Diane M. Strong
Worcester Polytechnic Institute

Yang W. Lee
Cambridge Research Group

Richard Y. Wang
Massachusetts Institute of Technology

10 Potholes in the Road to Information Quality

Poor information quality can create chaos. Unless its root cause is diagnosed, efforts to address it are akin to patching potholes. This article describes 10 key causes, warning signs, and typical patches. With this knowledge, organizations can identify and address these problems before they have financial and legal consequences.

When spring finally comes to New England, it brings with it gaping holes in roads. The larger of these potholes are patched, but the road can never be made smooth again. Smaller potholes, although a daily nuisance, are often not repaired at all. Something similar is happening in the area of information systems: People who need information to accomplish their tasks are finally being provided with easy online access to relevant information. But the information highway has potholes. Like New England drivers in the spring, information consumers must dodge quality problems, large and small, in their quest for high-quality information. Like road crews, the people who produce, store, and maintain information achieve minimal quality through a never-ending process of patching rather than repairing potholes.

These potholes along the path from collecting raw data to providing useful information are caused by what we call *information quality* (IQ) problems. We use the term "information" to refer to both data and information. "Data" usually refers to information at its early stages of processing, and "information" the product at a later stage. Rather than switching between terms, we use "information" to refer to data or information values at any point in the process.

ROAD TO INFORMATION QUALITY

Figure 1 illustrates the road to IQ. In the figure, the production and storage of information are conceptualized as an information manufacturing layer on top of an underlying information infrastructure.[1] Central to this is the concept of a production process that transforms data into information that consumers find useful.

We identify three roles within this information manufacturing system and we associate with each role a process:

- information producers generate and provide information, the raw material for the information manufacturing system;
- information custodians provide and manage computing resources for storing, maintaining, and securing information; and
- information consumers access and utilize information for their tasks. Utilization may involve additional information aggregation and integration.

Like potholes, IQ problems often arise unexpectedly and cause major negative consequences before they are resolved. When large potholes develop along this path, most organizations do little more than patch them. Smaller potholes, or those that are extensive but not highly visible, are ignored. There is a better way. Organizations do not need to wait for problems to appear before they address them. If they know what to look for, organizations can anticipate and handle IQ problems before they trigger a crisis. Doing

0018-9162/97/$10.00 © 1997 IEEE

so requires an understanding of the underlying causes of IQ potholes and how they are typically patched.

First, it must be understood that IQ problems encompass much more than incorrect information values.[2-3] They also include production problems and errors, technical problems with storage and access, and problems caused by the changing information needs of consumers. Proactive management of IQ means attending to information processes, technical computing details, and consumer requirements.

The goal is high-quality information, which we define as information that is fit for use by information consumers. This definition follows directly from the quality literature in which fitness for use is widely adopted as the criteria for high quality.[4-5] This means that usefulness and usability are important aspects of IQ—after all, this is how consumers view IQ. In fact we know that there are four major aspects to information quality and 15 dimensions underlying these aspects, as Table 1 shows.[3] A detailed definition of these quality attributes is outside the scope of this article, but we will introduce these terms below, in the context of the problems we observed.

Table 1. IQ Categories and Dimensions.

Category	Dimension
Intrinsic IQ	Accuracy, Objectivity, Believability, Reputation
Accessibility IQ	Accessibility, Security
Contextual IQ	Relevancy, Value-Added, Timeliness, Completeness, Amount of Information
Representational IQ	Interpretability, Ease of Understanding, Concise Representation, Consistent Representation

ROOT CAUSE ANALYSIS

In this article, we describe 10 key IQ problems we have observed, their typical patches, and our recommendations for long-term solutions. We base these observations on research that employed qualitative data collection and analysis techniques.[6-7] We interviewed information producers, custodians, consumers, and managers about IQ problems and their resolution. We also examined organizational documents and observed work practices. This data collection produced 42 mini-case studies of IQ improvement projects in three organizations.[8] We then analyzed these cases using content and pattern analysis techniques.[6]

While the 10 problems we identify here may not constitute a complete list of underlying causes of all organizational IQ problems, it does include what we found to be the key causes in three leading-edge organizations: an airline, a hospital, and a health maintenance organization (HMO). We selected these

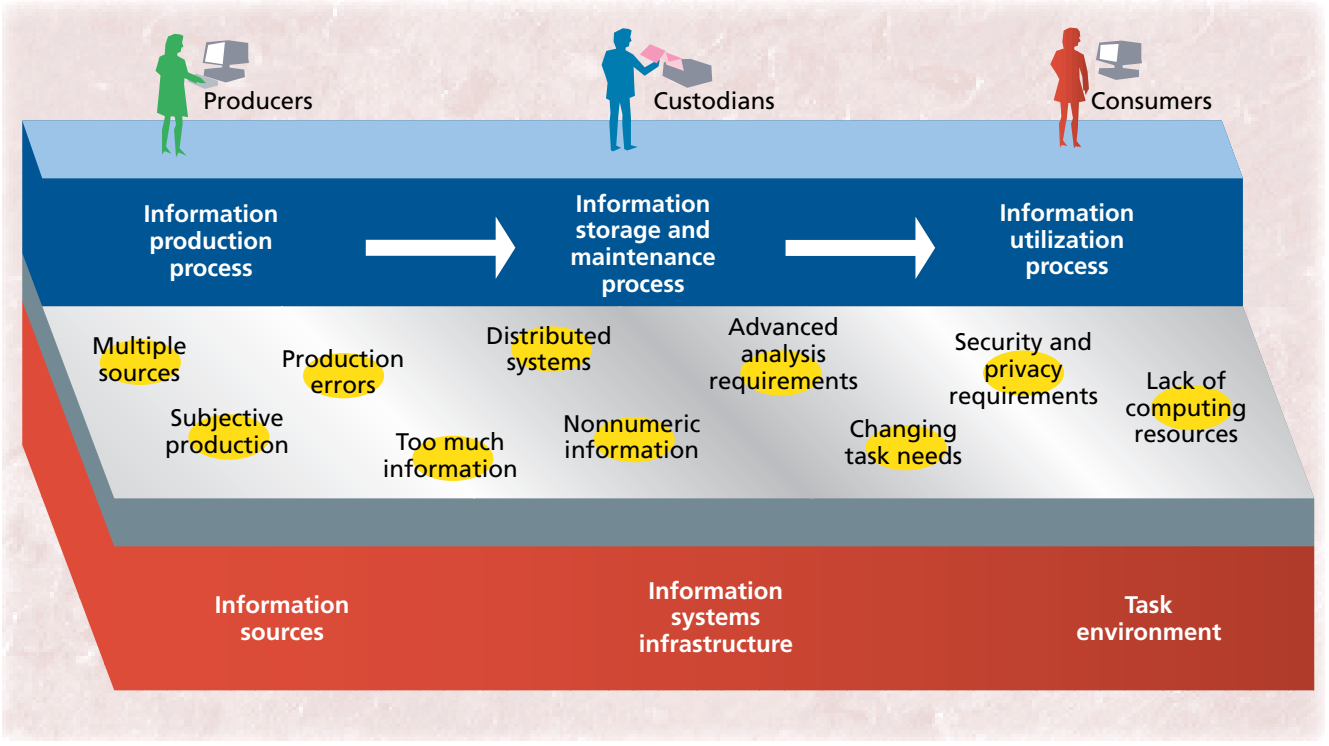
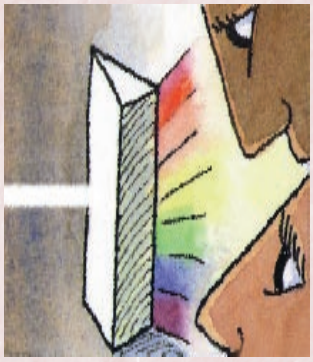


Figure 1. Ten key information quality problems in information manufacturing systems.

#1. Multiple sources of the same information produce different values.



Example: A hospital uses two separate procedures to assess illness severity.
Dimensions affected: Consistency and believability.
Organizational effect: Financial and legal problems.
Warning signs: Different systems developed for different purposes that require the same information.
Patches: Use only one of the systems. Download one set of information to consumers.
Problems with patches: Lose other purposes for which unused systems were developed.
Real solution: Develop common definitions and consistent procedures.

organizations as research sites because their IQ practices exceed the current state of IQ research and because they provide sufficient variance in their information management practices.

The remainder of this article presents the problems according to whether their root cause was attributable to information production problems, technical problems with storing and accessing information, and the information needs of consumers.

Information Production Potholes

Three potholes are the direct result of how information producers and processes work. Poorly designed and poorly managed information production processes generate poor-quality information as raw material.

Multiple sources of the same information produce different values. Database designers warn against storing the same data in multiple places because it is difficult to ensure that multiple copies will be updated consistently. Similarly, producing the same information using several different processes is likely to produce different values for the "same" information. For example, one hospital produces illness-severity assessments both by having a specialist evaluate patients upon admission and by having a nurse observe patients during their stay. Not surprisingly, the two assessments sometimes differ. But billing and other reporting contexts require one consistent value.

Similar information production situations arise frequently in organizations. It most often happens when systems designed for different purposes require the "same" information as input. When these systems are developed autonomously, the result is parallel but slightly different procedures for collecting the same information. This creates a problem with what is known as representational quality.

Some may view this kind of IQ problem as just a bump in the road, not a pothole. But multiple sources of the same information can cause serious problems. In hospitals, multiple values for the same information can create financial and legal problems. It can also contribute to incorrect assessments of the costs of various clinical practices. And consumers may stop using the information when inconsistencies lead them to doubt the believability, and hence the intrinsic value, of the information.

This kind of pothole is often not repaired, leaving multiple production procedures to continue to produce different information values. However, this problem may be covered (hidden from consumers) by selective use of the systems. For example, clinical consumers access clinical information and financial customers, financial information. In other situations, all users may gravitate to one source. Or information may be downloaded selectively to consumers.

Moving beyond patching requires a reexamination of information production processes. The hospital we studied arrived at a common definition for illness severity and a production process that would consistently produce it. And it changed its computer systems to implement this process.

Information is produced using subjective judgments, leading to bias. Information stored in organizational databases is considered to be factual. The process by which these "facts" are collected, however, may involve subjective judgments. Illness severity is one example. Another is the assignment of medical codes

#2. Information is produced using subjective judgments, leading to bias.

Example: Medical coders use judgment in selecting disease codes.
Dimension affected: Objectivity and believability
Organizational effect: Financial, legal problems.
Warning signs: Subjective information production activities.
Patches: More information for production rules to reduce information variance.
Problems with patches: Added rules are complex, subjective, and possibly inconsistent.
Real solution: More training, better rules, expert systems.

#3. Systemic errors in information production lead to lost information.

Example: Information missing because edit check would not accept; incorrect information because edit check would accept.
Dimension affected: Correctness, completeness, and relevancy.
Organizational effect: Information lost or distorted.
Warning signs: Purchase or develop new entry system that has excessive edit checks; pay entry clerks by volume and not by quality.
Patches: Enter information into a nonedited (comment) field and move to appropriate field by program patch; tell entry clerks not to enter incorrect information.
Problems with patches: Requires extra programming and more complex information entry.
Real solution: Statistical process control; process improvement; behavioral control and proper incentives.

to indicate a diagnosis and categorize medical procedures performed. There are rules for assigning these codes, but medical coders must still exercise judgment in selecting appropriate codes. Such subjectively produced "facts" are less objective and more suspect (of lower quality) than information produced without any human judgment.

Using subjective information to produce information is like building a road with poor materials and sloppy procedures. The resulting information will be avoided by consumers, just as they avoid bad roads. Thus, information collected using procedures that rely on costly human judgment may not provide sufficient organizational benefit to make it worth collecting.

These IQ potholes are often covered or patched. Covering involves hiding the extent of subjective judgment involved in creating the information. For example, consumers will tend to believe that information from computer systems is more factual than it really is. Patching involves adding more production rules to handle the variance in information produced from similar underlying facts.

A third, more extreme approach to addressing this problem is to eliminate human judgment. After all, one way to eliminate potholes is to eliminate roads. But we do not advocate the elimination of human judgment from information production. Doing so would severely restrict the information available to consumers because some information can only be produced subjectively. Thus, some form of patching is necessary. If you know how your information is produced, you can continuously improve the activities that involve subjective assessments, so that cracks in the road never develop into potholes. For example, most medical coders undergo a year of training and use comprehensive manuals that specify the rules for assigning codes. Continuous improvement involves more training, better rules, and advanced computer systems, including expert systems.

Systemic errors in information production lead to lost information. Because IQ involves more than just correctness or accuracy, we focus not on errors (mistakes), but systematic problems. Some errors in producing information, however, are actually systematic problems. For example, even those databases

#4. Large volumes of stored information make it difficult to access information in a reasonable time.

Example: Multiple-year trend analysis needed over 12 Gbytes of information for each year and need to analyze several thousand records out of several million.
Dimensions affected: Concise representation, timeliness, value-added, and accessibility.
Organizational effect: Excess time required to extract and summarize information.
Warning signs: Large amount of operational information with need for managerial or strategic analysis of this information.
Patches: Condense information using codes; create extracted subsets of information as needed.
Problems with patches: Difficulties for information consumers in interpreting codes; need for Post-its; not timely for analysis.
Real solution: Rewrite using GUI and power of client systems; analyze information needs and develop regular, frequently extracted subset of this information.

#5. Distributed heterogeneous systems lead to inconsistent definitions, formats, and values.

Example: Different systems in each division, each using a different format for diagnostic codes.
Dimensions affected: Consistent representation, timeliness, value-added.
Organizational effect: Inconsistent information that is difficult to access and aggregate.
Warning signs: Multiple systems across departments.
Patches: Consumers manage extraction from each system and aggregation of the information.
Problems with patches: Consumers do not understand data and file structures, creating burden on consumers.
Real solution: Data warehouse.

that require strict edit checking may be compromised if data-entry clerks change values in order to pass the edit check or simply skip fields. While errors can occur anywhere in information manufacturing processes, systematic errors during production are especially important because they affect the entire system.

For example, in our study the HMO clerks could not enter all the codes for outpatient surgery into computers because edit checks rejected those codes in hospital format, but accepted those in doctor's office format. This problem was discovered only when an analysis conducted using this information produced suspicious results. An investigation revealed that this routine nonentry had been occurring since the system had been purchased. Because no one had complained, the clerks assumed the information was not important. This information is now lost, since it would be costly—and perhaps impossible—to recover it.

This IQ pothole is caused by hidden problems in information production processes. Patching these potholes means fixing these problems as they are discovered. For example, the HMO used a temporary workaround: Clerks entered the information into a comment field until the vendor released a new version of the software that accepted all the codes. Moving beyond patching means understanding, documenting, and controlling information processes just as is done with manufacturing processes. Manufacturing control techniques, such as statistical process control, can be applied to information processes.[9]

Information Storage Potholes

Three potholes are the consequence of the difficulties in storing large amounts of varied information across different computer systems.

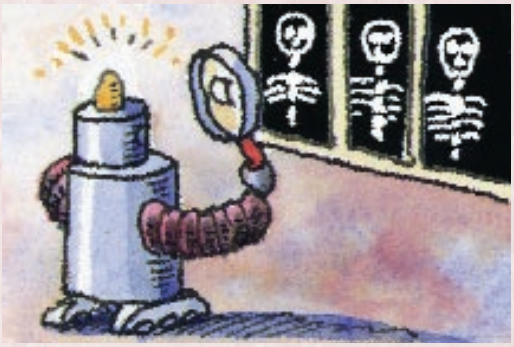
Large volumes of stored information make it difficult to access in a reasonable time. More information is not necessarily better. Large volumes of information present problems for those responsible for storing and maintaining the information and for those searching for useful information. Obviously this is a problem in the accessibility aspect of IQ, but it also affects the timeliness of the information and the value-added dimensions of information. A telephone company, for example, generates thousands of billing transactions hourly. Yet customers expect each company representative to have immediate access to their billing records so their questions can be resolved instantly.

Volume alone impacts the ability to use information: The HMO we studied generates more than a million patient records each year. To analyze disease trends over a several-year period may involve only a few thousand records, but these must be selected from millions. One hospital, which serves tens of thousands of patients each year, generates 12 Gbytes of operational information at the same time. Multiyear trend analysis thus still presents a problem.

#6. Nonnumeric information is difficult to index.

Example: Medical image information is relatively easy to store but difficult to access; doctors' notes are difficult to enter but relatively easy to access.
Dimensions affected: Concise representation, value-added, accessibility.
Organizational effect: Costly storage of information with potentially little benefit.
Warning signs: Much operational information is in the form of images or text.
Patches: Electronically store text and image information.
Problems with patches: Electronic storage can be costly on the input side with limited benefit on the output side.
Real solution: Assess the benefits of electronic storage and compare to costs of input and storage of information.

#7. Automated content analysis across information collections is not yet available.



Example: Difficulties in analyzing trends in image and text information: "Is pneumonia increasingly common in ICU patients?"
Dimensions affected: Analysis requirements, consistent representation, relevance, value-added.
Organizational effect: Analysis of image and text information stored in electronic form is very limited.
Warning signs: Operational and managerial decisions require analysis across multiple image or text files.
Patches: Use coding systems to summarize text; use various algorithms to analyze images like CT scans.
Problems with patches: Patches address only part of the problem and may generate new problems, for example, codes are difficult to interpret.
Real solution: Be aware that with electronic storage, information consumers expect analysis routines to compute trends across files.

Information systems professionals have standard techniques for storing large amounts of information and for providing efficient access. One such technique uses codes to condense textual information (a hospital patient's religion, for example, is coded as 1 for Protestant, 2 for Catholic, 3 for Jewish, and so on). While such concise representations are common, they can put an undue burden on consumers, who must interpret the codes. As one consumer said, "I judge the quality of a computer system by the number of Post-Its attached to the terminal." Post-Its hold small reminder notes, and their presence indicates the need for memory aids to interpret the actual meaning of displayed information. As systems are rewritten to use the processing power of client machines and graphical user interfaces, such problems are being reduced. But such problems are still common.

A common patch is to create extracted subsets of information into data warehouses. At the HMO, historical analyses are done using a weekend batch extract, an ad hoc solution that significantly increases the time required to access and analyze this information. The hospital created an additional database loaded with a subset of the yearly operational information over about a dozen years. This new database, which is stored on an easy-to-access client-server system, is updated weekly.

In designing and managing this database, the hospital has moved beyond patching to create a permanent solution to its information volume problem. Such a solution requires proactive analysis of what information is needed to manage the hospital and conduct medical research. This solution is only as effective as the analysis of which information is important to save.

Distributed heterogeneous systems lead to inconsistent definitions, formats, and values. Just as roads add to a nation's transportation capacity, so distributed systems add to organizational capacities: Distributed heterogeneous systems allow information to be accessed and analyzed almost as if it were in one location. Like roads, though, distributed systems have potholes. Among other things, it can take an excessive time to select and aggregate relevant information from several systems, adversely affecting the timeliness and value-added dimensions of IQ.

The most common IQ problem associated with distributed systems is inconsistent information—information that has either different values or different representations across systems. Different values are generated from multiple sources or created by inconsistent updating of multiple copies. Different representations (formats or codes) must be accommodated when autonomously designed systems are integrated. For example, each HMO division stores medical codes differently; some use decimal points and some don't. Every time this information is integrated, attention must be given to this problem. Routines to handle such problems, whether manual or computerized, represent patches.

Data warehouses are a current solution to distributed system problems. Rather than rewrite old, autonomously developed systems, a data warehouse is populated using extraction routines that pull information from old systems and resolve inconsistencies. This presents a centralized, consistent front end to distributed back-end systems.

Nonnumeric information is difficult to index. It is now possible to collect and store many types of information, including images. Storage is not the issue: Representing this information concisely, for the purpose of easy entry and access, is.

Detailed doctors' notes, for example, largely remain on paper because they are too expensive to convert to electronic form: Doctors are unlikely to enter their notes directly into a computer system, and having a third party decipher and key them is time-consuming. In cases like this, it is easier to store information as images—new image-making equipment has made this easy. However, retrieval of text and image data by content is very difficult.

As information technology storage and retrieval capabilities improve, organizations need to decide how much, if any, of this kind of information to store. Why store this information if it makes little or no improvement in the quality of information delivered to consumers? The HMO did decide to store all clinical information electronically, both text and images. In this case, the need to access a high volume of patient information at every doctor visit is sufficient to justify the costs of collecting the information electronically.

Information Utilization Potholes

Four potholes are the evidence of the failure to meet consumer requirements for advanced information analysis in the context of the changing task environment of information consumers.

Automated content analysis across information collections is not yet available. Providing consumers access to nonnumeric information distributed across systems is not enough. If the information is to have relevance and any added value, consumers must also be able to aggregate it, manipulate it, and use it to compute trends. A related problem is the difficulty of analyzing trends across information stored in multiple systems with inconsistent definitions, names, or formats.

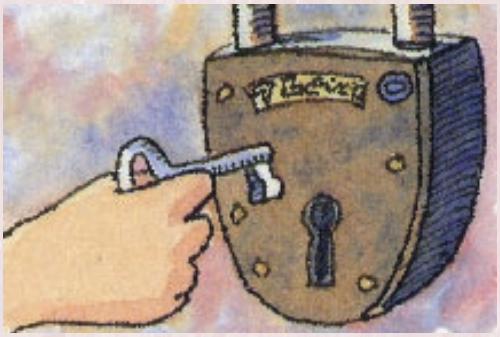
For example, while it may be convenient to store and access X-rays electronically, X-rays are not of much use if they can't be analyzed. A single X-ray is analyzed to spot diseases; a series of X-rays is analyzed to assess changes in health. Analysts and medical researchers want to analyze electronic images in the same way they analyze quantitative information, using statistical summary and trend analysis techniques.[8]

Providing storage and access to information without also providing analysis capability is like building a road that has so many potholes it is minimally useful.

#8. As information consumers' tasks and the organizational environment change, the information that is relevant and useful changes.

Example: The basis for insurance reimbursement to hospitals changes, requiring changes in information processes and systems.
Dimensions affected: Relevance, value-added, completeness.
Organizational effect: Mismatches develop between available information and what is needed for tasks.
Warning signs: Changes in consumers, consumer tasks, and competitive or regulatory environment of the organization.
Patches: Only when mismatch between information needs and available information becomes too large, revise information processes and systems.
Problems with patches: Information, processes, and systems lag far behind the needs of information consumers.
Real solution: Anticipate changes in consumer tasks and revise processes and systems before the mismatches become crises.

#9. Easy access to information may conflict with requirements for security, privacy, and confidentiality.



Example: Patient medical information must be kept secure and confidential, but analysts and researchers need access.
Dimensions affected: Security, accessibility, value-added.
Organizational effect: Mechanisms for security bar access, so the information provides less value.
Warning signs: Vital information is not accessible, within reasonable constraints.
Patches: Local solutions to security breaches and accessibility complaints as they occur.
Problems with patches: Patchwork solutions make every situation unique, which increases time to negotiate accessibility.
Real solution: Develop consistent policies and procedures for secure information.

For this reason, the hospital decided *not* to store all text and image data electronically, in contrast to the HMO's decision: Electronic storage is not sufficiently beneficial until they can automate the analysis of several X-rays and compute trends across many X-rays.

This pothole can be patched in several ways. Coding systems, for example, are essentially patches that enable the analysis of textual information. Routines to match fields across distributed systems assist with analyzing information across these systems. Patches like this create more problems. First, each patch is an inherently piecemeal solution because each routine must be tailored to specific fields. Solving the entire problem requires several patches. Second, these patches may be incomplete or they may cause other problems. Matching names across systems, using Soundex algorithms for example, will fail to find some correct matches. And systems for summarizing information in an analyzable way generates analysis problems if consumers are required to interpret the codes.

Solving these problems requires advances in information technology and its applications. Data ware-

houses and common data dictionaries are solutions to analyzing structured information across systems. Much progress is being made in algorithms for analyzing image information: computed tomography (CT) scans are analyzed by computing algorithms, for example. In general, the capability for storing new forms of information will always be developed before full capabilities for analyzing that information.

Consumers' needs for information changes. Information is only of high quality if it meets the needs of information consumers; only then does it satisfy the contextual aspect of IQ. Providing the information that consumers need is a difficult problem because there are multiple consumers, each of which may have differing needs. Furthermore, these needs change over time. While an initial solution to consumers' information needs may be satisfactory, over time the quality of the solution deteriorates. This is similar to building a road and not expecting potholes.

For example, health insurance companies used to reimburse hospitals on the basis of fees charged for individual procedures performed. Now the fees are fixed, according to the assigned disease code. This necessitates different processes for collecting and using information for billing and different management information needs for analyzing costs.

This pothole represents mismatches between the information provided and the information needs of consumers. When the pothole becomes large enough, it is patched by changing processes and computer systems. As mismatches develop, however, consumers develop various manual and computerized workarounds.[10] Workarounds can become routine, so that the road consists of many small potholes that never grow large enough to draw sufficient attention to be patched or repaired.

Moving beyond patching means planning for changes to information processes and systems and anticipating changing consumer needs before they become serious IQ problems. Systems should also be designed for flexibility in reporting. The hospital anticipated the change in reimbursement procedures and made changes to its processes and systems before the change was implemented. Hospitals that failed to do so suffered severe financial problems.

Easy access to information may conflict with requirements for security, privacy, and confidentiality. To information consumers, high-quality information is information that is easily accessible. Ensuring privacy, confidentiality, and security of information, however, requires barriers to access. In this case the accessibility and security goals conflict. For example, medical records are confidential, yet without access to them analysts can't do research and managers can't make decisions. Protecting patient privacy means limiting access to patient records by, for example, requiring legal department permissions to use patient records. To consumers, the need to obtain advanced permission is a barrier to information access.

Patching these potholes means finding ad hoc solutions to privacy, confidentiality, or security problems as they arise. For example, if a patient's HIV status is released unintentionally, new procedures are developed to prevent a reccurrence. These new procedures can be developed to minimize barriers to accessibility for legitimate tasks.

Moving beyond patching means developing privacy, confidentiality, and security policies and procedures for all information when it is first collected. These policies then guide the development of standard procedures to assess the need for access with minimal effort and delay. Consumers typically recognize the need for privacy, confidentiality, and security of information and are willing to abide by reasonable rules.

Lack of sufficient computing resources limits access. While most US knowledge workers have computers in their offices and communications lines are reliable, computing resources are limited. Skimping on resources can cost more money in the end. For example, the airline used unreliable communications lines with insufficient capacity to access and maintain a parts inventory. As a result, not all inventory transactions were recorded, leading to inaccurate and incomplete information. The HMO did not provide

#10. Lack of sufficient computing resources limits access.

Examples: Unreliable communications lines lead to incomplete information. Shortage of terminals reduces information value.
Dimensions affected: Accessibility, value-added.
Organizational effect: Lack of computing resources limit the quality of information.
Warning signs: User complaints about computing resources and their reliability.
Patches: Provide more computing resources as consumers complain or have consumers pay for their own computing resources.
Problems with patches: Computing resource allocation becomes a political process that lacks a rational basis.
Real solution: Develop technology upgrade policies so consumers know when to expect more resources.

terminals to all employees, which reduced access to information and reduced productivity.

It is likely that there will always be requests for newer, faster, better-equipped computers with higher bandwidth communication lines. As users complain, this pothole is patched by providing more computing power.

A longer term solution is to develop technology upgrade policies. For example, some universities have decided that computers in student labs should be upgraded every three years to the prevailing standard computer system. In addition, funding for more computing resources may be allocated to consumers' budgets to ensure the best use of funds by consumers. More efficient use of existing computers is likely when consumers are charged for computer use.

Until recently, there has been little awareness of the pervasiveness of IQ problems and their severe financial and operational costs to organizations. With the increasing dependence of organizations on the quality of information for managerial and operational decision making, patching IQ potholes is no longer a viable approach. Patching potholes is a reactive response. Organizations must learn to recognize the signs of potential pothole development and proactively develop solutions before problems arise. This requires knowledge of information manufacturing processes and understanding about why these processes perform or fail to perform as they should.[11]

Organizations that attend to the warning signs of IQ problems will provide their consumers a smooth road to high-quality information. Just imagine springtime in New England without potholes. Imagine being able to focus on where you are going and why, not the road. ❖

References

1. D.P. Ballou et al., "Modeling Information Manufacturing Systems to Determine Information Product Quality," *Management Science*, 1997, to appear.
2. R.Y. Wang, V.C. Storey, and C.P. Firth, "A Framework for Analysis of Data Quality Research," *IEEE Trans. Knowledge and Data Eng.*, Aug. 1995, pp. 623-640.
3. R.Y. Wang and D.M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers," *J. Management Information Systems*, Spring 1996, pp. 5-34.
4. E.W. Deming, *Out of the Crisis,* MIT Center for Advanced Engineering Study, Cambridge, Mass. 1986.
5. J.M. Juran and F.M. Gryna, *Quality Control Handbook, 4th Edition,* McGraw-Hill, New York, 1988.
6. M.B. Miles and A.M. Huberman, *Qualitative Data Analysis: A Sourcebook of New Methods*, Sage Publications, Thousand Oaks, Calif., 1984.
7. R.K. Yin, *Case Study Research: Design and Methods, 2nd Edition*, Sage Publications, Thousand Oaks, Calif., 1994.
8. D.M. Strong, Y.W. Lee, and R.Y. Wang, "Data Quality in Context," *Comm. ACM*, May 1997, pp. 103-110.
9. *Data Quality Control: Theory and Pragmatics*, G.E. Liepins and V.R.R. Uppuluri, eds., Marcel Dekker, Inc., New York, 1990.
10. D.M. Strong and S.M. Miller, "Exceptions and Exception Handling in Computerized Information Processes," *ACM Trans. Information Systems*, Apr. 1995, pp. 206-233.
11. Y.W. Lee, "Why 'Know Why' Knowledge is Useful for Solving Information Quality Problems," *Proc. Americas Conf. Information Systems,* The Association of Information Systems (AIS), Aug., 1996, Phoenix, Ariz, pp. 200-202.

Diane M. Strong *is an assistant professor in the Management Department at Worcester Polytechnic Institute. Her research centers on data and information quality, and she also studies software maintenance and software quality. She has published widely in leading journals. Strong is an organizer of the Conference on Information Quality. She received a PhD in information systems from Carnegie Mellon University.*

Yang W. Lee *is president of Cambridge Research Group, Inc., which specializes in information quality management. Her research focuses on organizational learning, information technology, and information quality management. Lee is an organizer of the Conference on Information Quality. She received a PhD in organizational learning and policy analysis from Massachusetts Institute of Technology.*

Richard Y. Wang *is codirector of the Total Data Quality Management Research Program and an associate professor at MIT Sloan School of Management. He has written extensively on issues related to the development of concepts, models, and methods that are fundamental to the information quality field. He received a PhD from Massachusetts Institute of Technology, Sloan School of Management with an IT concentration.*

Contact Strong at Dept. of Management, WPI, 100 Institute Rd., Worcester, MA 01609-2280, dstrong@wpi.edu; Lee at Cambridge Research Group, Inc., 225 Crafts Rd., Chestnut Hill, MA 02167-1421, ylee@tdqm.com; Wang at E53-320, MIT Sloan School of Management, 30 Wadsworth St., Cambridge, MA 02142-1347, rwang@mit.edu, http://web.mit.edu/tdqm/.