To Brooke Dickson
Office of Information and Regulatory Affairs
Office of Management and Budget
Washington D.C. 20503

informationquality@omb.eop.gov

Ms. Dickson,

As one of the leading authorities on the subject of Information Quality Improvement, I am most please to see this action to help assure the quality of information both maintained for internal use and for information provided to the public and others. Nonquality information has high costs internally in scrap and rework when poor quality information causes processes to fail. But it can have negative consequences to individual citizens whose lives and safety can be jeopardized when they "rely" on faulty information. Best-case scenarios are wasted time and money of trusting citizens. Thank you for your leadership in this area.

I am also pleased to provide comments on the proposed guidelines for implementing Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (Pub. L. 106-554) re government-wide guidelines that "provide policy and procedural guidance to Federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by Federal agencies." As to my credentials, I have specialized in information quality management for over ten years. My bio can be found at www.infoimpact.com > Who We Are > Bios. I have developed a complete methodology for information quality management called TQdM (Total Quality data Management). In it I have applied W. Edwards Deming's 14 Points of Quality directly to information quality. TQdM has been implemented in many organizations worldwide, including in some government organizations. I have described six processes for information quality management in my book, "Improving Data Warehouse and Business Information Quality," (IDW&BIQ) published by John Wiley & Sons, 1999. (The introduction and first chapter can be previewed at www.infoimpact.com, along with reviews by press, quality consultants and quality practitioners.) These address a pragmatic way to realize the components of Section 515. In my comments I refer to specific chapters or pages that apply.

First, I am pleased that the importance of quality information is being formally recognized within the Federal government.

Secondly, the ramifications of "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies" are significant. This effort is much more than simply implementing procedures that "should" make this happen. This requires a significant cultural transformation to a mindset that:
- Information is a product in and of itself, just as a manufactured automobile.

This is a very important procedure. This procedure should be made available not only to external persons, but there should also be a procedure or feedback loop for employees and contractors within the agencies to provide corrected information they may discover to the appropriate process owner. (See IDW&BIQ, p. 39.)

Documents should contain the name and contact information of the line manager responsible for the document for getting corrections to. Each line manager should have an information quality lead to provide quality guidance within that unit and to conduct information quality assessments of the information subject to quality measurement.

"II. Agency Responsibilities 3. Report...the number and nature of complaints...."

This has the potential to create fear in staff. Deming's Point 8 states, "Drive out fear." Point 8 of Information Quality states "Drive out fear of information production." Most of the time complaints are handled by pointing blame at "negligent" staff. Staff must not have fear of being "punished" if they should make mistakes. For staff to do a good job, they must have training and other resources to perform work properly.

Complaints should be seen not for punitive purposes, but as opportunities to identify areas for improvement. The goal is, of course, to have no complaints (depending on the quality standard for the information). Complaints should be used to identify process improvement initiatives to eliminate recurrence of the types of errors. The "Process Improvement" (Plan-Do-Check-Act) process is the most important of the information quality processes. This process "Plans" an improvement by identifying the root cause of the defect(s), defining an improvement(s), then "Does" implement the improvement in a controlled manner, "Checks" to assure the improvement worked without introducing negative side effects, and "Acts" to make the improvement(s) permanent and controlled. Process improvement is described in Chapter 9 "Improving Information Process Quality: Data Defect Prevention," IDW&BIQ, pages 285-310.

"III. Guidelines for Ensuring... 1. Overall, agencies should adopt a high standard of quality..."

This is consistent with Philip Crosby when he says, "there is absolutely no reason for defects in any product or service" ("Quality is Free," p. 47). This applies to information products as well. If information is worth gathering, it is worth gathering correctly and completely.

"Quality is to be ensured and established at levels appropriate to the nature of the information to be disseminated." This statement is correct. Not all information is equally important or with zero-defect requirements. The precision quality characteristic of a "start-time" field is less important for the start of an information meeting, but requires high precision of the start of each component in the sequence of the Space Shuttle launch countdown.

Information should be prioritized into classifications of importance for quality standards based on the impact and costs of errors or omissions. Costs and impact should be measured, or at least carefully estimated, based on its impact within the agency, within other agencies to which the information is to be disseminated, and to the public and private industry to which the information is to be disseminated. A classification for information prioritization is found in IDW&BIQ, page 157. Chapter 7 in IDW&BIQ,

"Measuring the Cost of Nonquality Information," pages 213-235, describes a process for quantifying the costs and risk of nonquality information.

"III. Guidelines for Ensuring...2. As a matter...agencies should develop a process for reviewing and documenting for users the quality (including the objectivity, utility, and integrity)..."

First, I recommend a new word for the term "user." The term "user" implies at best one who is not empowered, or at worst, someone with an addiction to a harmful substance. While popular in IT vocabulary, the term is an anachronism in the Information Age of knowledge work. Peter Drucker, the great management consultant, coined the term "knowledge worker" in the 1950s when he foresaw the impact of information technology on industry and society. (See IDW&BIQ, pages 25 and 494-5, and see the September, 2000 "DM Review" article, "What's In a Name," at www.inforimpact.com > Information Quality Resources > Newsletters and Articles for an electronic version.) In the context used here, the term "information consumer" seems appropriate.

Second, you identify several aspects of quality including objectivity, utility and integrity. These are what I call "characteristics" of information quality and others call "dimensions" of quality. There are several quality characteristics, including accuracy, completeness, precision, nonduplication, equivalence (in redundantly maintained data), timeliness, contextual clarity, derivation integrity, and usability, among others. They are described in IDW&BIQ, pages 141-153. Different types of information will have different important quality characteristics BASED ON the processes the information supports or the decisions the knowledge workers or information consumers make. The pertinent characteristics need to be assessed and reported for the information consumers, in order for them to have confidence in the information, or if there are some problems, the information consumers have the information that tells them, X% of the information is in error (quality), or the information source is one that may contain bias (objectivity) because it is presented from only one perspective. There may be more than one purpose to be made of a collection of information. Quality standards need to be established taking this into account. The principle is that if one purpose has higher costs or risk due to error, that quality standard should be adopted.

Third, it is most important for information consumers to have a "warranty" on the quality of information they are "buying," in the same way they need warranties on manufactured goods. The assessment of quality and posting of reliability is especially important in the information is incomplete, contains uncorrectable inaccuracies, or uses methods of collection or statistical analysis that allow for statistical margins of error. Thanks for this, from the public constituents.

Regarding Enterprise Architecture, it is important to have an Enterprise Information Architecture that allows enterprise data to be defined singularly for the entire organization and to maximize the ability to control information. Because information has utility for multiple purposes, both within and outside of agencies, the data must be modeled and maintained in a way independent of any single purpose or use. This requires quality of data definition and data design. Data should be organized and maintained in databases and collections by subject of information rather than by single

use. Utility of information is a function of the retrieval and presentation that meets specific use. By so doing, the value of the information is maximized. Data that is stored based on one use, is data that is sub-optimized for all uses. This is like the way grocery stores are organized, but category of food groups. Organizing data for one purpose is like organizing the grocery store by recipe!

The First process of TQdM is to measure data definition and information architecture (data model) quality. The characteristics of data definition quality and the assessment process are described in Chapter 5, "Assessing Data Definition and Information Architecture Quality," IDW&BIQ, pages 83-136.

"III. Guidelines for Ensuring... 3   As a matter of citizen review, agencies should establish administrative mechanisms to seek and obtain correction of information maintained."

(Discussed above under II. 2.)

IV. Agency Reporting Requirements. Discussion.

Re: "Monitor agency compliance...consider alleged instances of agency failure to comply, and recommend or take appropriate action." Again, this has the potential to create fear. If everyone has basic training in process improvement actions and are encouraged and empowered to participate, recommend improvements, it creates an environment of teamwork and stewardship of providing quality services to help one's "information customers." Information quality is the responsibility of everyone in every agency. But people must have the skills, training, and incentives to make quality happen. Management must create that environment in their agencies.

"IV.   Agency Reporting Requirements.   3.   The agency must draft a report...providing the agency information quality guidelines...."

The agencies need information quality (IQ) leaders trained in IQ processes and implementation to perform this within one year. These IQ leaders need to help lead the agency in the implementation of IQ processes and to help provide orientation and training to all within the agency. In addition to the book, there are seminars and conferences to assist in this:

- The Ninth IQ Conference, organized by Information Impact International, will be held in Baltimore September 30-October 4. There are tutorials and four tracks of presentations by leading authorities and IQ practitioners who have implemented significant IQ environments. There are six presentations by government entities. The IQ Conference   program   and   details   can   be   found   by   visiting http://www.infoimpact.com/conf.html.

- A public 3-day Information Quality Improvement seminar, developed by myself and co-sponsored by Information Impact International and SAIC will be held in Vienna, Virginia, December 3-5.   The course description and details can be found at http://www.infoimpact.com.   Student   comments   can   be   found   at http://www.infoimpact.com > Who We Are > Endorsements.

"IV. Agency Reporting Requirements. 4. The agency must submit this draft report...and post this report on the agency's web site...."

This is excellent. It is unconscionable the way some organizations have disclaimers denouncing any liability for information errors and omissions. I have called for organizations to not disclaim, but to identify what processes and techniques they do have to assure the quality of information they post.

"IV. Agency Reporting Requirements. 5. On an annual basis...submit a report...detailing the number and nature of complaints...."

I highly recommend this annual report contain, in addition to complaints, statistics on the number and nature of positive improvements implemented, the degree of improvement in quality (accuracy, completeness, timeliness, and other quality characteristics they are measuring) of information. Let us promote the positive side and provide recognition for the good people are doing.

"V. Definitions 1. 'Quality,' 'Utility,' 'Objectivity,' and 'Integrity'..."

These represent a good set of quality characteristics. See comments above. There will be some subsets of quality characteristics that must be considered based on various uses of information. For example, "Quality" consists of accuracy and completeness, and correct precision of data for a use. I call these quality characteristics "inherent" characteristics. If they meet these characteristics the information can support many uses. The "Utility" set of measures deal with what I have called "pragmatic" quality characteristics. The information is useful for my purpose. For example, timeliness (of a weather forecast) or presentation of information in a way that supports my use may be different for some other purpose.

"V. 1. A. "

Your point here about taking into account the perspective of the public is most important. Much data can have significant additional value if considered from, and designed for the public and to capture attributes to support the many potential purposes and uses. This does not mean to capture everything. But processes should capture the additional, valuable attributes to support the majority purposes. The major cost of information acquisition is getting to the point to start capturing the data. The cost to capture each additional attribute or fact of information is incremental from that point. The additional cost may be marginal, but can, in fact, put the agencies in a potential revenue generating position, if it so does to pursue that direction.

Re: "For disseminated information to be useful, the presentation should clearly reflect the quality of the information." In fact, presentation quality is one of the three components of information quality: 1) Definition quality—the data must be defined clearly and robustly; 2) Content quality—the content must be accurate, complete, and have the right precision; and 3) Presentation quality—the information must be presented in a way to allow the knowledge worker to take the right action or make the right decision.

"V   B

This is right on target. The source of information needs to be known for certain information in order for the public to determine whether it is "reliable" or not. For "controversial" information where the information has a degree of subjectivity and there are two distinct sides of an issue, it is best to have each side presented from a source authority of each "side."

"V. 1. B. ii. "
Re: "scientific and statistical information." For reliability of scientific information, the methods of the experiment or observation should be described along with disclosure of any variables that might "bias" the results. The calibration of the measurement instruments should be described along with any degree of error that might be induced by the precision of the measurement instruments.

"V. 1. C. ii. "
Security of the information is crucial. If data is "transferred" from multiple databases, the information consumer should be notified of the "lineage" (all files the data passed through) along with the control mechanisms to assure no data has been dropped from being moved from one file to another, and the control mechanisms to assure no mishandling or miscalculation or error introduction has taken place.

Other comments

- Nowhere did I see a reference of confidentiality or privacy of certain types of information. However, in the same way certain information does need to be protected. This is important characteristic of certain information types. Some individual information must not be distributed to the public. However, that same information may be used in aggregations that can be disseminated to the public without divulging confidentiality or identifying individuals.

- Nowhere did I see directives that provide the requirement for training personnel in basic quality principles and tools to help carry out this directive. Training and education make up two of Deming's 14 Points of Quality. Point 6, "Institute training," says that for workers to do a quality job, they must be trained in how to do it, that they can perform processes in a consistent, repeatable way. Point 13, "Institute a vigorous program of education and self-development," addresses the reality that it is not good enough to know today's skills—we must be continually learning tomorrow's skills. See IDW&BIQ, pages 393-395. Learning quality processes and techniques is an investment in tomorrow's skills. Quality processes and techniques enable everyone to look at the work they perform today—whatever that work may be—and find ways to do it better.

Thank you for this initiative and for the opportunity to comment.

Sincerely,

Larry P. English
President and Principal
Information Impact International, Inc.
871 Nialta Lane, Suite 100
Brentwood, TN 37027

PS: Please forgive any typos and awkward sentences. I just discovered this directive late last week, and have rushed to draft this to get it to you by the deadline. I normally would have my office proof it. Again, thanks. LPE.

Larry English can be reached at Larry.English@infoimpact.com.

About the Author:

Larry P. English is president and principal of INFORMATION IMPACT International, Inc., Brentwood, TN., and is the author of the widely acclaimed book, *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits* (preview it at www.infoimpact.com). English has chaired eight Information Quality Conferences, the next in Baltimore, MD, Sept 30 – Oct 4, 2001.

English provides consulting and education to organizations help them implement effective IQ processes to reduce costs of nonquality data and exploit the value of quality information.