DEVELOPING, IMPLEMENTING AND MONITORING AN INFORMATION PRODUCT QUALITY STRATEGY

(Practice-Oriented Paper)

Elizabeth M. Pierce
Indiana University of Pennsylvania
Eberly School of Business and Information Technology
empierce@iup.edu

Abstract The purpose of this paper is twofold. First it outlines the basic methodology for improving the design and manufacture of information products through the development, implementation, and monitoring of an information product quality strategy. Secondly, it shows how existing marketing techniques can be used to construct an information product quality strategy through the application of customer profiling and prioritization of information quality specifications and expectations.

Key Words: Information Product, Data Quality, Conjoint Analysis

1 INTRODUCTION & BACKGROUND

1.1 The Knowledge – Information – Data Hierarchy

To succeed in today's global, competitive environment, companies must achieve excellence in all their business processes. A key factor in achieving this excellence is whether individuals responsible for carrying out a business process have access to the knowledge they need for making sound decisions and properly conducting transactions. This knowledge must be conveyed and stored in some physical format, i.e. an information product. An information product is comprised of data items and it is the quality of each data item that is of importance to the individual using the information product. The quality requirements for the information product as a whole drive the requirements of the raw source data items and the semi-processed component data items needed to create the information product whose ultimate purpose is to convey and retain some aspect of the organization's knowledge base. To illustrate, consider the case of a retailer whose sales force requires knowledge about a customer's desires in order to complete a sale. The need to convey and store that knowledge compels the retailer to define a customer sales order, an information product composed of specified source data about the customer and his requested product(s), and service(s) that must be processed and arranged into an agreed upon form.

Ilkka Tuomi [1] observed that it is knowledge that drives information and in turn, it is information that drives data. His concept of a knowledge, information, and data hierarchy is illuminating because it implies that data and information quality must also be defined accordingly. A company must first identify the kind and quality of knowledge that it needs to conduct day-to-day operations and to make decisions.

Only then can the firm adequately specify the information products along with their quality criteria that will make it possible for the firm to retain and convey this knowledge. Once the quality criteria for the information is well understood, companies can then proceed to make sound decisions about how to model, represent, and process the raw data to meet the necessary standards. Thus in the example of a sales order, a retailer cannot adequately specify its data quality criteria for the different data components that make up the sales order until it fully comprehends the level of quality required for the sales order as a whole and that in turns depends on the quality expectations for the knowledge needed by the sales people to complete a sale to a customer.

1.2 What does it mean to treat information as a product?

In describing his knowledge-information-data hierarchy, Tuomi relies on information systems as the focus for improving knowledge management and organizational memory. Since the early 1990's however, individuals like Richard Wang, Yang Lee, Leo Pipino, and Diane Strong have advocated that organizations treat their information as an end-deliverable that satisfies consumer needs, i.e. an information product, rather than as some by-product of a computer system [2]. Under the information product-centric approach, information systems are still vital but their importance stems from the role that quality information systems play in creating quality information products. Wang et al.'s approach recommends that organizations adopt four principles for managing information as a product. [2]

(1) Organizations must understand their consumers' information needs
(2) Organizations must manage information as the product of a well-defined information process that incorporates technology as well as organizational behavioral factors.
(3) Organizations must manage the life cycle of their information products.
(4) Organizations should appoint an information product manager (IPM) to manage their information processes and resulting products.

1.3 Applying Manufacturing Principles to Information Product Improvement

Like a manufacturing system that uses an assembly line to convert raw materials into physical products, an information system can be viewed as an information manufacturing system that converts raw data into information products. The information product paradigm allows for proven principles from manufacturing to be applied to the improvement of the design, development, manufacture, and distribution of information products. For example, the product manufacturing field has an extensive body of Total Quality Management (TQM) literature with principles, guidelines, and techniques for improving product quality. These classical TQM principles can be adapted to the improvement of information products through the undertaking of five tasks [3].

(1) Articulate an information quality vision in business terms
(2) Establish central responsibility for information quality through the information product manager.
(3) Educate information product suppliers, manufacturers, and consumers
(4) Teach new information quality skills based on the cycle of defining information quality dimensions, measuring information quality metrics, analyzing root causes for information quality problems, and improving information products.
(5) Institutionalize continuous information quality improvement

1.4 Applying Marketing Techniques to Information Product Improvement

Like any physical product, an information product's quality is determined by two sets of manufacturing-related activities: Design & Development followed by Production & Distribution. While the TQM literature provides much assistance in improving the manufacture of information products, the marketing literature provides additional help in identifying customer-based quality requirements for information products as well as the relative importance of these requirements. These customer-based quality criteria must be balanced with other information product design constraints such as cost considerations, regulatory requirements, privacy restrictions, and standards such as uniform data formats in order to achieve a good information product.*

It should also be noted that for any given business process, a single information product is probably insufficient to meet the knowledge needs of all participants in the process. Just like Ford Motor Company must design a variety of vehicle lines each consisting of multiple brands and models with many customizing features and options to satisfy its customers' diverse set of transportation needs, companies must design entire information product lines that can be customized to meet the knowledge needs of its employees, customers, and vendors. Again, marketing techniques can be employed to aid information product managers in understanding the necessary quality requirements and in designing and developing a complete line of information products to support their organization's knowledge management and memory needs.

The remainder of this paper outlines how a firm can take advantage of existing manufacturing and marketing techniques in order to develop, implement, and monitor an information product (IP) quality strategy for the purpose of improving their information products so they are better able to fulfill the organization's knowledge management and memory needs. The marketing methodology used in this paper to create the information product quality strategy is adapted from Martin Christopher's writings on improving customer service through coordination with an organization's logistics and supply chain management efforts [5] [6].

2. DEVELOP THE CUSTOMER-BASED IP QUALITY STRATEGY

2.1 Create an Information Quality Mission Statement

Before one can begin developing an information product quality strategy, it is important to express an overall vision of what role information quality should play in the culture and shared values of the organization, followed by a clear definition of the precise information quality objectives, especially as they relate to the overall quality mix of the organization. The information product quality strategy should then document the mechanisms by which these objectives will be achieved for a specified information product or product line. Besides expressing the vision, the culture, and the shared values of the organization as well as defining the scope of the business, this mission statement should provide guidance and direction to individuals in the firm as to the actions they should take. In particular, the mission statement should explain how information quality fits into the company's intended strategic positioning. Defining the mission statement requires the involvement of the firm's top management with input from those responsible for implementing the programs that will support the stated mission. The mission

* It should be noted that other fields like graphics design, information systems auditing, communications media, publishing, and even specialty fields like designing websites for the blind contribute many useful ideas for improving the overall design of an information product.

statement should be reviewed on a regular basis since a company's mission statement must be adjusted to reflect changes in the competitive environment. [5] [6]

Although few examples of mission statements incorporating information quality exist, Albrecht and Zemke have summarized the prerequisites for a meaningful service mission statement which can be applied to the requirements for a successful information quality mission statement [4].

(1) It is nontrivial; it has weight. It must be more than simply a "motherhood" statement or slogan. It must be reasonably concrete and action-oriented.
(2) It must convey a concept or a mission which people in the organization can understand, relate to, and somehow put into action.
(3) It must offer or relate to a critical benefit premise that is important to the customer. It must focus on something the customer is willing to pay for.
(4) It must differentiate the organization in some meaningful way from its competitors in the eyes of the customer.
(5) If, at all possible, it should be simple, unitary, easy to put into words, and easy to explain to the customer.

2.2 Identify the Customers and Dimensions of Information Product Quality

As part of the development of an effective information product quality strategy, the information product manager must first identify the customers for the information product in question as well as obtain a list of data quality dimensions that are important to these customers.[†] This initial task, identifying those customers who are the primary sources of influence upon the information quality specifications and expectations, is crucial and often difficult. An information product may have customers both internal and external to the firm. In addition, an individual who receives an information product may only be acting as an agent for others. The information product manager will need to talk to both business and systems personnel to help discover exactly who uses a given information product. In some cases, the information product manager may even consider withholding the information product for a time to see if anyone complains. If not, perhaps this is an obsolete information product that should be retired.

Once a representative set of information product consumers has been identified, the information product manager can use either interviews or small questionnaires to elicit the importance they attach to the quality of the information product as a whole as well as the specific importance they attach to the quality of the individual data components that make up the information product. A 1996 study by Wang and Strong [7] provides a hierarchical listing of 15 data quality dimensions which are grouped into four major data quality categories: intrinsic, contextual, representational, and accessibility. Information product managers will find this list of data quality dimensions a useful starting point in engaging consumers in a discussion of the quality of their information product and its data components.

It is important to keep in mind that the purpose of this stage is to ensure that relevant and meaningful measures of data and information quality are generated by the information product consumers themselves. Once these data and information quality dimensions are defined, the information product manager can proceed to the next step of determining the relative importance of each one and the extent to which

[†] Although this paper mainly refers to developing an information product quality strategy for a single information product, the information product quality strategy could be developed for an entire information product line, especially when the information products share a similar set of customers, data items, and quality issues.

different types of information product consumers are prepared to trade-off one aspect of data quality for another [5] [6].

2.3 Establish the Relative Importance of these Information Quality Dimensions

Because customers will vary in their views as to which data components and quality dimensions are most important in determining a meaningful information product, the information product manager must formulate the information product quality objectives depending on the customers' answers to the following questions [5] [6]:

- From a customer's viewpoint, how important is a particular quality dimension compared to the other quality dimensions in the information quality mix? (Ex. For an information product consumer, how does the "time to access" the information product compare to the "ease of access"?)
- Which data components and their associated dimensions of data quality contribute most to the overall customer satisfaction with the information product?
- Which dimensions of data and information quality are seen as priorities by customers when they make their choice of information product supplier?

To find out the answers to these questions, the information product manager can employ several techniques:

Rank Technique: This technique works by asking a representative sample of consumers to rank order the list of data quality dimensions generated by the prior step from the "most important" to the "least important". The disadvantage of this technique is that it works poorly in situations where there are a large number of quality dimensions and data components for the consumer to consider. It also does not give the information product manager a sense of the relative importance of the data quality dimensions for the different data components that make up the information product [5] [6].

Weight Technique: This technique incorporates a rating scale. For instance, customers could be asked to place a weight from 1 to 10 against each data component's data quality dimensions according to how much importance they attach to each element. The disadvantage of this technique is that most respondents will tend to rate most of the data components' quality dimensions as highly important, especially since those data quality dimensions were originally generated on the basis of their significance to the customers [5] [6].

Trade off Technique: Often referred to as conjoint analysis in the marketing research field, this technique is based on the concept of trade-offs. To conduct this analysis, it is necessary to specify a list of levels or options for each information quality dimension under consideration. For example, for an information quality attribute like reputation which indicates the extent to which data are trusted or highly regarded in terms of their source or content, one might choose values such as "Highly Respected", "Average Respect", "Little Respect", and "No Respect". To find out how much information product consumers prefer an information product that possesses a certain combination of quality attribute options, one can use a variety of survey methods as listed below [16]:

- Self-explicated approach: Respondents rank the levels for each information quality attribute from most to least preferred and then assign an importance weight to each attribute. The self-explicated context puts emphasis on evaluating information products in a systematic, quality dimension-by-quality dimension manner, rather than judging the information product as a whole or in a competitive context.

- Traditional conjoint comparisons: Respondents are presented with two or more different quality profiles for an information product and are asked to rank (or rate) the different quality profiles according to preference.
- Choice-based conjoint analysis: Respondents are presented with two or more quality profiles for an information product and are asked to choose which quality profile they would use. Often this approach includes the option of choosing none of the above.

Regardless of the chosen method, the respondent's preferences can be examined individually, pooled (or averaged) into a single utility function, or broken into segments of respondents who have similar preferences. Statistical software written specifically for conjoint analysis can then be used to analyze the collected data.

To illustrate the basics of conjoint analysis, consider the following example adapted from Joseph Curry [8]. Suppose the Accounting department wants to learn more about its customers' quality preferences for the invoices that it produces. From experience, Accounting personnel know that their customers care mainly about three important quality features for this particular information product:

- Accuracy of the invoice
- Time needed to produce the invoice
- Cost to provide the invoice

Accounting personnel further know that there is a range of feasible alternatives for each of these features:

Accuracy of the Invoice	Time to produce the invoice	Cost to provide the invoice
100%	1 day	$0.65
90%	3 days	$1.25
80%	5 days	$2.50

Obviously, the ideal invoice would be 100% accurate, available in a single day, and only cost $0.65 to produce. However, it may not be possible to deliver an invoice that meets all those criteria. The Accounting personnel will have to trade-off one information quality dimension against another, but what is the appropriate trade-off? To find out, the manager of the Accounting department ranks his pair-wise conjoint comparisons of available trade-off choices as shown in Tables 1, 2, and 3.

Table 1: Accounting Manager's Rankings		**Time to produce invoice**		
		1 day	**3 days**	**5 days**
Accuracy Of the Invoice	**100%**	1	2	4
	90%	3	5	6
	80%	7	8	9

Table 2: Accounting Manager's Rankings		**Time to produce invoice**		
		1 day	**3 days**	**5 days**
Cost to Produce the Invoice	**$0.65**	1	4	7
	$1.25	2	5	8
	$2.50	3	6	9

Table 3: Accounting Manger's Rankings		**Cost to produce the invoice**		
		$0.65	**$1.25**	**$2.50**
Accuracy Of the	**100%**	1	2	3
	90%	4	5	6

Invoice	**80%**	7	8	9

Conjoint analysis figures out a set of values known as utilities for the different quality levels of invoice accuracy, availability, and cost so that when these values are added together for an invoice they reproduce the Accounting manager's rank orders. These utility values can then be applied to an invoice quality combination to get an estimate of the Accounting manager's preferences. These are several techniques that can be employed to estimate these utilities. One simple way to obtain these preference values is to treat the conjoint analysis as a multiple regression problem [17]. The respondent's ratings for the information product quality dimensions form the dependent variable. The quality characteristics of the information product (the attribute levels) are the independent (predictor) variables. The estimated betas associated with the independent variables are the utilities (preference scores) for the levels. The R-Square for the regression characterizes the internal consistency of the respondent. To obtain stable estimates of respondent utilities, most good conjoint studies collect 1.5 to 3 times more observations than parameters to be estimated.

Table 4: Accounting Manager's Utilities		**Time to produce invoice**		
		1 day – 50	**3 days – 25**	**5 days – 0**
Accuracy Of the Invoice	**100% - 100**	150 (1)	125 (2)	100 (4)
	90% - 60	110 (3)	85 (5)	60 (6)
	80% - 0	50 (7)	25 (8)	0 (9)

Table 5: Accounting Manager's Utilities		**Time to produce invoice**		
		1 day – 50	**3 days - 25**	**5 days – 0**
Cost to Produce the Invoice	**$0.65 – 20**	70 (1)	45 (4)	20 (7)
	$1.25 – 5	55 (2)	30 (5)	5 (8)
	$2.50 – 0	50 (3)	25 (6)	0 (9)

Table 6: Accounting Manger's Utilities		**Cost to produce the invoice**		
		$0.65 – 20	**$1.25 - 5**	**$2.50 – 0**
Accuracy Of the Invoice	**100% - 100**	120 (1)	105 (2)	100 (3)
	90% - 60	80 (4)	65 (5)	60 (6)
	80% - 0	20 (7)	5 (8)	0 (9)

Tables 4, 5, and 6 show one possible scheme for modeling utilities so that when these values are added together they reproduce the Accounting manager's rank orders for the invoice's quality dimensions. Notice that there is some arbitrariness in the magnitudes of these numbers even though their relationships to each other are fixed. Once a complete set of utility values that capture a customer's trade-offs are computed, they can be used to estimate a customer's preferences for an information product's quality mix (See Table 7). By applying the derived utilities in this example to two types of quality profiles for invoices that the Accounting department is considering, the utility results for each quality profile are obtained (See Table 8). Based on these results, one would expect the manager of the Accounting department to prefer Invoice 2 over Invoice 1 because it has the larger total utility value.

Table 7: Summary of Utilities for the Manager of the Accounting Department		
Accuracy of the Invoice	**Time to produce the invoice**	**Cost to provide the invoice**
100% = 100 Utilities	1 day = 50 Utilities	$0.65 = 20 Utilities
90% = 60 Utilities	3 days = 25 Utilities	$1.25 = 5 Utilities
80% = 0 Utilities	5 days = 0 Utilities	$2.50 = 0 Utilities

Table 8: Comparison of two quality strategies for an invoice

Accounting Manager	Invoice #1	Invoice #2
Accuracy of Invoice	100% = 100 Utilities	90% = 60 Utilities
Time to produce the Invoice	5 days = 0 Utilities	1 day = 50 Utilities
Cost to produce the Invoice	$1.25 = 5 Utilities	$2.50 = 0 Utilities
Total Utility	105 Utilities	110 Utilities

It is important to keep in mind that these particular results only reflect the preferences of the Accounting department's manager. The Accounting department should collect the responses from other individuals who have a stake in the quality of the invoices to see how their preferences compare to those of the Accounting department's manager.

2.4 Determine if "Customer Clusters" of Information Quality Preferences Exist

After the information product manager has determined the importance attached by different customers to the information product's overall quality mix, the next step is to see if any similarities of preference emerge. If one group of customers have a clearly distinct set of quality priorities then another then it would be reasonable to think of each of them as different information quality segments [5] [6]. One tool that is helpful for identifying customer segments is cluster analysis. Cluster analysis is an exploratory data analysis tool for solving classification problems. Its objective is to sort cases (people, things, events, etc) into groups, or clusters, so that the degree of association is strong between members of the same cluster and weak between members of different clusters [9]. The classification of each identified cluster is found by interpreting the similar values of the attributes shared by its members. Due to the iterative nature of the clustering algorithms, it is best to conduct this analysis using a data mining software module written specifically for cluster analysis.

2.5 Prioritize Information Quality Objectives for an Information Product

For the information product manager, the issue here is that often not all customers are equally important nor their information products equally vital. As a result, the information product manager will want to ensure that the highest information quality be given to key customers and their key information products. Pareto analysis can be used to help plot some measure of importance (e.g. % of profits, % of savings, etc.) by customer and by information product (e.g. % of information products/customers). The resulting curve is traditionally divided into three categories: the two 20 percent of products and customers by importance are the 'A' category; the next 30 to 50 percent are labeled 'B' and the final group is category 'C' [5] [6].

This A, B, C categorization can be used as the basis for information product inventory control whereby the highest level of quality is provided for the 'A' information products, a slightly lower level for the 'B' information products and lower still for the 'C' information products. The information product manager might also differentiate the information product holdings by manufacturing the 'A' items as close as possible to the customer while the 'B' and 'C' items can be produced further along the information supply chain or even be made candidates for outsourcing.

2.6 Design a Customer-Based Information Product Quality Strategy

Once the information product manager knows which consumers, data components, and data quality dimensions to focus on, he can then formulate the information product quality strategy. At each stage in the process from the first contact with the consumer through the manufacture of the information product to any after-manufacture support, it is important for the information product manager to establish precisely defined quality objectives for that information product. The design of the information product quality strategy should take into account the different needs of consumers so that the resources allocated to information quality can be used in the most cost-effective way. [5] [6]. At this stage, the information product manager may need to negotiate with the various consumer segments to match their specifications and expectations for the level of information quality to be provided with budgetary and other constraints for the information product.

The listing below gives an indication of the level of detail that is required for specifying the objectives of the information product quality strategy [5] [6].

Examples of pre-transaction elements

- **Written information product quality policy** (Is it communicated internally or externally, is it understood, is it specific and quantified where possible?)
- **Accessibility** (Is the information product manager easy to contact/do business with? Is there a single point of contact for discussing quality issues with the information product?).
- **Organizational structure** (is there an information quality management structure in place? What level of control do they have over the information manufacturing process?)
- **System Flexibility** (Can the organization adapt their information quality delivery systems to meet particular customer needs?)

Examples of transaction elements

- **Order cycle time** (What is the elapsed time from order to delivery of the information product? What is the reliability/variation?)
- **Data availability** (What percentage of demand for each information product can be met from currently stored data? What are the quality levels of currently stored data components?)
- **Order fill rate** (What proportion of orders for information products are completely filled within the stated lead time?)
- **Customer Driven Data Quality Requirements and Requests** (What are the customer-driven data quality specifications and expectations for the information product and its data components?)
- **Other Information Product Requirements and Requests** (Are there any other cost, regulatory, or standardization considerations that should be factored into the design and manufacture of the information product?)
- **Order status information** (How long does it take the organization to respond to a customer query with the required information? Does the information product manager inform the customer of problems or do the customers contact the information product manager).

Examples of post- transaction elements

- **Call out time** (If a problem with the information product is discovered, how long before the problem is fixed?)
- **Product tracing/warranty** (Can the organization track the location of information products once they are delivered to the customer? Are information products properly archived or disposed of once they are no longer needed?)
- **Customer complaints, claims, etc.** (How promptly does the organization deal with information product complaints? Does the organization measure customer satisfaction with their response?)

Once the information product quality strategy has been developed and recorded, it must be implemented and monitored. For the sake of completely describing the full usage cycle of the information product quality strategy, the remaining sections of this paper summarize the literature-to-date on how to deliver an information product that fulfills a specified set of quality requirements.

3. IMPLEMENT THE CUSTOMER-BASED IP QUALITY STRATEGY

If the information product quality strategy prescribes how to satisfy customers' quality specifications and expectations for an information product, then the implementation program is the step-by-step plan for putting the information product quality strategy into action. The information product quality strategy should become part of the requirements analysis used to develop or modify the information system that will ultimately manufacture the information product. It is at this stage that the vast literature on how to construct a sound information system driven by customers' requirements is applied. From the design of the conceptual model and system specifications through the database implementation, software engineering, and hardware selection and then on to quality assurance, acceptance testing, changeover, and finally the operation and maintenance phases, the information systems development group should be guided by the customer-based information product quality strategy.

4. MONITOR THE CUSTOMER-BASED IP QUALITY STRATEGY

During the operation and maintenance phase of an information system's life, the information product manager will need to continuously monitor the system's output in order to assess how well the information products being produced by the information system are meeting the consumers' quality specifications and expectations. If the information product manager discovers gaps between the quality of an information product being produced and the consumers' quality criteria, then this signals that it is time to revisit the information product quality strategy and to implement changes in the information system that manufactures the information product. The information product manager will find the following techniques valuable for accomplishing these tasks.

4.1 Document the Process that Creates the Information Products

During the course of the information system's development, documentation about the system's inputs, processing, and outputs should have been created. Program flowcharts, data dictionaries, entity-relationship diagrams, workflow models, and data flow diagrams all contribute knowledge that the information product manager can use to better understand the system that is manufacturing the information product. Additionally within the last few years, a new tool called an IP-Map has emerged that is specifically designed to help the information product manager to visualize the data sources, processes, systems, business units, and organizations involved in the creation of the information product [10]. An IP-Map consists of a set of constructs combined with metadata that systematically represent the manufacture of an information product. The IP-Map's constructs are similar in some respects to those found in data flow diagrams, but the IP-Map's constructs allow more differentiation of the stages involved with the information product's manufacture such as inspection or cross-over of organizational boundaries [10]. This systematic representation of the information product's creation allows the information product manager to more easily locate sources of quality problems such as bottlenecks, conflicts of ownership, timeliness, completeness, or accuracy issues.

4.2 Measure Quality Levels of Information Products

An often quoted phrase is "If you can't measure it, you can't manage it". Good metrics allow the information product manager to describe over time the quality levels of the information product and its data components, a necessary condition for making and sustaining information product quality improvement. But although metrics can indicate to the information product manager that a quality problem exists, they cannot determine by themselves either the source or the solution to the problem. Compounding these limitations is the fact that establishing an effective and efficient method for collecting and using metrics is often difficult. Without thorough planning, it is easy to fall into the trap of measuring the wrong thing in the wrong place or time using the wrong measurement device.

Thomas Redman [11] uses the term *measurement system* to refer, collectively, to all activities that make up measurement, from deciding what to measure, to specifying the measurement protocol, to presenting results, to using results to make improvements. His measurement system is built around five components.

(1) Business Requirements: This component addresses the questions of who will use the measurements and for what purpose.
(2) What to Measure: This component examines the issues of where measurements will be taken, what data will be included, how will the measurement be defined and what scale will be used for reporting results.
(3) Measurement Protocol: This component specifies how measurements will be taken, whether sampling will be employed, who will take them, where data will be stored, and how results will be summarized and presented.
(4) Management Action: This component asks how individuals will use the measurements. Will metrics be used to execute control, to identify root causes of data quality and eliminate them, etc.?
(5) Integration/Evolution: This component emphasizes that appropriate measures fit into their appropriate organizational contexts.

By designing a measurement system that addresses these five components, the information product manager will have the metrics he needs to monitor and control the information manufacturing process so it is in line with the information product quality strategy.

4.3 Identify Gaps in the Quality of Information Products

If there are gaps between an information product's quality and its customers' specifications and expectations, how can the information product manager identify the source and the solution for the quality mismatch? Several types of tools including information quality tools, total quality management tools, data profiling tools, and the benchmarking methodology can help.

4.3.1 *Information Quality Tools*

Recently several tools have been developed specifically for measuring, analyzing, and improving information quality. These tools include the following:

- **PSP/IQ Model**: Many products today have both service and product quality attributes. Take for example, the automobile. The value associated with owning a particular manufacturer's automobile is enhanced by services like roadside assistance, financing options, and maintenance plans that come with the vehicle. The PSP/IQ model developed by Lee, Strong, Kahn, and Wang [12] illustrates how the multiple dimensions of information quality can be categorized as contributing either to the product or service quality of an information product. This categorization provides a framework for the information product manager to verify if the information product quality strategy incorporates both product and service quality aspects and whether or not these quality aspects can be assessed against a formal specification or customer expectation.

- **IQ Assessment Survey**: This questionnaire is designed to measure information quality along the dimensions important to information consumers and managers. Several of these dimensions together measure information quality for each quadrant of the PSP/IQ model [12]. Two techniques can be employed for interpreting the assessments captured by the questionnaire [12]. The first technique compares an organization's information quality to a benchmark from a best-practices organization. The second technique measures the distances between the quality assessments of different stakeholders of an information production system. With these analysis techniques, information product managers can investigate the quality of their information products and determine appropriate areas for improvement.

4.3.2 TQM and SPC Tools

As mentioned before, the field of product manufacturing has created an extensive list of tools and techniques to support total quality management [13]. Many of these tools and techniques can be applied to the analysis and improvement of an information product so it meets the objectives of the information product quality strategy. Below is a list of the most common TQM tools available to aid information product managers in discovering and rectifying the root causes of data quality problems.[‡]

- **Histogram**: This chart shows the nature of the distribution of a variable's values by plotting the frequency of occurrence against their respective data values.
- **Control Chart**: This chart plots the variation in a process's measurements over an extended period of time.
- **Process Flow Diagram**: This diagram documents the process flow and interaction among the process steps.
- **Check Sheet**: This cross tabulation table lists the frequency of defects by categories such as location, time, or cause.
- **Scatter Diagram**: This diagram plots the relationship between two variables of interest.
- **Pareto Diagram**: This chart is essentially a sorted histogram. It highlights problem areas with the most frequent occurrences.
- **Cause and Effect Diagram**: This diagram displays all contributing factors and their relationships to a particular problem area.
- **Defect Concentration Diagram**: This diagram helps to determine whether the location of the defects on a product conveys any relevant information about the potential causes of the defects.
- **Force Field Analysis**: This diagram identifies factors that may influence a problem area either positively or negatively.
- **Design of Experiments**: This statistical technique is used to analyze the magnitude and direction of factors that affect a process's variation.

[‡] Researchers and practitioners are continuing to discover and explain the root causes of poor quality data. One of the best examples of this type of research is the article: Strong, D. M.; Lee, Y. W.; and Wang, R. Y., "10 Potholes in the Road to Information Quality", *IEEE Computer*, 30, 8 (August 1997), pp. 38-46.

- **Acceptance Testing**: This is an audit tool used to ensure that the output of a process conforms to requirements.

4.3.3 *Data Profiling*

Another practical technique for identifying the source of data quality problems is to use data profiling [14]. Data profiling seeks to identify problems like duplicate records, missing values, inconsistent formats, domain problems, etc. by employing tools like SQL to verify column, entity, and referential integrity as well as business rules. Once a data quality issue has been uncovered, the examiner then considers how this problem may impact the business processes that rely on this data as well as the part that data entry practices may play in propagating these data quality issues. The results from a data profiling study assist the information product manager in recommending improvements to the way an organization's information systems process their source data so that the resulting information products are consistent with the desired quality strategy.

Due to the work-intensive nature of data profiling, information product managers may wish to use a software tool to help automate the analysis. Several vendors now offer software tools capable of performing many data profiling checks. One of these tools, the Integrity Analyzer, is available from www.crg2.com [3]. From its menus, individuals can elect to evaluate entity integrity, referential integrity, column integrity, and user-defined integrity of a specified data set. In addition, the tool is able to perform frequency checks, to browse tables, and to build data views.

4.3.4 *Competitive Benchmarking*

Benchmarking is a process that evolved from the practices of site visits, reverse engineering, and competitive analysis. Benchmarking works by comparing and measuring a company's processes against those of a consenting best-in-class performer. Benchmarking aids a firm in determining where it stands relative to best-in-class practices and assists the firm in making rapid performance improvements where it lags far behind [15]. As the field of information quality expands, information product managers may find benchmarking against companies with highly developed information quality programs a constructive way to identify and correct lapses in their own information product quality efforts.

5. DISCUSSION & CONCLUSIONS

It is important to note that the Develop-Implement-Monitor steps in establishing an information product quality strategy is not a linear prescription, rather it is a cyclical one. Firms that already have a set of information products and systems in place may wish to start with the Monitor stage to first evaluate if there are any gaps between their consumers' quality specifications and expectations and the information products currently being delivered. Based on that assessment, firms can then refine their information product quality strategy and develop a set of proposed changes which are then implemented in the organization's information systems. Monitoring is continued to verify if those changes did indeed result in improving the quality of the information products.

It is also important to note that improving information products is an on-going endeavor. Just like Ford Motor Company must continually re-evaluate their products lines to meet changing technological capabilities and consumers' expectations, organizations must continually review their information products to keep them current with innovations in information technology and changing knowledge needs. Lee et al. [3] perhaps express it best when they write that achieving data and information quality should be considered a journey, not a final destination. In that line of thinking, the concept of an Information Product Quality Strategy and its associated techniques presented in this paper should be considered a proposal. The next step will be to develop a more detailed and elaborate case study that builds upon the work presented in this paper. Such an empirical study incorporating multiple users, information products, and quality features would greatly help to validate the approach recommended by this paper.

REFERENCES

[1] Tuomi, I., "Data is More than Knowledge", *Journal of Management Information Systems*, 16, 3 (Winter 1999-2000), pp. 103-117.

[2] Wang, R. Y., Lee, Y.W., Pipino, L.L., Strong, D.M., "Manage Your Information as a Product", *Sloan Management Review*, Summer 1998, pp. 95 - 105.

[3] Lee, Y.W., Pipino, L.L., Funk, J.D., Wang, R.W., *Journey to Data Quality*, MIT Press, Forthcoming.

[4] Albrecht K. and Zemke, R., *Service America*, Dow –Jones Irwin, 1985.

[5] Christopher, M., *Logistics and Supply Chain Management*, *2nd Edition*, Prentice Hall, 1998.

[6] Christopher, M., *The Customer Service Planner.* Butterworth-Heinemann Ltd., 1992.

[7] Wang, R. Y. Strong, D.M., "Beyond Accuracy: What Data Quality Means to Data Consumers", *Journal of Management Information Systems*, 12, 4 (Spring 1996), pp. 5-33.

[8] Curry, J., "Understanding Conjoint Analysis in 15 Minutes", www.sawtoothsoftware.com, Accessed on May 27, 2004.

[9] www.clustan.com, Accessed on May 27, 2004

[10] Shakaranarayan, G. Ziad, M., Wang, R. Y., "Managing Data Quality In Dynamic Decision Environments: An Information Product Approach", Forthcoming in *Journal of Data Management*.

[11] Redman, T., "Measuring Data Accuracy: A Framework and Review", *Information Quality – Part of the Association of MIS Monograph Series*, M.E. Sharpe, Forthcoming 2005.

[12] Lee, Y.W., Strong, D.M., Kahn, B.K., Wang, R.Y., "AIMQ: a methodology for information quality assessment," *Information & Management*, 40, 2002, pp. 133-146.

[13] Montgomery, D.C., *Introduction to Statistical Quality Control*, John Wiley & Sons, 2000.

[14] Olson, J.E., *Data Quality: The Accuracy Dimension*, Morgan-Kaufmann, 2003.

[15] Rao, A., Carr, L., Dabolena, I. et al., *Total Quality Management: A Cross Functional Perspective*, John Wiley & Sons, 1996.

[16] Orme, B., "Which Conjoint Method Should I Use?", www.sawtoothsoftware.com, Accessed on August 21, 2004.

[17] "Analysis of Traditional Conjoint Using Excel: An Introductory Example", www.sawtoothsoftware.com, Accessed on August 21, 2004.