

454 Shotwell Street
San Francisco, CA 94110
August 16, 2004

Rick Otis
Deputy Assistant Administrator
Office of Environmental Information
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue NW, Mail Code 2810A
Washington, DC 20460

Kristin Tensuan
Program Outreach and Communications
Office of Environmental Information
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue NW, Mail Code 2822V
Washington, DC 20406

VIA E-MAIL AND PAPER MAIL

EPA Docket No. OEI-2004-0002, "Public Forum on the eRulemaking Initiative's
www.Regulations.gov Web site and the Development of the Federal Docket
Management System"

Dear Mr. Otis and Ms. Tensuan,

Thank you for the opportunity to attend your office's public meeting on the eRulemaking Initiative on August 2 in San Francisco. We were impressed by the prototypes you presented and appreciate your efforts to encourage public participation in the design of the system itself, as well as in future proceedings that will be docketed through it. As we promised, we are writing to you to present our views about features that should be present in the Federal docket management system and about some related questions that were raised at the meeting.

Many advocacy organizations and other users of Federal regulatory dockets share these proposals; they are partly a result of discussions we had with other advocates at Prof. Stuart Shulman's workshop on rulemaking held at the George Washington University in June.¹ Providing these features will serve the purposes of the eRulemaking Initiative by improving timely access to regulatory information and enhancing public participation. We view these suggestions as closely aligned with the eRulemaking Initiative's vision of "help[ing] citizens easily access and participate in a high quality, efficient, and open rulemaking process" and its goals of "increas[ing] the amount, breadth, and ease of

¹ At Prof. Shulman's workshop, we saw an earlier presentation about the eRulemaking Initiative and discussed questions of functionality with a number of other members of the advocacy community. We look forward to Prof. Shulman's forthcoming report on the workshop and encourage you and your colleagues to review it when it becomes available.

citizens and intergovernmental access and participation” by “enhanc[ing] public access to information on federal rulemaking.”²

I. Bulk Data Retrieval

The Federal docket management system should support the retrieval and download of an entire docket or record set. Such a bulk retrieval feature can be seen merely as a special case of programmatic docket output, which is discussed below. However, this feature is particularly important to those who comment regularly, as they will frequently need to read and re-read the entire record in a proceeding.³ It is also of interest to journalists, researchers, and those preparing challenges to regulations.

Bulk retrieval of an entire docket is the most-requested feature among representatives of advocacy organizations with whom we have conferred. Existing government systems typically lack this feature; they usually permit an entire docket history to be *displayed*, but do not allow the associated docket items to be *downloaded simultaneously*.

The utility of bulk retrieval is such that advocates have already created screen scrapers for some existing docket systems on a case-by-case basis.⁴ (A screen scraper is a program that uses a display intended for human beings to extract information automatically into a standardized format. Ordinarily, a screen scraper must be written for each display or interface from which information is to be obtained, and must be updated if the interface changes in any significant way.) This activity is wasteful of both advocates' resources and government resources, but the lack of bulk retrieval functionality in most government information systems has made it appear the most appealing alternative to the advocacy community today.

An officially supported bulk retrieval feature would go a long way toward eliminating this waste and improving public access to regulatory information. Presumably, such a feature would allow interested parties to retrieve an entire docket history in the form of a single archive file, but other approaches exist.⁵ This feature would be more efficient on

2 eRulemaking Initiative, “eRulemaking Initiative Public Forum” presentation, August 2004.

3 Agency practice differs with regard to what interested parties may do after the formal public comment period for a proposed regulation has closed. The Federal Communications Commission, for example, regularly provides a formal reply period during which parties may submit formal responses to one another's initial comments. Many agencies permit the submission of *ex parte* letters or presentations after formal comments have closed. Reviewing the entire record to date is standard practice for advocates preparing evidence and arguments to rebut or amplify others' claims. It can also be essential for understanding the historical context of a particular rulemaking.

4 For example, one of us created a primitive screen scraper for the FCC's Electronic Comment Filing System (ECFS) in 2003 to allow the automated download of about 5,000 docket items from a proceeding of interest to us. This effort was wasteful in the sense that it is unlikely to work with any future versions of ECFS and in the sense that the FCC could easily have provided equivalent access to this data by means of a public, supported interface. We are aware of other anecdotal evidence that advocates are regularly writing screen scrapers for other government systems.

5 The system might, for instance, allow all items in a docket to be downloaded in a single ZIP archive file, which could then be extracted on the requester's own computer. It might also be possible to support bulk retrieval by other means. If a docket contained 100 PDF files of 5 pages each, a bulk retrieval feature could allow the download of a single 500-page PDF file containing the entirety of the docket's text. Software could “stamp” each page to indicate from which underlying docket item it had been

the government side because it would avoid the need to perform repeated searches and generate redundant user interfaces for a person or screen scraper intent on downloading *everything*. It would be more efficient on the advocacy side because it would eliminate the need to develop screen scrapers or ask human beings to perform the tedious work of downloading an entire proceeding record by hand.

Allowing bulk retrieval is useful to any advocate, including those in the legal profession, as well as to scholars and journalists who desire a comprehensive view of exactly what was submitted in a particular docket. It also facilitates creating alternative presentations and collections of regulatory information, which may be useful to constituencies from libraries to advocates to publishers of government information.⁶

II. Public Programmatic and Scriptable Docket Input and Output

The Federal docket management system should allow and encourage the public to use and develop third-party software to get information into and out of dockets. To this end, it should publish and support interfaces for public input and output, such as *web services* interfaces.

By providing standardized interfaces to the Federal docket management system, the eRulemaking Initiative will allow any member of the public to create new software that adds functionality to the system.⁷ That means that people can implement desired functionality when the need arises – including, for instance, creating disability

taken. Finally, web services interfaces might help here, too, by allowing the docket management system to return an appropriately formatted complete list of URLs corresponding to the contents of a docket; that list could then be parsed readily by machine and all the referenced dockets could be downloaded mechanically. A hypothetical XML file might contain references to all docket events, a limited amount of metadata, and associated URLs for retrieval of the documents corresponding to the events. It would then be feasible for any programmer to create a useful docket browsing or downloading application.

⁶ Some participants in your office's recent public meeting seemed to hold the view that it is most desirable for all users of a docket to have recourse to the official public government version of a docket, rather than obtaining access through some other source or republisher. We submit that it has always been the case that interested parties have obtained government records through third parties. In any case, officially supported bulk access is a matter of efficiency and timeliness, not a qualitative change in the legal standards applicable to public access. The data in the Federal docket management system, subject to certain exceptions, will generally be a matter of public record, and the public will be entitled by law to obtain it and to re-use it for a wide variety of purposes. In other contexts, the government routinely permits third parties to package and even to sell for profit data obtained from government systems (such as the SEC's EDGAR system, the datasets about American demographics collected by the Bureau of the Census, and data about the natural world collected by the United States Geological Survey and the National Weather Service). We believe this republication is entirely proper and amply demonstrates that no catastrophe results from the availability of multiple sources of public records. In the context of rulemaking proceedings, we do not believe that attempts to contract the scope of public access to public records would be warranted. The view that reuse and republication of such records is appropriate is expressed, among other things, in the policy of the copyright law, which precludes Federal government entities from obtaining copyright in their original works. See 17 USC §105.

⁷ Compare, for example, the Google web APIs at <<http://www.google.com/apis/>>. By providing an open interface to the functionality of the Google search engine, Google's developers have allowed independent developers to create new applications based on Google, including applications Google had not previously anticipated. Like other web services, Google uses standards like SOAP to enable a programmatic interface.

accessibility interfaces beyond those included in a web browser.

Of particular interest to us as an advocacy organization is the practice of creating *action centers* for on-line public advocacy. Advocates have done this for some time and action centers, whether home-grown or created by contractors, are now widely deployed.⁸ We believe that action centers have an important role to play in facilitating public comment. Action centers can benefit significantly from the availability of programmatic interfaces, such as web services interfaces, for input into the Federal docket management system. Their role in the comment process bears more detailed discussion, and we intend to address this question separately in the future.

In general, programmatic interfaces leave room for third parties to develop useful tools that are beyond the original scope of the eRulemaking Initiative, or that would have been difficult to anticipate beforehand. This sort of extensibility is appropriate to a modern database-backed web site, and offering it will enhance the value to the public of the regulatory data in the Federal document management system.

III. Flexible Automated Notification

We were happy to see that the docket management system will include a notification system to help the public keep track of certain regulatory events. We believe that a notification system is an essential feature, and we encourage the eRulemaking Initiative to try to make its notification features as flexible as possible. Notifications are particularly important because they can help reduce the previously immense cost of tracking the Federal Register and individual dockets for occasional references to topics of interest. In the past, this cost has produced significant asymmetries between "large" and "small" organizations tracking an issue; those who could afford to hire law firms to monitor regulatory activity became aware of important events and made their voices heard, whereas those who lacked these resources did not. The on-line document management system can help narrow this gap, and automated notifications are one important means of doing so.

As an initial matter, notifications should be available through RSS, not only through e-mail. RSS is the emerging standard for publicizing all kinds of regularly-updated events and documents, such as news feeds, newsletters, weekly columns, and even syndicated comics.⁹ It is now the state of the art for notification applications on the Internet. Dozens of client software applications support RSS and generally allow a user to synthesize multiple news sources in a single user interface.¹⁰ Millions of Web sites support RSS as a means of announcing changes or the availability of new documents or news items. Some state governments have begun using RSS as a way to publish

⁸ One example is EFF's Action Center at <http://action.eff.org/>. Most of our action center campaigns are directed at legislatures, but some have been directed at agencies. We and other action center operators (and our contractors who develop them on a commercial basis) can benefit directly from interfaces that allow integration between action centers and regulatory docket systems.

⁹ On RSS and deploying it, see, for example, Ben Hammersley, *Content Syndication with RSS: Sharing Headlines and Information Using XML* (Sebastopol, CA: O'Reilly and Associates, 2003).

¹⁰ Supporting RSS yields instant compatibility with all these applications and others that will be written in the future, and makes notifications more useful to an enormous prospective audience.

government information.¹¹

The Federal government should follow suit. **Any notification feature in the Federal docket management system that can be accessed through e-mail should also be made available through an RSS feed.**

Docket users may wish to receive – and the system should support – a wide variety of notification types. For example, it would be useful to us to be able to

- receive a notification whenever a particular agency issues an NPRM;
- receive a notification whenever a particular agency issues a final rule;
- receive a notification whenever a particular agency issues any docket whatsoever;
- receive a notification whenever any agency issues an NPRM mentioning a particular keyword;
- receive a notification whenever any agency issues a final rule mentioning a particular keyword;
- receive a notification whenever a particular party comments in any docket;¹²
- receive a notification whenever any item is filed in a particular docket;
- receive a notification whenever any item is filed in a particular docket and containing a particular keyword;¹³ or
- receive a notification whenever any item is filed in any docket mentioning a particular keyword.¹⁴

The sheer variety of these applications suggests that, ideally, **the Federal docket management system should allow notifications to be generated that correspond to any search request the system can accept.** In the prototype we saw on August 2, search terms were extremely flexible, but notification support was somewhat less flexible. Notifications could be much more useful if they took on the full flexibility of searches.¹⁵

11 For example, the Commonwealth of Virginia describes its use of RSS at

http://www.virginia.gov/cmsportal/services_869/information_2656/index.html; we are advised that Virginia also provides or will provide RSS feeds for individual state agency regulatory dockets.

12 All advocates may have repeated conflicts with other advocates over particular issues about which they disagree. In some cases, it could be helpful to us be able to learn when certain entities comment with any agency, because we are extremely likely to want to reply. We might, however, have been previously unaware that a particular proceeding implicated any of our interests. (For example, an issue of copyright policy might be raised in a docket that had previously concerned only the technical details of a broadcast format.)

13 Many proceedings have a large number of topics and commenters may have an interest in only one or two. Filtering docket events may thus be important.

14 We may be interested when someone cites a particular statute or regulation, regardless of the agency to whom that statute or regulation has been cited (especially if we advocate reforms of that statute or regulation or are challenging or applying it in the courts). We may also be interested to know whenever someone mentions a particular topic in any advocacy before the Federal government. It is easy to imagine, for example, that an environmental advocate might want to know whenever perchloroethylene is mentioned in any proceeding (not merely proceedings devoted to that subject). For that matter, a manufacturer of perchloroethylene might well want to obtain the same information.

15 It is possible that some notification types would be more difficult to implement than others because of the burdens they place on system resources. We thus recognize that, depending on system architecture and available resources, it might not be possible to attain the ideal of making every possible search available as a notification. We encourage the eRulemaking Initiative and its contractors to think creatively about how to support as many notification types as possible.

We also suggest that **notifications should be available at both the docket event level (the filing of a new item within a docket) and the docket creation level**, in the same way that searches will be able to operate either on docket items or on dockets themselves.

IV. Miscellaneous

It should be possible to obtain a permanent URL for linking directly to a particular docket or to a particular document. This is already possible in some existing agency docket management systems, and is extremely useful for purposes of reference and public discussion. Then, instead of describing what may be a complicated search procedure, a writer can simply refer people directly to the original source document using a standard web link. For example, a journalist might refer to a filing (or a rule) in an on-line article, and could allow readers to follow a link to the actual materials under discussion.

The eRulemaking Initiative should establish a clear and public privacy policy for its users that exceeds statutory requirements for the collection of information from the public. We understand that some people are afraid to comment on, and even to research, controversial topics because they don't know what the government will do with the information they provide. Other people legitimately fear retaliation from their employers, prospective employers, or even social acquaintances for taking a position on a controversial issue or disclosing sensitive information to regulators.¹⁶ The Federal docket management system should try to mitigate these concerns by having a strong and clearly stated privacy policy and by avoiding retaining any more information about its users than is necessary to accomplish its purposes.¹⁷

Thank you for the opportunity to comment on the functionality of the Federal docket management system. Please let us know if can be of any further assistance.

Sincerely,



Ren Bucholz
Activism Co-ordinator



Seth Schoen
Staff Technologist
Electronic Frontier Foundation

¹⁶ For instance, even mundane research by an employee into workplace safety regulations or the law relevant to unionization might provoke an employer's anxiety.

¹⁷ Regulations.gov has an existing privacy policy at <http://www.regulations.gov/privacy_and_use.cfm>, which we presume will be reviewed or augmented in the course of Phase 2 of the Initiative.