



**PUBLIC.RESOURCE.ORG** ~ *A Nonprofit Corporation*

---

## **Open Source America's Operating System**

**"It's Not Just A Good Idea—It's The Law!"**

July 14, 2009

The Honorable Vivek Kundra, Chief Information Officer  
Office of Management and Budget  
Executive Office of the President  
The White House  
Washington, D.C. 20500

The Honorable Aneesh P. Chopra, Chief Technology Officer  
Office of Science and Technology Policy  
Executive Office of the President  
The White House  
Washington, D.C. 20500

Dear Mr. Kundra and Mr. Chopra:

I am writing to request your assistance in making available at no charge and for bulk access two of the most important legal databases maintained by the executive branch:

- Publications in the Federal Register System, maintained by the National Archives and Records Administration (NARA).
- Patents, maintained by the U.S. Patent and Trademark Office (USPTO).

While the U.S. government maintains a minimal web presence for both databases, those web sites are only useful for casual browsing. In both cases, the underlying source code for the documents is only available for substantial fees.

A yearly subscription to the Code of Federal Regulations for bulk access to the "SGML" source code with images is \$17,000/year. The same \$17,000 fee applies to other NARA databases such as the Federal Register. While there are PDF versions of the Federal Register and text versions of the Code of Federal Regulations available for browsing, it is impossible to easily download them in bulk, and the underlying source code which could be used for creating new versions of these documents is prohibitively priced.

Likewise, the U.S. Patent and Trademark Office makes a web site available for casual searching and browsing, but the only bulk access to patent data is limited to the first page of a patent. To get the full text of current and historical patents requires a very substantial fee. For example, the Patent Grant Data/XML v. 4.2 ICE (Current Calendar Year Subscription) (EIP-5300P-OL) costs a breathtaking \$39,000.

These fees are so substantial that they actively discourage the use of these key U.S. government databases by public interest groups and scholars, limiting access to a few well-heeled corporations. In particular, at Public.Resource.Org, we would make much more extensive use of these databases if we could afford access, helping fulfill our mission of making America's primary legal materials available to the public.

Our desire to work with this data is shared by many other groups, including our colleagues at the Sunlight Foundation, Columbia University, Cornell University, University of Colorado, Harvard University, Northwestern University, Stanford University, and GovTrack.US. All of these academic and nonprofit groups have notable track records for providing innovative uses of government data, and the lack of bulk access to these databases has greatly discouraged development of new applications.

Patents and “the law” have a very special place in our system of government, being the only two executive branch databases specifically called out in the U.S. Constitution:

- The very purpose of the patent database is to “Promote the Progress of Science and useful Arts.” The very essence of a patent is publication, and deliberately restricting access goes against the explicit language of the Constitution. While we are sympathetic with the desire of the U.S. Patent Office to derive revenue from the sale of these bulk feeds, such a policy runs directly contrary to their primary mission. Filing fees imposed on those that seek economic gain from the public through the issuance of a patent are more than sufficient to make up any revenue shortfall created by making bulk data available at no cost.
- Likewise, the purpose of the Federal Register system is to provide a systematic vehicle for notification and publication of regulations that are enacted by the government. Restricting access to this data by putting it behind a series of \$17,000 pay walls yields less than \$200,000 in annual revenue to the government, yet is costly enough that only a few well-heeled corporations have access. The public interest simply can’t afford to play.

Initiatives such as Data.Gov have been very successful and you are both to be applauded for the dramatic change in philosophy in the U.S. Government when it comes to release and dissemination of information. However, it is my worry, a worry shared with my colleagues listed above, that any progress on releasing the USPTO and NARA databases in bulk will become entangled in bureaucratic delay, and I am writing to urge that you make these crucial documents of our democracy available sooner rather than later.

Respectfully yours,

Carl Malamud  
President & CEO  
Public.Resource.Org

cc: Michael L. Wash, Chief Information Officer, Government Printing Office  
John B. Owens, II, Chief Information Officer, U.S. Patent and Trademark Office  
Raymond A. Mosley, Director, Office of the Federal Register