

## CHAPTER 2. TIQM® ASSESSMENT PROCESS

### 2-1 Overview

- A. This chapter is concerned with the first stage of HUD's information quality improvement: assessment. Assessment is the first step to achieve expected levels of information quality necessary for HUD to serve its constituents properly.

### 2-2 Select Information Group Candidates

- A. This step provides a guide for Program Areas to use when selecting data elements. With limited time and resources available, it is not feasible to correct every data element in every location, and to analyze and improve every process that produces it. Therefore, a set of criteria for selecting and prioritizing which data elements to assess and improve must be developed. This is the process of determining the scope of a project.<sup>5</sup>

#### B. Determine Scope Based on Business Needs

1. Documenting the scope of effort is necessary to provide direction for information quality improvement and data correction. To obtain the most value from these efforts, it is necessary to assess business needs, taking into consideration the entire information value and cost chain that may be affected by information quality. To determine enterprise-wide business needs accurately, it will be necessary to conduct interviews with knowledge workers in each of HUD's business areas to find out how they and their information stakeholders outside of HUD are using the information and to determine their quality expectations.

#### C. Identify Information Group to be Assessed

1. Once the business needs have been defined, identify the data necessary to support those business needs. Collect information from knowledge worker interviews in each of HUD's business areas to determine how they use the information in the performance of their jobs, and how the data supports the business need. The objective is to determine information where assessment and improvement could yield significant tangible benefits.<sup>6</sup>
2. The following example in Figure 2.1 illustrates how to document the data necessary to support the business needs.

<b>Data Element Scope Worksheet Example</b>			
Information Group	Data Element (Table or Record Name)	Within Scope? (Y/N)	Rationale for Inclusion or Exclusion (process and decision requiring it, and consequences if data is defective)
Inspection	Inspection Date	Y	<ul style="list-style-type: none"> <li>• For selecting inspections within the last fiscal year.</li> <li>• Printed on the report.</li> </ul>
	Inspection Report Completion Date	N	<ul style="list-style-type: none"> <li>• Not applicable to this report or indicator.</li> </ul>

Data Element Scope Worksheet Example			
Information Group	Data Element (Table or Record Name)	Within Scope? (Y/N)	Rationale for Inclusion or Exclusion (process and decision requiring it, and consequences if data is defective)
Human Resource	Inspector First Name, Middle Initial, Last Name	N	• Not applicable to this report or indicator.
Property	Property ID	Y	• To distinguish a particular property in the computer system.
	Property Name	Y	• Printed on the report.
	Property Street Address, City, State, Zip	Y	• Printed on the report.
	Property Contact Phone Number	N	• Not applicable to this report or indicator.
Organization	Regional HUD Office Street Address, City State, Zip	N	• Not applicable to this report or indicator.

**Figure 2.1: Illustration of a Data Element Scope Worksheet**

- D. Identify Information Value and Cost Chain
1. For the Information Groups and Data Elements in the scope of the assessment, determine all business processes, applications, and people who create or update a group of data along with the dependencies between the processes that create or update the same data. This is to identify all points at which information quality can be impacted.
  2. The information value and cost chain includes (source: *Improving Data Warehouse and Business Information Quality*, p. 160-162)
    - a. All database and files, including paper documents, in which the data is stored from its point of origin to its last database of use;
    - b. All business processes and application programs that create, update, or delete data;
    - c. All replicate, extract and interface programs that copy data from one database and transform it and propagate it to another database;
    - d. All information producer and knowledge worker roles in the process;
    - e. All application programs that retrieve the data.
- E. Identify Information Stakeholders
1. For the selected information group, identify the categories of information stakeholders. These stakeholders include
    - a. The information producers (including Program Areas, support offices, other federal agencies, state and local governments, and lenders) that create or maintain the information;
    - b. The knowledge workers who use it, including Program Areas and support offices within HUD;

- c. The end customers, including Program Areas, support offices, the Executive Branch, Congress, GAO, OMB and taxpayers who require it.

For each of these categories, identify key contacts and maintain this information in an authoritative repository for future reuse.

F. Identify Information Quality Objectives and Measures

1. Establish the information quality characteristics to be measured in the information group to be assessed. There are two sets of characteristics: data content and presentation. The data content quality characteristics are described in Figure 2.2.

Characteristic	Quality Characteristic Description	Example of Non-Quality Data
Validity	The degree to which the data conforms to its definition, domain values and business rules.	A U.S. address has a state abbreviation that is not a valid abbreviation (not in the valid state abbreviation list).
Non-Duplication	The degree to which there are no redundant occurrences or records of the same real world object or event.	One applicant has multiple applicant records (evident when an applicant gets duplicate, even conflicting, notices).
Completeness	The degree to which all required data is known. This includes having all required data elements (all facts about the object or event), having all required records, and having all required values.	An indicator for spouse is set to "yes", but spousal data is not present.
Relationship Validity	The degree to which related data conforms to the associative business rules.	A property address shows a Michigan zip code, but a Florida city and state.
Consistency	The degree to which redundant facts are equivalent across two or more databases in which the facts are maintained.	The same applicant is present in two databases or systems and has different name, address, or dependents.
Concurrency	The timing of updates to ensure that duplicate data stored in redundant files is equivalent. This is a measure of the information float (the time elapsed from the initial acquisition of the information in one file or table to the time it is propagated to another file or table).	On Monday, an applicant's change of address is updated in the Applicant record of origin file, but the record is propagated to the main Program database after the weekend cycle (Friday night). That record has a concurrency float of 5 days between the record-of-origin file and the record-of-reference database.
Timeliness	The degree to which data is available to support a given knowledge worker or process when required.	A change of address is needed to schedule an inspection but is not available to the field office, and the inspector leaves without the proper information.
Accurate (to reality)	The degree to which data accurately reflects the real-world object or event being described.	The home telephone number for a customer record does not match the actual telephone number.

Characteristic	Quality Characteristic Description	Example of Non-Quality Data
Accurate (to surrogate source)	The degree to which the data matches the original source of data, such as a form, application, or other document	An applicant's reported income on the application form does not match what is in the database.
Precision	The degree to which data is known to the right level of detail (e.g., the right number of decimal digits to the right of the decimal point).	The summary amounts in congressional reports are rounded to the nearest \$1,000.00 and do not include amounts in the hundreds, tens, dollars or pennies. However, the amounts will be aggregated in dollars and cents and then rounded to the nearest \$1,000 to avoid rounding errors.
Derivation Integrity	The correctness with which derived data is calculated from its base data.	The summary of accounts for a given district does not contain all valid entries for the district.

(Source: *Improving Data Warehouse and Business Information Quality*, p. 142-143)

**Figure 2.2: Characteristics of Information Content Quality**

The presentation quality characteristics are listed in Figure 2.3.

Characteristic	Quality Characteristic Description	Example of Non-Quality Data
Accessibility	A measurement of the degree of ease-of-access interested knowledge workers have to the data they require.	The planning analyst needs the current account of insurance per jurisdiction, but the information is not available unless a programmer extracts it.
Contextual Clarity	The degree to which presentation of the data enables the knowledge worker or end customer to understand the meaning of the data and avoid misinterpretation (intuitiveness).	Applicants report incorrect, or have missing, annual income on the form due to an improper label.
Usability	The degree to which the information presentation is directly and efficiently usable for its purpose.	Statistical information that would be easily understood if presented in a table format is provided in several paragraphs of text.
Rightness	The characteristic of having the right kind of data with the right quality to support a given process.	All the application information is present, but the credit report is missing, so the underwriting process cannot be executed.

(Source: *Improving Data Warehouse and Business Information Quality*, p. 142-143)

**Figure 2.3: Characteristics of Information Presentation Quality**

- G. Determine Files and Processes to Assess
1. Depending upon the assessment objectives, information may need to be measured at different points in the information chain (see Figure 2.4 below). Identify the system(s) that capture, maintain or use the information group, and assess the same information in all applications and files or databases. There is a

tendency to assess the quality only in the circle of influence (the owned application or database); however, the critical impact to the Department occurs when the information is not of the expected quality and is shared across applications and Program Areas. If there are resource or time constraints, it is better to reduce the number of data elements in the assessment but include the entire value chain for the information being assessed. This means that the assessment must include all relevant databases, applications, files, and interfaces. In all cases, the approach to be taken must be defined and documented.

Assessment Objective	Assessment Point
1. Understand state of quality in the database.	The entire database or file. This should be a data source that supports major business processes.
2. Ensure effectiveness of a specific process.	The records output from the processes within a time period being assessed but prior to any corrective actions.
3. Identify data requiring correction.	The entire database or file. This should be a data source that supports major business processes.
4. Identify processes requiring improvement.	The records output from the processes within a time period being assessed, but prior to any corrective actions.
5. Ensure concurrency of data in multiple locations.	A sample of records from the record of origin that must be compared against equivalent records in the downstream database. If data may be created in the downstream database, extract records from both and find the equivalent records in the other.
6. Ensure timeliness of information.	A sample of data at the point of origin. These must be compared against equivalent data from the database from which timely access is required.
7. Ensure effectiveness of data warehouse conditioning process.	A sample of data from the record-of-reference. These must be compared against equivalent record(s) in the data warehouse.

(Source: *Improving Data Warehouse and Business Information Quality*, table 6.2, p. 165)

**Figure 2.4: Information Quality Assessment Point by Assessment Objective**

#### H. Prioritize Data Elements Supporting Business Need

1. Once the data elements necessary to support the business needs have been identified, prioritize the data elements. A simple high-medium-low scale may be used. Knowledge workers who understand how the data meets their requirements best make this determination.
2. As stated in Section 515 Guidelines, "The more important the information, the higher the quality standards to which it should be held." Factors making a data element high priority might be
  - a. Importance to key decision making,
  - b. Internal or external visibility,
  - c. Impact on financial reporting,
  - d. Operational impact of erroneous data (e.g., wasted time or resources).

Figure 2.5 illustrates how to document the data element prioritization:

Data Element Prioritization Worksheet			
Information Group	Data Element (Table or Record Name)	Data Element Priority (High, Medium, Low)	Rationale for Priority (process and decision requiring it, and consequences if data is defective)
Inspection	Inspection Date	High	This is critical to determine if the inspection had been performed within a year.
	Inspection Rating	High	This is critical to determine if the property passed inspection.
	Inspection Comments	Low	Not critical for this report.
Property	Property ID	High	This is the identifier of the property information in the computer system.
	Property Name	Medium	This is an important characteristic of the property but is not indispensable for the report.
	Property Street Address, City, State, Zip	Medium	This is an important characteristic of the property but is not indispensable for the report.
	Property Owner First Name, Last Name	High	This is a determinant characteristic of ownership. It is required to assess if proper practices are in place.
	Property Owner Middle Initial	Low	Not critical for this report.

Figure 2.5: Illustration of a Data Element Prioritization Worksheet

### 2-3 Assess Data Definition and Information Architecture Quality

- A. In this step, determine the quality measures for data definition and information architecture. Also, evaluate the structure and definition of the information under assessment. Finally, develop or refine definitions and structures that are missing or have defective definition or structure. In the case of defective definition or structure, recommend improvement in the data development and maintenance processes that created the defective definitions and architectures.<sup>7</sup>
- B. Identify Data Definition Quality Measures
  1. Identify and, if necessary, define the essential and critical quality characteristics of data definition and information architecture. These are the minimum HUD requirements for
    - a. Data names, including business term, abbreviated name, database or file name, standard screen name and standard report name;
    - b. Definition;
    - c. Valid value set (including value definitions) or reasonable range of values;
    - d. Business rules for data integrity;
    - e. Physical specifications of the data.

These quality characteristics must be in place for ensuring effective communication among information producers and consumers or knowledge workers, and data resource management and application development personnel.<sup>8</sup>

- C. Assess Data Definition Technical Quality
1. Assess the data definition for conformance to data standards and guidelines. Determine whether the data definition conforms to the minimum established standards.<sup>9</sup> Obtain a comprehensive and concise definition for each data element in the information group. This definition must contain an agreed-upon statement or rule about the data content of the data element and its representation, the business rules that govern its data integrity, and the expected quality level based on the entire value chain of the data element.
  2. The data element definition must be known, documented, understood, communicated, and validated by all business areas in the value chain. It should be documented for each record of origin and other metadata for each data element or groups of data elements as described in the illustration in Figure 2.6 below.
- D. Assess Information Architecture (IA) and Database Design Quality
1. Assess the information architecture (or logical data model), the database design (implementation model), and the physical implementation of the data structures against modeling, design, and implementation of best practices, in accordance with three kinds of IA assessments.<sup>10</sup> An *Information Architecture Completeness Assessment* will determine whether the data model has all required entity types and attributes to support the business processes. An *Information Architecture Correctness Assessment* determines whether the data model truly reflects the real world entity types, attributes, and relationships. Finally, an *Information "Chaos" Assessment* determines which instances of data redundancy in Program Area proprietary files, application software package files, or other storage mechanisms are controlled, and which are not controlled.
- E. Assess Customer Satisfaction with Data Definition and Information Architecture
1. Measure customer satisfaction with the definition of the information products based on the knowledge workers' assessment. The deficiencies discovered in this step are critical input to the process discussed in Section 3-2. In this case, the processes that can be improved are the data definition and application development processes.
- F. Develop or Improve Data Definitions
1. In cases where the data to be assessed in the subsequent steps lacks or has defective definition and /or information architecture, develop correct definitions and/or information architecture to ensure that subsequent tasks can be executed. Interact with representatives of the business and IT areas across the value and cost chain to arrive at appropriate definitions and architecture (see Section 2-2(E) for a discussion on the identification of the stakeholders).<sup>11</sup>
  2. To achieve a new or revised definition, first develop the necessary common terms and business concepts and then use them to define the entities, data elements and relationships. The terms, entities, data attributes and relationships

must be coordinated and validated by the stakeholders across their value and cost chains.

3. The template for a data definition as a Business Concept is given in Figure 2.6 below.

Business Concept	Name of the business object to be defined.
Status	[Status] as of x/x/02 [the date of the last revision]
Definition	A succinct, complete business description using common terms or previously defined terms – if new terms are needed, add them to the Business Concept Control Chart and reference in the next section.
Related Business Concepts	<ul style="list-style-type: none"> <li>• A short-cut or hot-link to a previously defined term.</li> <li>• A description or annotation of an undefined term.</li> </ul>
Data Integrity Rules	A bulleted list of data integrity rules (things that “can” or “must” be present for accuracy, completeness).
Unique Identifier	Free form text, bullets, or whatever can be used to provide insights on the appropriate way to uniquely identify each instance of the data entities associated with the business concept.
Life Cycle	Free form text, bullets, or whatever can be used to provide insights on the states and transitions associated with the business concept.
Classification	Free form text, bullets, or whatever can be used to provide insights on the intrinsic classification necessary to understand the business concept.
Domain	<ul style="list-style-type: none"> <li>• A bulleted list of all applicable domain values. It must be all-inclusive, however if it is not known at the time, a sample can be used while the definition is in process; indicate which approach is used.</li> <li>• Can be a diagram.</li> </ul>
Special Usage	Free form text, bullets, or whatever can be used to highlight real-life cases of the use of the business concept.
Examples	Free form text, bullets, or whatever can be used to provide real-life or mock-up illustrations of the business concept
Issues & Concerns	<p>Proposed definition concerns:</p> <ul style="list-style-type: none"> <li>• A bulleted list of issues or concerns associated with the definition proposed in this document.</li> </ul> <p>Existing definition concerns:</p> <ul style="list-style-type: none"> <li>• A bulleted list of issues or concerns associated with the existing definition, or definitions.</li> </ul>
Background Documentation	Free form text and bullets as needed. For specific references to existing document, use the name and date of the document.

**Figure 2.6: Illustration of the Business Concept Template**

G. Improve Data Development Process

1. If there is a pattern of missing or unsatisfactory data definitions that would be required in order to implement effective edit and validation routines, or if data is defined with multiple meanings (overloaded data), then the data development and/or data maintenance processes are probably broken. If so, recommend a process improvement initiative to improve the defective process (see Chapter 3).

This improvement must be done prior to the next project requiring new data to be defined and implemented.

#### 2-4 Determine Desired Quality Standards for Prioritized Data Elements

- A. Perform in-depth analysis upon the data elements that are within scope. This analysis should identify the information value and cost chain of each data element or group of data elements, describe the element's quality characteristics, and determine the element's quality standard for each quality characteristic.<sup>12</sup>
- B. Define Information Value and Cost Chain for Data Element(s)
1. In order to assess data adequately, identify all record(s) of origin for the data. Currently, there are cases where data elements entered in an initial database are updated in a second or even a third system. In cases where redundant data is identified, it must be corrected in every database in which it is stored.
  2. The following template (Figure 2.7 below) should be completed for each system identified as a record of origin.

Data Element By Record of Origin System						
Data Element Business Name	Record of Origin System Name	Physical Data Element Name	Definition	Field Type	Length	Create/Update
Inspection Date	DQ1	LAST-INSPECTED	The date the most recent property inspection took place.	Numeric	8	Create, Update
Inspection Rating	DQ1	INSPECTION-RATING	A classification indicating the relative condition of the property at the time of the inspection.	Numeric	3	Create, Update
Property Owner First Name, Last Name	DQ1	OWNER-NAME	The First and Last Name of the person registered as legal owner of the property.	Alpha-numeric	40	Create, Update
	DQ2	OWNER-FORMAL-NAME	The fully formatted name of the owner. In the case of individuals it is the combination of the First, Middle and Last Name. In the case of organizations it is the legal name.	Alpha-numeric	50	Update

Figure 2.7: Illustration of a Data Element By Record of Origin Worksheet

Once the record of origin has been identified, determine where other read-only versions of the data are located in the organization. Strategies can then be formulated regarding the data validation and correction of those data sites.

- C. Identify Accuracy Verification Sources
1. In order to verify the accuracy of a data element value, it is vital to identify the most authoritative source from both surrogate and real-world sources from which to assess and confirm the accuracy or correctness of the data value. The most

accurate of these is the real-world source, since the surrogate sources have the potential to contain errors.

- D. Determine Applicable Quality Standard for Each Data Element
1. Determine and document the criteria for information quality according to the quality criteria discussed in Section 2-2(F). In setting the desired level of information quality, all criteria should be assessed as to relevance and level of importance. Information that meets the criteria are considered quality; those data elements that do not meet the criteria are considered "defective" and must be corrected or discarded. The example in Figure 2.7 illustrates how to state data element quality criteria.
  2. Accuracy criteria descriptions must explicitly name the data validity source that is the basis of the data in the record of origin. If no data validity source is available, then the method of determining accuracy must be described.
- E. Determine Quality Standards (Compliance Levels)
1. Having defined the quality criteria for each data element, determine what percentage of the data must comply with the specifications. This percentage will be the measuring stick referenced when an organization performs an internal quality assessment. Additionally, the assessment team will use the compliance levels when auditing.
  2. The compliance percentage should be stated for each criteria specification. For example, a 100% Validity compliance target means that no data can deviate from the validity criteria. A 98% Accuracy level means that at least 98% of the data must meet the Accuracy criteria. If a data element meets *all* stated quality compliance targets, then the data element passes and is categorized as "quality compliant."
  3. Figure 2.8 is an example for documenting the data element compliance targets as well as data exceptions.

Data Element Quality Criteria Worksheet for Record of Origin System: <i>DQIS (P150)</i>												
Information Group / Data Element		Validity	Not Blank	Completeness	Relationship Validity	Consistency	Concurrency	Timeliness	Accurate (to reality)	Precision	Accurate (to surrogate source)	Derivation Integrity
Inspection Date	<b>Quality Criteria</b>	Can be blank (not inspected) or a valid date since 1922.	-	-	With Inspection Rating: both blank or both not blank.	-	-	-	-	-	Must match the inspection date on the inspector's log.	-
	<b>Compliance level</b>	100%	-	-	99%	-	-	-	-	-	99.5%	-
	<b>Exceptions</b>	-	-	-	-	-	-	-	-	-	-	-
	<b>Findings</b>	97% compliant; 3% are before 1922	-	-	4% missing when rating present.	-	-	-	-	-	5% did not match inspector's log.	-
Inspection Rating	<b>Quality Criteria</b>	Can be blank or numeric.	-	-	With Inspection Date: both blank or both not blank.	-	-	-	-	-	Must match the inspection rating on the inspector's log.	-
	<b>Compliance level</b>	95%	-	-	100%	-	-	-	-	-	100% for 1996 and later.	-
	<b>Exceptions</b>	-	-	-	-	-	-	-	-	-	Include only 1997 to current date.	-
	<b>Findings</b>	100% compliant.	-	-	4% present when date missing.	-	-	-	-	-	1% did not match inspector's log.	-
Property Owner's Name (First & Last)	<b>Quality Criteria</b>	Not blank. No special characters except hyphen, comma or period.	-	-	-	-	Must reflect changes received by HUD within 10 working days.	-	-	-	Must match owner's name in local authority's document of the assistance contract.	-
	<b>Compliance level</b>	99%	-	-	100%	-	-	-	-	-	98%	-
	<b>Exceptions</b>	-	-	-	-	-	-	-	-	-	-	-
	<b>Findings</b>	87% compliant; 13% are blank.	-	-	-	-	Current process takes up to 45 days to verify official change before updating the system.	-	-	-	Due to resource limitations, verified only lowest inspection ratings; found 22% names misspelled.	-

Figure 2.8: Illustration of Quality Target Compliance

## 2-5 Assess Current Level of Information Quality

A. A vital step in information quality improvement is to assess the *current level* of information quality. When selecting the data records for the assessment, acquire a representative, or statistically valid, sample to ensure the assessment of the sample accurately reflects the state of the total data population, while minimizing the cost of the assessment. To be a statistically valid sample, "every record within the target population has an equal likelihood of being selected with equal probability."<sup>13</sup> When properly conducted, a random sample of records provides an accurate picture of the overall information quality of the database.<sup>14</sup> In certain circumstances, purposive samples may be usefully substituted for random samples. For example, if only active cases are of interest, the sample may include all active cases.

## B. Extract Random Sample of Data

1. Select a sample size for each point of measure (database, file, transaction flow) based on the following formula:

$$n = ((z \times s) / B)^2$$

where:

**n = the number of records to extract.**

**z = a constant representing the desired confidence level** (e.g., the degree of certainty, expressed as a percentage, of being sure about the estimate of the mean).

**s = an estimate of the standard deviation of the data population being measured** (e.g., the degree of variation of errors within the data population).

**B = the bound or the precision of the measurement** (e.g., the variation from the sample mean within which the mean of the total data population is expected to fall given the sample size, confidence level, or standard deviation).

2. Next, generate the random samples. It is critical that data be sampled in a way that ensures the data is unchanged from the point at which the sample is extracted.
3. Provide a sampling report indicating the sample size for each information group, the total universe of data and the portion to be sampled, the sampling criteria and an explanation of the reason why this will produce a statistically valid sample.<sup>15</sup>

## C. Measure Information Quality

1. Analyze the information in the samples against its target criteria based on the data definition and information architecture (Section 2-3(F)) as well as the defined quality standards (Sections 2-4(D) and 2-4(E)). The assessment should be performed against the established specifications, compliance targets, and data exceptions. Different data elements may require different assessment techniques for the various criteria. For each information group, either automated or physical data assessments -- or both -- are performed.

2. For accuracy assessment or certification, the authoritative source for the data element must be specified. This may be a hard copy document, a physical inspection of the real object or event the data represents (or a review of a recording of an event), or information from an external source considered accurate, or an official document (such as a certified land survey) that is considered to be accurate.

D. Validate and Refine Data Definitions

1. The data definitions and architectures may be adjusted based on facts discovered during the measurement process. In such cases, the assessment team may execute the information architecture and data definition process described in Section 2-3(F) to arrive at the appropriate revised definitions.

E. Establish Statistical Control

1. For processes that acquire, produce or maintain mission-critical information, it is imperative that they be in a state of statistical control. That is, with respect to the mission-critical information they produce, their results are predictable and the quality of the information is in line with the initial agreed upon levels. Examples of how and when statistical process control of data quality is performed will be available on the EDMG team web site.

## 2-6 Measure Non-Quality Information Costs

- A. The objective of this step is to identify the cost of non-quality information for the information groups or data elements under assessment. Non-quality information costs are assessed in three areas: process failure costs, information scrap and rework costs, and lost or missed opportunity costs.<sup>16</sup>

1. **Process failure costs** – the result of a process, such as distribution of funds, which cannot be accomplished due to missing, inaccurate, incomplete, invalid, or otherwise non-quality information.
2. **Information scrap and rework costs** – incurred when a knowledge worker has to waste time handling or reconciling redundant data, hunting for missing information, verifying data, or working around broken processes.
3. **Lost or missed opportunity costs** – HUD may be missing out on opportunities to greatly improve the lifestyle of communities due to non-quality information, or may be directing funds toward areas of lesser need.

B. Identify Business Performance Measures

1. Information has value to the extent that it enables the enterprise to accomplish its mission or business objectives. In order to determine if a process or information set adds value to the organization, it is important to understand
  - a. The business vision and mission,
  - b. The business plans and strategies,
  - c. The strategic business objectives.

Business performance measures are identified directly from goals set forth in HUD's APP.

C. Calculate Information Costs

1. Identify what percent of information systems and data development and maintenance is value-adding and what percent is cost-adding, performed solely to improve information systems productivity and effectiveness. Determine the relative costs of application and data development in the three categories of information development expense:
  - a. **Infrastructure:** These are the costs of developing and reusing databases and applications that create and maintain the information resources.
  - b. **Value-adding:** These are the costs associated with the access and retrieval of information for use to add value.
  - c. **Cost-adding:** These are the costs of redundancy of both application and data development and maintenance.

D. Calculate Non-Quality Information Direct Costs

1. Quantify the cost of non-quality information to determine the business impact of information quality problems, raise awareness of the importance of information management, and establish a benchmark for measuring information quality initiatives. This process identifies the categories of costs of poor-quality information and calculates the costs of information scrap and rework.

E. Measure Lost Opportunity Costs and Information Value

1. Identify the indirect costs derived from lost or missed opportunities. These can be revenue not realized, missed service to the public, or the costs of underpaying or overpaying for the services rendered. These costs are usually very large and difficult to estimate.

2-7 Interpret and Report Information Quality State

- A. Once data has been assessed, the results will be analyzed, interpreted and clearly presented in a format easily understood by knowledge workers, information producers, and process owners. The results will include all assessments of all components (definition, content and presentation). Also, the results will describe findings and recommendations in the quality standards (expected levels of quality), actual quality levels, and information costs, especially non-quality information costs. Each report will include a cover sheet, a summary and detail section, and an assessment procedure report for each information group (see example Figure 2.9 below).<sup>17</sup>
- B. The quality assessment summary and detail sections will describe the current level of information quality and then will make recommendations for succeeding project(s). These recommendations will address approaches for correcting the data errors identified and for changes to systems, procedures, training, and technology that will help to ensure the appropriate level of quality for the data. The Final Information Quality Assessment Report will describe the following:
  1. A description of the approach, such as:
    - a. The Source system(s) included.
    - b. The Assessment criteria.
    - c. The Specific HUD participants.
  2. Current level of information quality:
    - a. General conclusions about the information quality of assessed elements.
    - b. Information quality defects found.

- c. The Assessment results (number and types of errors found, level of confidence in the results, any other issues).
3. Recommendations to close the gap between current information quality levels and target information quality standards:
  - a. For Data Corrections, indicating appropriate approaches for the errors identified.
  - b. For Information Quality Improvements, identifying types of changes that may be made to systems, procedures, or technology to ensure the appropriate level of data quality.
  - c. High priority tasks or areas of concern to be addressed first.
  - d. Start date.

<b>ASSESSMENT PROCEDURE REPORT</b>	
<b>Information Group Name:</b> _____	<b>Date Assessed:</b> _____
<b>Time Period Covered From:</b> _____	<b>To:</b> _____ <b>IQ Analyst:</b> _____
<b>File(s) Sampled:</b> _____	<b>Processes Sampled:</b> _____
<b>Sampling Procedure:</b> <input type="checkbox"/> Representative Random Sample	
<input type="checkbox"/> Purposive Selection Sample (Purpose): _____	
<b>Sample Size:</b> _____	<b>Sample Percent:</b> _____ % <b>Standard Deviation:</b> _____
<b>Confidence Level:</b> <input type="checkbox"/> 90% <input type="checkbox"/> 95% <input type="checkbox"/> 99%	<b>Precision:</b> _____
<b>Assessment type:</b>	<input type="checkbox"/> Electronic <input type="checkbox"/> Third-party corroboration
	<input type="checkbox"/> Physical to surrogate source <input type="checkbox"/> Survey
	<input type="checkbox"/> Physical to real object / event
	<input type="checkbox"/> Other: _____
<b>Quality Characteristics Assessed:</b>	
<input type="checkbox"/> Completeness of values	
<input type="checkbox"/> Reasonability tests / distribution analysis	
<input type="checkbox"/> Validity: conformance to business rules	
<input type="checkbox"/> Accuracy: correctness of values to: <input type="checkbox"/> Source: _____	
<input type="checkbox"/> Surrogate: _____	
<input type="checkbox"/> Non-duplication of records	
<input type="checkbox"/> Timeliness of data availability	
<input type="checkbox"/> Equivalence and consistency of redundant data	
<input type="checkbox"/> Usefulness and value-adding	
<small>Source: <i>Improving Data Warehouse and Business Information Quality</i>, Figure 6.9, 190, adapted for HUD</small>	

**Figure 2.9: Illustration of an Assessment Procedure Report Template**

[THIS PAGE INTENTIONALLY LEFT BLANK]