

## Going Beyond the Law: It's not just the right thing to do—our success depends on it

By Gerald W. Gates  
U.S. Census Bureau<sup>1</sup>

Keywords: Ethics, Privacy, Data Stewardship

Statistical agencies are subject to laws that ensure that personal information is protected from unauthorized disclosure and use and thereby maintain the trust of respondents by putting them in control over who sees their information and determine how it is used. Without this trust, response and, ultimately, the quality of the data will suffer. These laws have always been supplemented by professional ethical codes that provide some direction when the law is not specific. Today, as technology changes agencies' ability to gather, process, link, tabulate and disseminate information, it is even more critical to consider ethical codes of practice when management decisions impact individual privacy.

I have on the wall of my office a cartoon clipped from the newspaper that shows a traveler entering a vast wasteland. He happens upon a signpost that directs him left in the direction of "legally right" and right in the direction of "morally right." The caption reads "Inevitable Intersection on the Road of Life." We are all faced with these decisions and the choice we make can have dramatic impacts on our future. For statistical agencies the decisions we make can have real consequences on real people and, ultimately, will impact the ability to do our work. Following just the law may be expedient in the short-term, but establishing and following an ethical framework will help ensure that agencies make decisions in a way that supports their long-term viability.

The U.S. Census Bureau operates under several statutes that protect privacy and confidentiality. Title 13, United States Code, requires that the agency maintain confidentiality without exception and limits data use to the statistical purposes for which it was collected. The Privacy Act of 1974 requires that the agency inform the respondent about the purposes and the uses of the data and publish information about the existence of databases maintained on individuals. The Paperwork Reduction Act of 1980 also requires informing respondents about the

purpose, use and protections afforded the data and requires agencies to limit direct collections. The E-Government Act of 2002 requires that agencies conduct Privacy Impact Assessments, requires agencies to ensure adequate security for sensitive information, and permits limited sharing of statistical data while providing uniform protections for all data collected by statistical agencies.

Codes of professional ethics and guiding principles are well established and, in recent years, have been updated to reflect the impacts of technology. These codes are supported by legal requirements but extend those requirements to support fairness, honesty, integrity, and equity. They document the shared professional values and experiences of one's colleagues. They are designed to provide high standards and the criteria to measure performance. Finally, much like laws, they strive to support uniform practices across the profession. One of the first sources of professional ethical values for statisticians related to privacy is found in the Privacy Protection Study Commission Report of 1977, that highlights the importance of functional separation as a fundamental protection for limiting the use of statistical information. That report also set a very high standard for statistics that justified many of the exemptions afforded these data by the Privacy Act and subsequent legislation. Namely, the report noted that statistics are not concerned about the identity of individuals (i.e., individual identities are not important to statistics). Other notable ethical guidelines related to privacy for statistics include the International Statistical Institute's Declaration of Professional Ethics (1985)<sup>2</sup>, the National Academy of Sciences' report on Privacy Lives and Public Policies (1993)<sup>3</sup>, the United Nations' Fundamental Principles of Official Statistics (1994)<sup>4</sup>, and the American Statistical Association's Ethical Guidelines for Statistical Practice (1999)<sup>5</sup>.

Given that we have laws and ethical codes, it would seem that decisions related to collecting, processing, linking, and disseminating data are

pretty straightforward. On the contrary, these decisions are complicated by tensions created when trying to maximize response while preserving privacy and when trying to protect confidentiality and meet the needs of data users. Following ethical principles may mean giving up some control over program decisions, as well as adding complexity to survey procedures that impact costs.

Taking the “glass is half full” approach, we could also view the application of ethical principles as supporting the work of the agency. Defining one’s ethical principles and the policies and practices that apply them can be persuasive to a reluctant public, especially as they are bombarded with news about how technology, as exemplified by Internet scams and data mining, is making privacy obsolete. Adopting privacy principles and the framework to apply them accomplishes several objectives. It ensures consistent decision making that focuses on how the public may perceive the action. It promotes transparency in the operations of the agency. It minimizes the risk of surprises by addressing potential problems before they occur. Finally, it promotes a long-term view that will sustain the agency into the future.

At the Census Bureau, ethical principles are applied through our Data Stewardship program, that recognizes the legal obligations to collect and disseminate high quality data while respecting individual privacy and protecting respondent confidentiality. Importantly, Data Stewardship extends the legal requirements to include: ethical responsibilities to limit collections to what is a Necessity for our mission; Openness about our policies and processes; Respect of our respondents; assurances that personal information remains Confidential. Each of these values is expressed in terms of the agency’s Privacy Principles. A formal governance structure is in place to develop and set policy related to these principles. The program is led by the Data Stewardship Executive Policy Committee, chaired by our Deputy Director and supported by three staff committees: The Privacy Policy and Research Committee, the Disclosure Review Board, and the Committee on Administrative Records Policies and Procedures. Information on the Census Bureau’s Data Stewardship program is found at <http://www.census.gov/privacy/>.

Data Stewardship policies are relevant to management decisions related to data collection, data processing, and data dissemination. For data collection, policies address informed consent, respect of respondents’ privacy, and consideration of the sensitivity of certain questions for certain individuals, and the need for safeguarding personally identifiable information. Ethical considerations include being respectful of the person’s time and considering responsibilities to children and other sensitive groups.

For data processing, Data Stewardship means limiting access to personally identifiable information, to ensure that only those with a need-to-know can see the data. This is accomplished through formal employee training, prohibitions on unauthorized browsing, and use of access control lists and file level auditing. Where record linkage is involved, policies provide for project-level review and approval, compliance with formal agreements, and use of protected control numbers instead of personal identifiers. Ethical considerations include being transparent about these activities and alleviating fears that these linked data will be used for non-statistical purposes.

Data dissemination decisions involve the determination that published data are safe from re-identification by potential intruders. All data products and public releases must meet standards established by the Disclosure Review Board. Proposed data releases that do not obviously meet those guidelines are reviewed and approved using a checklist developed by the Census Bureau and subsequently adopted for government-wide use by federal statistical agencies. Where useful data cannot be publicly released due to disclosure concerns, options for controlled access at Census Bureau secure remote sites is considered if the proposed projects meet legal requirements under Title 13. Ethical concerns include addressing the perception that published data may not appear to be safe. Also, when custom data products are created for a requestor from nonpublic or public data files, the Census Bureau’s policy is that these requests and the data themselves may be subsequently shared with anyone else as a public data product. Openness about the activities ensures that the public is afforded the opportunity to comment on the statistical uses of these data as they may impact themselves as members of a group.

Despite the protections afforded through the Data Stewardship program, privacy issues remain illusive and require continual review and attention. In my opinion, the five most significant privacy issues facing statistical agencies today relate to administrative records and record linkage, government reuse of private sector data, data breaches, researcher access to non-published data, and concerns about group privacy. It seems like every day the public is exposed to media reports of government agencies mining personal records, sharing data, and reusing private sector data. Concerns raised by these events present great challenges in convincing potential respondents that we are good data stewards.

I will discuss briefly two of these issues as they affect federal statistical activities. First, regarding administrative records and record linkage, the main privacy concerns focus on the level of consent required or expected by the individual whose information is linked. Is the right to consent diminished by the fact that legal protections limit the risk of harm? Is the agency trusted to comply with the protections and use limitations they promise? Is the government trusted to continue this promise into the future? Finally, are the file existence and its uses transparent to those persons involved?

The second issue is group privacy. The law and the Census Bureau's mission require the dissemination of data to the public. The law also protects the confidentiality of individuals who participate in surveys and censuses, but does not protect them as members of the groups to which they belong. In fact, statistics are intended to be descriptive of groups (e.g., young/old, urban/rural, Black/white). Program and policy uses of statistics can either benefit or harm individuals, not for who they are but because they belong to a class of people who are impacted by government programs. Statistics are a public good and, in my opinion, should not be regulated. Nevertheless, since groups can be singled out in these data it is important that the existence of the data not be a secret. As I mentioned, Data Stewardship policies provide that we publicly release the names of requesters of special data products and the nature of data they requested, so that a public debate can occur over the potential program uses.

In conclusion, there are many tough choices agencies make that impact their ability to collect and disseminate high quality statistical data. These choices invariably also impact the public's view of how its privacy is protected. Since perceptions may have real consequences, we need to make sure we consider them in any decisions to collect, process, or disseminate data. We do that by complying with the law, adhering to ethical principles, and continually monitoring public opinion to see how we are doing.

### Notes and References

<sup>1</sup> *The views expressed are those of the author. This report is released to inform interested parties and to encourage discussion.*

<http://www.census.gov>

<sup>2</sup> *International Statistical Institute's Declaration of Professional Ethics (1985)*

<http://isi.cbs.nl/ethics.htm>

<sup>3</sup> *National Academy of Sciences' report on Privacy Lives and Public Policies (1993)*

<http://newton.nap.edu/catalog/2122.html>

<sup>4</sup> *United Nations' Fundamental Principles of Official Statistics (1994)*

<http://unstats.un.org/unsd/goodprac/bpabout.asp>

<sup>5</sup> *American Statistical Association's Ethical Guidelines for Statistical Practice (1999)*

<http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics>