AMERICAN INSTITUTE OF BIOLOGICAL SCIENCES

3 August 2001

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

RE: Proposed guidelines for ensuring and maximizing the quality, objectivity, and integrity of information disseminated by federal agencies (66 FR 34489; June 28, 2001)

Dear Ms. Dickson,

The American Institute of Biological Sciences appreciates the opportunity to comment on the above-referenced proposed guidelines. The American Institute of Biological Sciences comprises 79 scientific societies with a collective membership of over 240,000 scientists in disciplines spanning all of biology -- from basic to applied, from molecular to organismal, from agronomy to zoology. Many of those scientists are employees of federal agencies while all of us benefit from the scientific information disseminated by federal agencies.

AIBS is eager to work with OMB to assure that scientific data disseminated by federal agencies is reviewed in an appropriate manner. As a preliminary matter, however, we note that OMB has only six weeks between the end of the comment period and the deadline for the issuance of the final guidelines. That is wholly inadequate to address the many substantial problems posed by the proposed guidelines. We suggest that OMB seek permission from Congress to extend the September 30 deadline. Further, we suggest that OMB then consider entering into negotiated rulemaking procedures, to include the scientific community and others who are concerned about the effect these guidelines might have on the free flow of information - which is a basic tenet of Circular A-130 (Section 7b). The scientific community holds considerable expertise on the issue of data quality and could help OMB craft guidelines that meet the directives of the new law without erecting barriers to the dissemination of information.

We agree that the quality of data and data interpretation disseminated by the federal government are important. However, with regard to scientific data, we suggest that the scientific community has well-established procedures and standards that are appropriate for use by federal agencies. In fact, as OMB itself notes in the explanatory text, some federal agencies, including the U.S. Geological Survey, have long-established, rigorous peer review procedures for all information disseminated by that agency. We agree that these agencies should be permitted to continue to rely on these procedures and mechanisms.

1444 Eye Street, NW • Suite 200 • Washington, DC 20005
Telephone / 202-628-1500 • Fax / 202-628-1509
e-mail: admin@aibs.org • web: www.aibs.org

OMB has asked for comments on whether the proposed guidelines strike the appropriate balance between the ability of agencies to disseminate information and the development of mechanisms to ensure that recipients can have confidence in that information and force corrective action on agencies. We do not feel that OMB has achieved that balance. We are extremely concerned about several issues raised by the proposed guidelines. Chief among these concerns are the following:

- The guidelines do not provide meaningful definitions of the key terms "quality," "integrity," "utility," and "objectivity"
- Implementation of the guidelines, without significant safeguards to prevent misuse and abuse of the complaint mechanism, will have a chilling effect on government-generated science
- The guidelines, by insisting that scientific data must be substantially reproducible upon independent analysis of the underlying data, have the potential to cause extreme delay in the publication of scientific data
- The guidelines have the potential to cause significant delays in the publication of science information in peer-reviewed journals

We will address these issues first.

Definition of terms

The guidelines are vague in defining and quantifying the terms "quality," "integrity," "utility," and "objectivity." In fact, the actual guidelines set a standard that contradicts the standard in the explanatory text setting out the "underlying principles." The latter reports that "OMB has designed the guidelines so that agencies will meet *basic information quality standards*." (emphasis added). The former (Section III, 1) states that, "*agencies should adopt a high standard of quality*." We suggest that the "basic standard" is more appropriate.

That same section provides that quality is to be established at levels appropriate to the nature of the information to be disseminated. We are at a loss to understand how quality is dependent on the nature of the information. We appreciate that OMB may be concerned about allowing agencies to define quality in a manner appropriate to kind of information disseminated by that agency, but are concerned that the vagueness of this standard will lead to time-consuming, burdensome disputes - possibly in the form of litigation.

Section V (Definitions) sheds no light on the definition of terms, nor the quantum of each attribute required. It appears that OMB is stating that if the information meets some undefined level of integrity, utility, and objectivity, then it meets the quality standard. We concur that these are components or measures of quality.

That being the case, though, the guidelines set no standard for utility. How many people must find information useful before it is considered to have utility, and over what period of time? The guidelines suggest that utility is measured by "whether the information is useful to all users of the information, including the public." This is a tautology. Presumably, someone who uses information finds it useful. Further, a requirement that the information be useful to all recipients is unrealistic. Some who seek information on a given subject will find some information useful

and other information of little interest, depending on the specific interest of the recipient. If the standard is utility to all recipients, then no information will be disseminated.

With regard to scientific information, and particularly basic scientific research, the utility of information might not be known for decades...or longer. Is an agency supposed to defer dissemination until the utility is known?
The guidelines also fail to define the term "integrity" or describe how it is to be assessed. Parts B and C seem to be alluding to research misconduct in the form of deliberate manipulation or misinterpretation of data. On the assumption that integrity, as used by OMB in these proposed guidelines, refers to the manner in which the data were collected, analyzed, and interpreted, it would seem that integrity refers to the absence of research misconduct. We point out that on 6 December 2000, the White House Office of Science and Technology Policy published government-wide research misconduct standards that were derived from lengthy and arduous discussions among the scientific community and a number of federal agencies. The research misconduct policy, in our view, sets a very high standard for integrity. It seems pointless for OMB to require a separate, undefined integrity analysis. There is no need for two parallel systems to assure integrity. Further, most cases of research misconduct come to light after the data are published, either when someone comes forward to report wrongdoing or when someone tries, but fails, to repeat the experiment. These events might happen years after the publication, if at all. In order to assess integrity, agencies might have to wait indefinitely to see if anyone comes forward with claims of fraud or other research misconduct. As for reproducibility, there are any number of reasons why such an effort might fail without invalidating the results of the original experiment.

We are particularly concerned by OMB's singling-out of scientific information in (V, B, ii). The law that the guidelines implement does not mention scientific or statistical information. Therefore, it would appear that these guidelines go beyond that which was mandated by Section 515. Further, the standards set for scientific and statistical information are extremely problematic.

The proposed guidelines set reproducibility as a standard of quality (V,B,ii,a). This is a problematic standard. First, it appears that the requirement of reproducibility is limited to the analysis, rather than the production of data. The language should be revised to make clear that this is the case. Second, reviewers rarely have access to the experimental or observational data, making a re-analysis impossible. The peer review process involves a scrutiny of experimental and statistical methods, not a repetition of those processes. Second, the burden imposed on federal agencies to analyze all data twice before releasing it would divert significant staff time and budget away from the conduct of other research.

The word accuracy is also used in this section. All observations and measurements carry a degree of imprecision in that is considered acceptable in science. Further, the scientific method allows for some error, by way of eliminating observations that are clearly anomalous ("outliers" in statistical parlance) when conducting statistical analyses. There is no such thing as a perfect dataset. Therefore, the use of the word accuracy must be modified by the understanding that it does not imply complete precision and that some degree of error is scientifically accepted. This

is particularly true in the study of natural systems, which, due the high degree of variability in conditions, cannot be conducted with the same degree of repeatability as controlled studies in laboratories.

The guidelines also demand a lack of bias. In both experimental and observational data collection, bias is addressed with appropriate methodology. It is the purpose of the peer review process to critique study methods. If the peer reviewer considers the study and the report to be acceptable, then it should be assumed that the information is free of bias.

The word "reliable" is also offered as a standard of accuracy for scientific information. However, the guidelines do not define this word or suggest how it might be assessed. Reliability can only be measured after dissemination and use. Further, reliability is dependent on the purpose for which the information is used. If, for instance, a member of the public used National Weather Service long range forecasts to determine when to buy and sell crop futures, it could hardly be suggested that the National Weather Service data were unreliable because the crop future prices did not match the trader's expectations.

The use of the terms "data" and "underlying data" are also problematic. We would suggest that OMB consider adopting the definition of "data" used in Circular A-110 to implement the Shelby amendment. Otherwise, data that are not available under Circular A-110 can be accessed under the data quality guidelines. This problem is eliminated if OMB makes it clear that these guidelines do not apply to grantees.

It is also unclear whether OMB is referring only to underlying data or to the interpretation of data. Interpretation does not fall within OMB's definition of information as "facts, data, or opinions." It is far more difficult to force the definition of quality on interpretation of data.

Finally, we are also concerned about the interpretation of the word, "dissemination." It should be made clear that dissemination is not deemed to occur simply because a study or paper is referred to or included in the bibliography of a government publication or document.

Chilling effect on scientific information

The single most serious provision is that requiring agencies to establish administrative mechanisms to allow "affected persons" to seek and obtain correction of information maintained and disseminated by the agency that does not comply with the OMB guidelines. The potential for abuse of this section by individuals and organizations whose motive is to prevent the dissemination or use of data demands that safeguards be put into place, not only to assure the free flow of information promoted by Circular A130, but also to avoid unwarranted burdens on the federal agencies.

We suggest the following safeguards:

a) This requirement should be prospective only. How can OMB or an agency require that information disseminated 10 years ago meet standards established in 2001? Also, the potential impact of retroactive application is enormous in terms of burden to federal agencies' budget and

staff. As was the case with the Shelby amendment, OMB should make clear in the guidelines that application is prospective only.

b) Second, we submit that under these guidelines, information that does not meet agency standards would not be disseminated. Therefore, agencies should be permitted to respond to complaints with a two-tier process. The first tier would simply require the agency to show that it followed its own established procedures before releasing the data. Only if it can be shown that the agency did not follow these procedures, should the "affected person" be permitted to attack the substance.

c) The "affected person" should not be permitted to attack the substance without a showing that an equally qualified scientist (s), i.e., a peer, has found fault with the integrity or objectivity. Data publication should not be delayed or data retracted everytime someone who is opposed to the potential commercial impact or the potential regulatory use of the data decides to attack the underlying data. There should be a burden of proof on the challenger to show that there is serious scientific merit to the challenge before the agency and the scientist(s) are required to go through this procedure. The experts supporting the challenge, and the challengers themselves should be required to disclose any financial or political interest (other than payment for time and expenses as is routine in the industry) they have in the subject matter to which the data pertain.

In science, findings are challenged by scientists who design and conduct studies and publish their findings, subject to the same rigorous peer-review process as the original research. Without adequate safeguards, these guidelines will undermine the value of scientific information by allowing that information to be assailed without insisting on the same degree of rigor and objectivity required of the original science.

d) The "correction" requirement requires definition. Would the original investigator be required to repeat the study? At whose expense? We suggest that the cost should be shared by the challenger, who, absent this law, would be required to conduct independent studies to address scientific findings. If the challenger can force the government agency to repeat a study, the challenger is, in essence, getting a "free ride" by avoiding the cost of scientific study.

e) Correction in the absence of a serious scientific challenge is, in essence, data tampering. The guidelines, as written, would compel a government agency to alter data or data interpretation, without scientific justification.

Other issues

1. Interaction with the journal peer-review process: This process is the "gold standard" and should suffice to meet the data quality review standards. Although some agencies, including the USGS and the EPA already require internal review prior to submission to journals, a second layer of review - either before submission for publication or after acceptance - could cause significant and unnecessary delays in publication for the author and publication schedules for the journals. Therefore, OMB guidelines should state explicitly that the journal peer-review process or other external peer review processes connotes a satisfactory level of quality and that the

agency need not undertake or procure another review. Further, the guidelines should state that submission to a journal or peer review panel does not constitute dissemination.

2. Other external reviews – many agencies require external peer review prior to publication. These external peer reviews are the equivalent of the journal peer review process and should be considered *de facto* to meet the data quality standard.

3. Qualification of reviewers: OMB and agency guidelines should make clear that the reviewers must have sufficient credentials.

4. Reporting requirement: We suggest that OMB require agencies to report the cost of compliance, particularly with regard to the complaint resolution process. Congress should know the cost of this measure, much as Congress insists on an assessment of the costs of other regulations.

5. Applicability to grantees: It would appear from the language of the Paperwork Reduction Act and Circular A-130 that this additional requirement will not apply to recipients of federal grants. We suggest that OMB make this clear in the guidelines. Also, OMB should make clear if, and under what circumstances, the guidelines will apply to contractors.

We thank you for your careful consideration of our comments, and hope that this analysis proves helpful to you in developing the final version of the guidelines.

Sincerely,

Ellen Paul

Ellen Paul
AIBS Public Policy Representative