

COMMENTS IN RESPONSE TO "PROPOSED GUIDELINES FOR ENSURING AND MAXIMIZING THE QUALITY, OBJECTIVITY, UTILITY, AND INTEGRITY OF INFORMATION DISSEMINATED BY FEDERAL AGENCIES"

Issued by the Office of Management and Budget, Executive Office of the President
Comments due August 13, 2001

SUMMARY HEADINGS

OVERALL COMMENTS

DEFINITIONS

CITIZEN REVIEW

SPECIAL ISSUES OF SECONDARY DISSEMINATIONS ORGANIZATIONS

SPECIAL ISSUES OF STI

SPECIFIC QUESTIONS ASKED BY OMB

OVERALL COMMENTS

The perspectives expressed in these comments are a synthesis of ideas based on work I have done with federal scientific and technical information managers over many years in our efforts to continually improve the quality and accessibility of scientific and technical information. Government providers of STI have long-standing procedures and processes for ensuring the highest quality of the information they provide. I share the goal of promoting an equally rigorous level of quality across the entire federal government. However, the guidelines proposed by OMB raise many concerns about how this might be accomplished. Although the principles offered are helpful to set the proper context to the quality guidelines, the flexibility creates uncertainty and does not put enough bounds on definitions and processes.

I hope highlighting these themes is useful in OMB's deliberations on the revision of the Quality Guidelines. I appreciate the opportunity to provide input to this important deliberation.

2. DEFINITIONS

Definitions become extremely important in these guidelines because they set the stage for the "citizen review" process and are the basis for agencies to determine what information is covered and the criteria by which they will be judged to meet the guidelines. This could be a process for the public good or could also become a nightmare for bureaucracy in a litigious world where information literacy and information dissemination as well as special interest activity is an increasingly important part of our national fabric.

The guidelines are intended to cover all types of information that the government collects, maintains, and disseminates from weather forecasts to economic statistics. The draft acknowledges that different standards of quality, integrity, utility, and objectivity may be needed for each case. In fact, research and monitoring information is singled out as having a system or standard that may be looked at by other information functions when developing their standards. But the guidelines seem to suggest that agencies must develop their own standards within one year and those standards can be challenged by anyone. It would seem that agencies need more concrete definitions and criteria in the OMB guidance in order to develop these standards for quality. Otherwise, challenges and appeals could proliferate and standards of quality could end up being determined by bureaucratic procedures and legal process.

In dealing with quality issues of scientific data sets, the definition of quality is often addressed in the metadata by giving quality indicators without changing the nature of the original information. Quality indicators provided in metadata, which could include level of peer review, among other things, are under development. This process of characterizing quality in metadata in itself is difficult to define and implement, but it could allow the user to better understand the data and what should be expected of it and what it should be used for. It might help to have a consistent approach rather than unclear definitions across agencies.

Regarding the specifics of the definitions of the terms "Quality," "Utility," "Objectivity," and "Integrity." These are "characterized" rather than clearly and concretely defined. The concepts by which they are characterized remain very vague and open to interpretation. It is recognized that this is a difficult challenge but without better clarity, the guidelines are ambiguous and make them unusable.

Utility – Whether the information is useful to users, including the public.

Many of our information organizations serve a wide variety of users. Specific information may be "useful" to only one of these audiences, while other information products and services may be of interest to all of them. It is unclear what metrics might be used to determine "utility," particularly in view of the potential for citizen review. Adding the concept that information should be helpful to, easily interpreted and clearly understood by **the intended audience** might help focus on the multi-user problem.

Objectivity - Whether the information is "presented in an accurate, clear, complete, and unbiased manner." Further, with respect to scientific research information, the results must be substantially reproducible upon independent analysis of the underlying data.

This issue of reproducibility is problematic in some areas of science. See the item on special issues in STI.

Integrity - Whether the integrity of the information has not been compromised by unauthorized access or revision

Note the use of technological means discussed in Item 4 under "Responsibility and Adjudication in Secondary Dissemination" addresses this.

Transparency. The word transparency should also be defined. This is not a term that is commonly understood.

3. CITIZEN REVIEW

It is clear that many of the concerns, especially related to the citizen review, are inherent in the requirements of the law. The premise of minimizing burden and impact is welcome and one that needs careful attention. The OMB guidelines can facilitate the objectives of the review process with guidelines that set limits to the potential burden on agencies. This is especially true in dealing with "nuisance" or "unsupported" requests by citizens "to seek and obtain correction of information." The guidelines appear to be written as if any challenger has standing and the burden of proof is on the agency to demonstrate "quality". This must be balanced by the responsibilities of citizens not to abuse these guidelines. Therefore, guidance on the "responsibilities of the citizen" in making such requests might help minimize the burden on the nation. It is in this context that the following concerns are expressed.

According to the proposed guidelines, it appears agencies would be required to respond to public complaints from individuals who might consider a particular point of view that is expressed as biased. Given the breadth of the contents of some dissemination organizations and the fact that the complaints could come from any source, whatever their particular level of expertise or own personal biases, agencies could be faced with responding to a large number of complaints that may themselves have limited or no scientific merit. As noted below under Item 4, the traditional role of a library is to make available the full range of information on a subject that meets its criteria for building a quality collection. There are special concerns for such organizations under the current guidelines (see below).

Care must be taken to ensure that the right for citizen review and appeal is used to address the intended issue of data quality, rather than abused to achieve political ends. It should not be *carte blanche* to tie up agency information dissemination products unreasonably. It should not be used to inhibit information dissemination by a few or to stop legitimate progress on issues and the deliberative process. There is some concern that the more critical the agency's data to key societal issues, the more likely there would be challenges to address. Thus, the more relevant the science, the higher the burden of cost and, potentially, the more chilling the effect of a *carte blanche* by the public. Anything the guidelines could include to anticipate these potential problems would make them more useful to agencies as well as serve the public good. A revision of the guidelines could help to ameliorate this potential negative situation.

Problems of Quality Upgrades.

There is a requirement that administrative mechanisms be established so that the public may obtain correction of information maintained and disseminated that does not comply with OMB guidelines. It is important to give some guidance on what is expected. Some sense of the requirements and limits on this might be helpful in the guidelines.

Regarding corrections, an interesting issue arises with regard to such things as errata changes, addendum, superseded documents, and document withdrawal. As users are required to be kept anonymous on the web (persistent cookies are not permitted to be used by Federal agencies), agencies have no way to inform those who already received information when subsequent changes are made to the original document. The guidelines might want to address issues of proactively providing corrections to information and consider this a valid general exception to the use of persistent cookies. When agencies formerly kept distribution lists on its technical reports, it could disseminate corrections easily to recipients of the faulty material. Technology enables all kinds of new opportunities to allow agencies to provide updates on information disseminated.

Unfunded, Burdensome Procedures And Reporting Requirements

In addition to the policy issues raised above, the guidelines raise the potential for significant administrative burdens for agencies. Although some of this is, again, inherent in the law, any guidance that could help bound the additional burden would be helpful.

Government Restrictions on Information Reuse and Quality

Given the current emphasis on ensuring quality of government information, we might introduce another concern. The STI agencies have long voiced concerns regarding the assurance of quality in the reuse of government information by non-governmental entities. If there is a desire to improve and guarantee the quality of government information, we recommend that these

guidelines (and a revision to OMB Circular A-130) provide the means where government agencies can assure that organizations reusing and further disseminating government information are required to meet the same standards, if that information is advertised as “government information or derived from government information.” This is particularly important when such information can directly impact the health and welfare of citizens.

4. SPECIAL ISSUES OF SECONDARY DISSEMINATION ORGANIZATIONS

Agencies disseminate information that they generate and, therefore, they should have responsibility for the quality. However, much information that is disseminated is authored or published by sources outside the government but used and disseminated in governmental processes. For example, science agencies disseminate foreign-produced scientific and technical information in the US as part of international exchange programs.

As the guidelines now stand, no distinction is made between government information created *de novo* by the government, and information created by others and disseminated as a service to citizens -- such as a library might do. As long as the government makes it clear who the producer of that information is (indeed, it should ensure integrity in that regard), and that the government is not responsible for the creation of the information being disseminated, the government should not have further responsibility for the quality of that information.

Within the government there are a number of variations of these types of “secondary dissemination” organizations, such as libraries, information centers, and clearinghouses. Many have missions and responsibilities to disseminate everything of a given class of material, which would potentially not meet the quality “definitions” (see “Library Case”) offered.

It would be helpful to have a clear exemption for secondary dissemination organizations with items they don’t produce internally or, perhaps, the definition of government information should be narrowed. It is too broad when it includes anything agencies process or disseminate.

It should be made clear whether the proposed guidelines apply to the results of federally sponsored contracts or grants where the information products are disseminated outside of the channels of government control; i.e., journal articles by commercial publishers. We would advocate that these also be exempted from conformance to the guidelines as long as the awards of these funds call for the adherence to quality standards which are well established in the research community.

The “Library Case” and Censorship Concerns

Further issues related to the question of secondary dissemination organizations specifically address the role and professional standards of federal libraries, from our National Libraries to federal libraries in departments and field centers.

Open access to scientific and technical information that allows for the scrutiny of peers and encourages reproducibility are necessary to detect and correct poor quality/erroneous data. It has always been the library role to ensure information on all sides of questions and issues are made available as part of our first amendment rights. This is also critical to assure even-handed debate regarding the applicability of specific research findings to social policy. Open and uncensored access to environmental regulatory data, educational practices, energy technologies as well as to health care findings, are essential for a robust American scientific enterprise and an informed citizenry. The guidelines should particularly note the importance of full and open access to information. We must be careful not to burn people at the political stake for advocating in today’s

vermacular like the sun is at the center of the solar system or that the earth is, indeed, round. A metaphor offered by a poetic colleague).

Responsibility and Adjudication in Secondary Dissemination

Of significant concern to some secondary organizations is the issue of responsibility and adjudication. Just as a library cannot be the guarantor of every volume in its collection, neither can information centers be responsible for the currency and accuracy of all the reports in its collection. Secondary organizations handle what originating activities provide in the context of defined mission responsibilities. Although these entities are not responsible for the information content quality, they are, however, responsible for maintaining the integrity of the information provided to them. Much of this can be done by technology; e.g., using a type of hash total for each document that is put into an online electronic document storage system and verifying output documents against this hash.

Secondary disseminators should not be held responsible for responding to citizen complaints about information that was not created by them and should not have to maintain a record of complaints or report about the "objectivity, utility, and integrity" of items within their collections. Clarity on this point in the guidelines would help relieve considerable burden on these types of organizations. In fact the definition of government information may be too broad for the purpose of these guidelines.

5. SPECIAL ISSUES OF STI

One can understand the concerns of both the scientific community and the potentially affected public with regard to quality issues. For the science agencies, much of their scientific data collection and publication is directed toward and used by a technically sophisticated community. Although this information is publicly disseminated and available, the public is not the primary audience.

Peer Review

Data objectivity and integrity in the scientific process are priorities since research results are used to make decisions that affect lives, including everything from engineering bridges to medical diagnosis and treatment. The scientific method is rigorous and objective by its nature. It is not reasonable to try to recreate or match it by any new processes in response to these guidelines. The peer review process is a key mechanism for determining whether the results of scientific research merit publication. Peer review may take the form of editors and reviewers, as in scientific journals, or there may be advisory boards of experts, or there may be internal reviews by management in scientific organizations, which is a form of peer review. Peer review and the open scientific process have held this country in good stead over decades. Although the proposed Guidelines allow agencies to develop their own guidelines using existing processes such as peer review, it would considerably reduce the burden if information that had already been peer reviewed in the scientific community was specifically exempted from the guidelines.

Data Reproducibility

With regard to the issue of reproducibility, for some areas of science, the idea that quality is equated with reproducibility is problematic. For example, scientific information that is event-based and temporal cannot be reproduced. It isn't possible to reproduce precisely the weather yesterday so that the information can be recollected. This is one of the major problems with

massive data collection via satellite or instrument. The precise conditions cannot be duplicated because it is dealing with the natural environment. Precision and accuracy of instruments (as well as the more traditional + or – used opinion polling is important.) The specific reference to reproducibility in scientific research should be reconsidered and redefined.

Timeliness Versus Quality Balance

There are some cases where preliminary release of STI is both appropriate and necessary. In scientific research and communication, the use of pre-prints and pre-print servers (with enabling Internet technology) is becoming increasingly common. This is a case where the purpose of posting on web sites is for the review as well as the dissemination process.

Sometimes it is necessary to make decision based on what information can be gathered in a particularly timeframe. In such cases, the best available information may be as critical, or more critical, than the “best possible information.” In such instances, the guidelines should not impede the release of useful information as long as the preliminary nature of the product is acknowledged. This “preliminary release of data” must be managed in such a way to protect scientists from undue penalties associated with the release of such information.

Surveys and Scientific Research

Survey (collection) information, when subjects are volunteers and the work is being carried out for the purpose of scientific research, should be exempted from the approval requirement since the funding review process would already have evaluated the quality of the data collection methodology.

6. SPECIFIC QUESTIONS ASKED BY OMB

Should OMB guidelines devote particular attention to specific types of information or information dissemination products? If so, identify areas where specific focus should be directed; explain why the focus is needed or is desirable; and describe any guidelines that you recommend for those areas.

See Item 4, Special issues of secondary dissemination organizations and Item 5, Special Issues of STI.

Should OMB develop specific guidelines to address information that agencies disseminate from a web page? Is there any need to adapt these guidelines to the agency use of a web page? If so, what guidelines are needed?

Many of changes in the way information is created, managed, and disseminated have been driven by an increasingly “cradle-to-grave” electronic environment. Today, web pages are simply one form of an information dissemination product. No specific guidance regarding Web site quality should be given through these guidelines.