

*Comments from Sarah Cohen, Knight Professor of the Practice of Journalism and Public Policy, Duke University*  
*"From Town Criers to Bloggers: How Will Journalism Survive the Internet Age?"*  
*March 2010*

These workshops at the Federal Trade Commission reflect an important recognition that the federal government itself can, indeed, have an effect on the amount and quality of public affairs, accountability and investigative journalism about government and powerful institutions.

Last summer, I became the first Knight professor in a new field tentatively called "computational journalism," which is geared at finding, commissioning and developing methods and tools to reduce the cost and difficulty of this expensive and important genre in reporting. If we needed no reminding of the need for original reporting from working journalists, a recent ad-hoc study of 121 stories on one subject in Google News showed that only 13 had any original content, and those came from only seven sources.

In my 15 years as an investigative reporter and editor, including more than a decade with The Washington Post, I depended on both technology and public records to explore and document stories of importance to the public. Examples in federal programs included a yearlong examination of waste and duplication in farm subsidy programs, spending by local and state governments of homeland security grants, and federal oversight of local water systems.

My comments on the future of journalism come from that background, and as such identify some concrete steps the federal government might take to encourage this kind of journalistic enterprise without influencing its content.

From my own experience and from conversations with colleagues still pursuing public interest journalism, I've come to the conclusion that the single biggest step the government can take is to improve and institutionalize the administration of the federal Freedom of Information Act, and encourage its adaption in state and local government programs that use federal funding. The effort, time and often cost involved in obtaining public records remains one of the highest barriers to public affairs journalism, not only but especially when that pursuit involves legal action.

My comments are not geared at evaluating the Act or President Obama's effort in transparency, but are focused instead on specific actions that could be taken and would make the process less dependent on presidential commitments or actions:

- Institutionalize the release of common records used to monitor agency activity. Currently, FOIA requests are required to obtain records like correspondence logs, desk calendars of cabinet and sub-cabinet level officials, payroll records of political appointees and contracts, grants and their audits. Although these could fall under a broad definition of "frequently

request records” in the 1996 E-FOIA law, I know of no agency that routinely releases any of these outside the FOIA process. Anticipating the release and planning for them as part of the routine administration of the agency would not only reduce work and effort for reporters, it could also reduce backlogs in FOIA offices.

- Centralize, update and document the “major information system” section of agency FOIA websites, including dates of most recent updates and names of officials to call or write for further information. Some of the difficulty and cost in FOIA stems from a requester’s incomplete understanding of how records are held, where they reside and what they contain. Clearing this up in advance will smooth the path toward transparency.
- Build openness into new or re-tooled administrative information systems and reviews of records systems, even when they were not intended for public use. Examples include personnel records, calendar and email systems, spending records, client records, inspection and compliance records and benefit records. New systems must already – and appropriately -- be certified as secure and able to protect private information. But there is no certification that the public portions of the records can be extracted. The same problem exists in the routine re-approval of information collections by the Office of Management and Budget under Paperwork Reduction Act rules. Making transparency one of the priorities in these reviews will reduce the cost and difficulty of extracting information later. A few key examples:
  - Email systems, which are still not used to archive “records” under the meaning in the law. There is no efficient way to search them in most federal agencies.
  - Contracts and grants, which comingle proprietary commercial or competitive information with the publicly releasable pieces. For instance, even applications from local agencies applying for the Education Department’s Race to the Top funds were redacted before release. Clearly segregating the portions of these records that may not be made public in application forms, requests for proposals and other documents would go far in reducing backlogs and unacceptable redactions.
  - New data systems, which are not designed for maximum transparency. Changes might include replacing private or proprietary unique identifiers such as Social Security or DUNS numbers; adding indicators on comment or narrative fields that allow an agency employee to mark individual records that cannot be released in full and must be reviewed; and splitting fields into releasable and un-releasable parts, such as 5-digit Zip Code distinguished from addresses or the extended 4 digits when necessary. This practice was instituted in Florida in the 1990s, resulting in much cleaner and

clearer release of databases while reducing the effort in extracting the records.

- Remind those who do business with the government that, with the exception of the most private personal information such as tax and health records, their information may be subject to Freedom of Information Act requests. This would reduce the requirement in some agencies that subjects of records review a release after the request has been made but before the records are produced. This is especially common with business records of contractors and subjects of inspections or consumer complaints. Even when the records are public by law, such as the Consumer Product Safety Commission complaint system, agencies will not release any information not because it is private, but because they have alerted the subject of a complaint.
- Create a clearer path to resolve FOIA requests within agencies that have languished for more than a few months without resorting to lawsuits or an appeal to the Office of Government Information Services. The most frequent complaint from reporters is that they cannot find anyone in an agency to look into their case or provide information or estimates on when the documents might be produced or how to reduce the effort involved. Last year, a reporter in Kansas City received a letter providing an estimate of the cost to produce 700 letters that were responsive to a request on correspondence about human trafficking. He was unable to get the records because he could not find anyone to take his call or e-mail agreeing to the fee. Delay and an inability to negotiate and discuss requests with agencies can lead to inaccurate stories, undiscovered stories and a reluctance to use public, rather than leaked, records.
- Encourage state and local governments to follow suit by including transparency requirements in federal grants and ensure that state and local governments release records that may have federal implications. For example, a reporter in North Carolina must approach the regional Health and Human Services office in Atlanta to obtain nursing home inspection reports conducted by the state agency for Medicare even though the records are in the state's files and are easily produced.

There is another, less obvious and more controversial, way that the federal government could help reduce the cost and difficulty in public affairs reporting. That is, by treating public affairs reporting as a "public good" for the purpose of using the fruits of federally funded research and possibly even for funding itself.

As an example, one issue that raises the cost and difficulty of investigative reporting is a barrier that most other industries have removed: The difficulty of scanning paper documents and converting them into searchable form. This OCR

process can put such a strain on the small computers or laptops available to most reporters that, for example, it could take 4 days of a computer's time to recognize the 10,000 pages of the calendar released by Secretary Clinton from her days as First Lady.

In state and federally funded computing centers may have both the technology available and the occasional excess computing capacity to make the service available to reporters at little or no cost. Treating the non-commercial side of public affairs reporting as a "public good" might make facilities such as these available to journalists. I have suggested using the same definitions that distinguish "commercial" from "public interest" FOIA requests as a way to possibly open up these taxpayer-funded facilities.

Along a similar vein, some federally funded research projects result in software and tools that are provided back to the federal government or are used in academia, but are sometimes not freely available. For example, a recent National Science Foundation research project on text analysis of public comments for regulatory reviews in the federal government is owned and released by two universities. However, in its present form, it cannot be adapted to the needs of reporters and is not available in a form that could be used to analyze, for example, local government or other important text collections. The need to attract developers by offering them patents and copyrights in cases like this might be balanced by some acknowledgement that the public has already paid for the development and should have access to the fruits of its investment, at least for activities that meet a common definition of a public good.

I am one of the journalists who is uncomfortable asking for federal funding, special rules or other favors from government. However, I have worked for publications that have accepted postal rate reductions and other content-neutral assistance. In that spirit, the last consideration might be for the funding of research into methods that journalists and other fields need to share, such as improvements in methods of analyze large documents collections, network analysis, the needs of historians to understand original documents, or devices for doctors to improve their note-taking in an age of advancing medical technology. All of these fields have access to federal grants, while research by journalists do not. It might be useful to consider adding transparency and research into public affairs – including journalists, citizens, NGOs and others – as a category of federal assistance that does not dictate content, but can ease the use of technology for publishing.