Programs

Government Relations

Government Relations

Home
About
S&T Newsletter
S&T Bill Trackers
Top Issues
Policy Briefs
Publications & Events
AAAS Statements & Letters
Policy Alert Blog
R&D Budget Analysis
Working With Congress

AAAS Policy Brief: Data Quality

Issue Summary | **Resources**

Contents

Introduction

Inserted into the Fiscal Year (FY) 2001 Omnibus Appropriations Bill was a two sentence rider concerning information released by the federal government. The Data Quality Act, as it later came to be known, tasked the Office of Management and Budget (OMB) with ensuring that federal agencies implement data quality standards and public correction mechanisms for the information they disseminate. The initial Act and OMB's subsequent implementation stirred considerable controversy within the scientific community, who saw the Data Quality issue as an attempt by businesses to hamper the government's regulatory apparatus.

Back to top

Origins of the Data Quality Act

Paperwork Reduction Act of 1995

The origins of the Data Quality movement can be traced back to the historic Paperwork Reduction Act (PRA) of 1995. At that time, the rapid growth of the internet was steadily increasing the capability of government agencies to disseminate information to the public. Encouraged by these changes, yet still cautious of the new medium,[1] lawmakers added the following section to the PRA:

> With respect to information dissemination, the Director [of OMB] shall develop and oversee the implementation of policies, principles, standards, and guidelines to--(1) apply to Federal agency dissemination of public information, regardless of the form or format in which such information is disseminated…

In order to address some of these new information sharing issues, OMB subsequently revised the policy document Circular A-130, "Management of Federal Information Resources." Some in Congress, however, deemed the changes insufficient.

Back to top

FY 1999 Omnibus Bill

Several years later, in the House report on the FY 1999 Omnibus Appropriation Bill for OMB, the following language appeared:

> The Committee urges the Office of Management and Budget (OMB) to develop, with public and Federal agency involvement, rules providing 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.....The OMB and agency rules shall contain administrative mechanisms allowing affected persons to petition for correction of information which does not comply with such rules; and the OMB rules shall contain provisions requiring the agencies to report to OMB periodically regarding the number and nature of petitions or complaints regarding Federal, or Federally-supported, information dissemination, and how such petitions and complaints were handled...

The requests did not make it into the official Conference Report,[2] and OMB chose not to comply by the stated deadline of September 30, 1999. This decision upset several legislators and industry lobbyists who remained determined to promote a Data Quality law.

Back to top

FY 2001 Omnibus Bill

Several years later, in a move reminiscent of the Shelby Amendment, Rep. Jo Ann Emerson (R-MO) of the Appropriations Committee inserted the Data Quality Act into the FY 2001 Omnibus Bill without debate or discussion. Her rider, added as Section 515 of the Treasury and General Government Appropriations Act, rephrased the previous requests to OMB, but added to them the full force of law.

Many environmental, scientific, and public interest groups were dismayed to hear of the new law, which they were never given a chance to oppose. Suspicious of the motives and origins of the Act, which to them seemed dangerous and unnecessary, they launched a campaign against what they considered an industry-backed attempt at deregulation.[3]

Back to top

OMB's Proposed Guidelines on Data Quality

As mandated by the Data Quality Act, OMB subsequently crafted agency information regulation guidelines and posted them in the Federal Register of June 28, 2001. The proposal advised federal agencies to "adopt a high standard of quality" by ensuring the objectivity, utility, and integrity of information they disseminate.

In order to comply, agencies were required to disclose the sources of their information (objectivity), while ensuring that data stay protected from unauthorized tampering (integrity). In addition, scientific information had to be generally useful and "substantially reproducible upon independent analysis of the underlying data."

OMB also incorporated the administrative measures laid out in the Data Quality Act. These included requirements for agencies to develop data quality reviewing procedures as well as public correction mechanisms for inadequate data. The Chief Information Officer (CIO) of each agency was held responsible for dealing with information quality complaints and ensuring their proper documentation.

A deadline of one year was set for each agency to develop appropriate guidelines.

Back to top

Public Response to OMB's Guidelines

Correction Mechanism

During the legally-required public comment period, OMB received 100 responses, mostly from academic and scientific groups.[4] Many of the respondents worried about the public correction mechanism and wondered where the burden of proof would fall in data quality disputes. Some scientific and watchdog groups felt "the burden of proof should lie with those making the challenge"[5] and that "the government should invariably err on the side of providing public access."[6] Fearing frivolous complaints, especially in the environmental arena, they suggested OMB require proof of scientific merit and disclosure of political and economic ties for all data quality cases.[7]

Some groups like the U.S. Chamber of Commerce, however, felt exactly the opposite. They believed that "the burden should not be on the private sector and the general public to seek out and investigate the quality of the information being released by federal agencies…Congress intends this burden to fall on the federal government."[8]

Back to top

Utility

Parties on both sides of the debate tended to agree on the need for clearer definitions of key terms in the proposal. The vague demand for utility of data, for instance, drew criticism from scientists because data are often only useful to a few people (or only in the far future).[9]

Back to top

Reproducibility

Perhaps the most debated aspect of quality was the notion of "reproducibility." As the Association of American Medical Colleges (AAMC) pointed out, scientists generally consider peer review to be the "gold standard" for quality, not reproducibility.[10] To add to this, computer models and long term health studies may not even be reproducible.[11] AAAS, questioning whether the entire Human Genome Project should indeed be repeated, concluded that the reproducibility requirement "oversimplifies the work of many research fields."[12]

Delving even deeper into the possible harms of the reproducibility requirement, the American Psychological Society (APS) warned that "many independent analyses of underlying data would breach the privacy protections in place that are required before research is allowed to proceed."[13] This might inhibit subjects from participating in critical research.

Back to top

Associated Costs

The irony of the Paperwork Reduction Act's role in the Data Quality proposal was not lost on everyone. The considerable amount of paperwork necessary to comply with OMB's requirements represented an "unfunded mandate…[that] will siphon off agency resources from other activities…"[14] Nowhere did the proposal mention how any of the oversight work and data correction would be paid for. In response to this complaint, John Graham, Director of OMB's Office of Information and Regulatory Affairs, countered

that the proposal provided a mechanism to dismiss frivolous complaints, and that in the long run agencies would save money by decreasing judicial and political opposition.[15] Still, some commenters requested that OMB estimate the cost and administrative burden of compliance for federal agencies.[16]

Back to top

Enforcement

Looming above all of these complaints were questions as to the strength of the Data Quality guidelines. OMB Watch, a public interest group, argued for weak enforcement, claiming that "although Section 515 requires an 'administrative mechanism' to permit correction of information not in compliance with the guidance, nowhere does it contemplate that this mechanism has anything to do with withholding or delaying dissemination."[17] Groups concerned with public health agreed, calling for consideration of medical emergencies. Without alterations, concluded the AAMC, the initial proposal "could seriously threaten public health and safety by tying the hands of the agencies with an unrealistic and impossible standard of 'quality'"[18]

The U.S. Chamber of Commerce, on the other hand, advocated a strong Data Quality requirement. They suggested that OMB screen agency data rules and provide more guidance on correction mechanisms.[19] Supporting this stringent view was the Center for Effective Regulation (CRE), who lobbied for a stronger proposal. They called for replacing the weak 'should's' of the proposal with stronger 'shall's' in addition to including in the scope of the rules information prepared for a federal agency by non-governmental organizations.[20]

Back to top

OMB's Final Guidelines on Data Quality

Burden of Proof

After considering comments from the public, OMB issued a revised proposal in the September 28, 2001 Federal Register. On the crucial issue of burden of proof for the correction mechanism, they seemed to come down on the side of federal agencies and the scientific community:

> Agencies, in making their determination of whether or not to correct information, may reject claims made in bad faith or without justification, and are required to undertake only the degree of correction that they conclude is appropriate for the nature and timeliness of the information involved...

This leeway also allowed agencies to appropriately deal with research subject confidentiality. Additionally, permission was granted to the agencies to place an employee other than the CIO in charge of the oversight process.

Back to top

Reproducibility Revisited

After receiving many comments on the vague notion of reproducibility, OMB attempted to clarify the requirement. In the revised proposal, replication of original data was only required for "influential" information which might have an important effect on public or private policies or technologies. In addition, formal and independent peer review was deemed sufficient in most cases. Finally, OMB reiterated that the call for reproducibility would not override other confidentiality protections.

Acknowledging the complexity of the issue, however, OMB asked for further comments on the reproducibility standard.

Back to top

Response to Further Comments

After receiving responses, OMB issued its final guidelines on January 3, 2002.[21] Based on the feedback it received, OMB made several notable changes to the guidelines. Peer review was considered sufficient in most cases, but it was still "rebuttable" if questioned by a petitioner. Similarly, publication in a scientific journal was not deemed sufficient to guarantee quality.

In making these recommendations, OMB took into consideration the difficulties of replicating certain types of data:

> Agencies may identify, in consultation with the relevant scientific and technical communities, those particular types of data that can practicably be subjected to a reproducibility requirement, given ethical, feasibility, or confidentiality constraints.

OMB also pointed out that it was not necessary to actually reproduce data, but rather to ensure transparency of methods and to show that the data could be reproduced. As to the question of who should do the reproduction when necessary, OMB suggested "a qualified party, operating under the same confidentiality protections as the original analysts." OMB also encouraged agencies to perform "robustness checks" to see "whether a specific statistic is sensitive to the choice of analytic method."

OMB justified these requirements by praising the ideal of data transparency. "The more important benefit of transparency is that the public will be able to assess how much an agency's analytic result hinges on the specific analytic choices made by the agency."

Finally, OMB took into consideration public health concerns, making allowances for "the timely flow of vital information from agencies to medical providers, patients, health agencies, and the public." To protect these concerns, the proposal provided for an emergency waiver process for extreme circumstances.

Back to top

Implementation of OMB's Guidelines on Data Quality

Following OMB's adoption of the finalized Data Quality guidelines, agencies were allowed one year to comply. Each agency drafted individual Data Quality regulations and submitted them to OMB, while at the same time providing a chance for public comment.

Since the implementation of the Data Quality Act, scores of formal petitions have been filed. The Center for Progressive Regulation and the Center for Regulatory Effectiveness provide detailed accounts of the current state of these Data Quality cases from two opposing perspectives.

Back to top

Endnotes

[1] In a 2002 speech, John Graham, Administrator of OMB's Office of Information and Regulatory Affairs, proclaimed that "the quality of information disseminated via agency web sites was a particular concern at the time."
[2] A conference report is "the final version of a bill proposed by House and Senate conferees." C-SPAN Glossary
[3] See, for instance, the Center for Progressive Regulation's perspective on the Data Quality Act.
[4] The "Public Comments and OMB Response" section in OMB's explanation of the Data Quality guidelines.
[5] Comments from the American Psychological Society (APS).
[6] Comments from OMB Watch.
[7] Comments from OMB Watch.
[8] Comments from the U.S. Chamber of Commerce.
[9] Comments from the Federation of American Societies for Experimental Biology (FASEB).
[10] Comments from the Association of American Medical Colleges (AAMC).
[11] Comments from the American Association for the Advancement of Science (AAAS).
[12] Comments from the American Association for the Advancement of Science (AAAS).
[13] Comments from the American Psychological Society (APS).
[14] Comments from the Center for Progressive Regulation.
[15] From a 2002 speech by Graham.
[16] Comments from the Association of American Medical Colleges (AAMC).
[17] Comments from OMB Watch.
[18] Comments from the Association of American Medical Colleges (AAMC).
[19] Comments from the U.S. Chamber of Commerce.
[20] Comments from the Center for Regulatory Effectiveness.
[21] OMB made subsequent corrections in a June 13, 2002 document and an Oct 1, 2002 document.

Updated April 7, 2005