identification and intervening calculations, transformations, aggregations and other steps) required to establish the lineage, hence authorization, compliance, authenticity and non-duplication, of the information.

Neutrality

The concept of information neutrality (objectivity, lack of bias) in the way subject matter is represented is sometimes identified as a separate attribute, especially in the context of financial reporting; however, in this publication, the attribute of neutrality is considered to be subsumed under the attributes of completeness, currency, accuracy and validity.

Exclusions

A number of important issues related to information are excluded in this publication because they do not relate to information integrity. For example, confidentiality and privacy are important information-related concerns but they are not addressed in this publication, although the framework and some of the enablers and controls presented here might be applicable to these issues. Similarly, information quality issues such as relevance and usefulness include concepts that go well beyond information integrity concerns related to its fit for a given purpose, but they are not addressed in detail in this publication which focuses primarily on information integrity attributes, enablers, risks and controls. For example, perceived relevance and perceived ease-of-use may affect how some users perceive information quality and how effectively they use the information, but these aspects of information are separate from its integrity or representational faithfulness.

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 possessing representational faithfulness. In other words, all are necessary, but none are sufficient by themselves to warrant the label. A numerical entry for a credit limit that is not authorized is not representationally faithful; an annual sales figure that omits one month of sales but is otherwise accurate is not representationally faithful; and so forth.

The core attributes of information integrity are threatened by risks in each of the three domains of content, process and IS environment. These risks are mitigated by information integrity enablers and controls corresponding to the risks in each of the three domains. Because of the limits of information processing systems, absolute completeness, currency, accuracy and validity are not achievable. Thus, representational faithfulness is subject to some degree of imperfection, with the tolerable degree of imperfection (the concept of materiality) being defined differently in different domains and contexts. For example, in a financial reporting context, an omission or misstatement is material if it would influence decisions that users make on the basis of the financial information of a specific reporting entity.[6] In contrast, in an integrated reporting context, materiality is used as a screen for including or omitting information about matters that substantively affect the organization's ability to create value over the short, medium or long term.[7]

[6] FASB Statement of Financial Accounting Concepts No. 8, Chapter 1, The Objective of General Purpose Financial Reporting, and Chapter 3, Qualitative Characteristics of Useful Financial Information, 2010 QC11 p. 17..

[7] International Integrated Reporting Council (IIRC). *Assurance on <IR>: an exploration of issues*, 2014, p. 20.

Enablers and controls can help humans and software to assess the degree of representational faithfulness possessed by an information item or information set so that it can be brought within an acceptable range of (im)perfection.

Distinction Between the Categories in this Framework and Traditional Classification of General and Application Controls

The split between General Controls and Application Controls harkens back to the early days of computing when there was a clear distinction between the computing centre and the applications that it ran. The assumption was that a general control was consistently applied to all applications across the board. Thus, for audit efficiency, general controls could be evaluated once and those evaluations could be used with all the applications. The applications of audit significance would have to be evaluated individually and those evaluations would need to be integrated with the evaluation of general controls.

While the distinction between General and Application controls may still be valid in some circumstances, it is less compelling today in IS environments characterized by various types and degrees of outsourcing, reduction of in-house system development in favour of package software and the embedding of controls that previously might have been called general controls within applications or sub-systems; e.g., logical access controls, segregation of incompatible functions, controls over system changes (for example, to application system control tables, user generated reports) and back-up and recovery features.

To be sure, the controls that were previously classified under the category of General Controls are still important; but, they often do not operate uniformly across the enterprise, its various systems andl applications. In a given entity, different systems may be acquired or implemented with different system development lifecycle (SDLC) controls. A system may operate across more than one IS environment with different availability controls in each environment. Different systems may be subject to different security controls even if operating in the same IS environment. Outsourced systems may have different operations controls governed by a variety of service level agreements, and so on. So the traditional approach of evaluating general controls once for the entire entity may no longer be appropriate in many cases. Instead, the manager or auditor should focus on achieving or evaluating control objectives by significant content, process and system without entertaining the illusion that those controls can operate or are operating uniformly across the organization.