



May 7, 2008

James C. Duff, Director  
Administrative Office of the U.S. Courts  
One Columbus Circle NE  
Washington, D.C. 20544

RE: Public Access to Opinions of the U.S. District and Bankruptcy Courts  
E-Government Act of 2002 and Reauthorization

Dear Director Duff:

I am writing to you in your capacities as Director of the Administrative Office of the U.S. Courts and as Secretary to the Judicial Conference of the U.S. concerning the E-Government Act of 2002 (the "2002 Act") as it pertains to the meaningful access to the judicial opinions of the nearly 200 U.S. district and bankruptcy courts.<sup>1</sup> In this letter, I provide an overview of the Act, the implications of the proposed reauthorization of the 2002 Act, and the accessibility of these opinions. I urge that the federal judiciary take the final small steps needed to permit the public to have unrestricted access through search engines and public access law web site to all lower court federal judicial opinions. I ask that you forward a copy of this letter to the appropriate committees of the Judicial Conference and that you arrange for a meeting to discuss the issues mentioned herein.

I will discuss the barriers that still exist as to the accessibility of opinions of the lower federal court and include with this letter specific steps that could be taken to make these opinions more widely available to the public through free sites operated by public interest groups, private companies, and law schools, and searchable and retrievable by the public through search engines such as Yahoo and Google.

## SUMMARY

- The E-Government Act of 2002 required the federal judiciary to make all judicial opinions, published and unpublished, available in searchable format.

---

<sup>1</sup> As founder of HyperLaw, Inc., I have argued for broader public access to judicial opinions since 1991, when HyperLaw released the first CD of United States Supreme Court opinions and then in 1993 the first CD of opinions of the United States courts of appeals. Subsequently, we were engaged in litigation with West Publishing Co. that established that West could claim copyright neither to its internal page citations nor to the text of court opinions as corrected and enhanced by West. In the 1990s, we were involved in efforts to create a public domain citation and wider dissemination of federal court opinions and met and or provided formal presentations to the Judicial Conference, Congressional committees, the Administrative Office, the American Bar Association, the Department of Justice, the Executive Office of the White House, and the American Association of Law Libraries.

- Nonetheless, because of the way in which the federal judiciary implements access to judicial opinions, public access law web sites and search engines have been frustrated in accessing and/or collecting the opinions of the nearly 200 U.S. district and bankruptcy courts and making them freely available and searchable in a meaningful manner by members of the public.
- Although no doubt there are many opinions available on the web, it is simply a myth that comprehensive collections of all judicial opinions of the nearly 200 U.S. district and bankruptcy courts are available on the Internet, either through public access law sites or through commercial search engines. Tellingly, even WestLaw and Lexis, which provide free access to Supreme Court and court of appeals opinions of recent years, do not provide open and free access to the district and bankruptcy court opinions.
- Pending E-Government legislation would require that opinions be indexable and searchable by search engines such as Yahoo and Google.
- For over seventeen years, the Administrative Office and the Judicial Conference have been considering creation of a repository of the opinions of the nearly 200 U.S. district and bankruptcy courts, but deferred action because of system limitations.
- In the past, opinions were not in electronic format, uniform naming of opinion files was limited by technology and systems, and the judiciary did not have a document management system to accommodate the requirements for a repository of opinions. These limitation no longer exist. As well, at least at the court of appeals level, there is no longer a distinction between published and unpublished opinions.
- The federal judiciary's Case Management/Electronic Case Files (CM/ECF) and PACER systems now collect and systematically maintain all documents filed in a case, which, necessarily includes all judicial opinions. Thus, the federal judiciary does indeed have a repository of opinions, although obscure and not directly accessible as such.
- Each judicial opinion (and every other case filing) is identifiable within CM/ECF by the name of the court, the docket number of the case, the docket entry number, and the date of the opinion.
- The CM/ECF system easily could, but does not, assign a unique persistent public file name to each judicial opinion. Search engines would require persistent file names. Bulk downloading by public access sites requires uniform persistent file names. Issues of authenticity when citing these opinions could be avoided by assigning unique file names.

- Users of CM/ECF with passwords have free access to opinions designated by the courts as written opinions, but these opinions are behind fire walls, bulk downloading is not allowed, and not all opinions are in a searchable format. Further, not all opinions are properly designated as written opinions to be made available to the public via this facility.
- Further efforts are required by the judiciary to assure that courts designate in the CM/ECF system, not just some but all judicial opinions, whether or not published, unpublished, precedential or not precedential.
- The federal judiciary budget for technology exceeds \$400 million a year.
- Because all case documents and related data are now routinely collected and stored in systematic manner by CM/ECF, only relatively minor technical and administrative issues need to be resolved in order to make all of the opinions of the nearly 200 U.S. district and bankruptcy court available to search engines and available for bulk download by public access law sites and the public.
- The Administrative Office of the United States Courts should implement the minor technical enhancements so that such accessibility is provided.
- The Judicial Conference of the United States should commit to assuring that all U.S. court judicial opinions are appropriately designated and that accessibility of the opinions by search engines and other public access law sites be assured, as quickly as possible.

## **BACKGROUND AND CURRENT STATUS**

Since 1991, the issues of providing public access to the lower federal court opinions has been complicated by the mechanics of dissemination of the opinions and, as well, a lack of uniform methodology for the naming of opinion computer files, often earlier referred to as "electronic citations." Today, the structural framework to allow meaningful access to these opinions is in place, thanks to the CM/ECF (aka PACER) program.<sup>2</sup>

---

<sup>2</sup> In this letter, I will use CM/ECF to describe the systems, including PACER, used by the judiciary to automate case management, case filing, and public access. Many use the term PACER to apply to all federal judiciary information systems. PACER (Public Access to Court Electronic Records) was the initial system developed by the Administrative Office, which describes PACER as "an electronic public access service that allows users to obtain case information from federal courts." Case Management/Electronic Case Files (CM/ECF) is "an electronic case management system that provides federal courts with enhanced and updated docket management capabilities, including the option of permitting case documents to be filed with the court over the Internet." Almost all district and bankruptcy courts and most courts of appeals have converted to the CM/ECF software system, as the earlier PACER system is being phased out.

The CM/ECF system now collects and hosts all such opinions. Thanks to CM/ECF, authoritative versions of all federal lower court opinions can be uniquely identified and located using the court name, the docket number of the case, and the docket entry number of the opinion document. This letter will in no way suggest any action as to adoption of a citation format, although it will recommend the inclusion of this relevant identifying information in the opinion document file metadata. Traditional views of citations are being reconsidered.<sup>3</sup> What is important is to have a way to uniquely describe a court opinion, and then to find the authoritative version. CM/ECF is now able to provide that functionality with minor modification.

It is worth noting that seventeen years ago, in July 1991, the Library Program Subcommittee of the United States Judicial Conference Committee on Automation and Technology issued a Draft Report on the establishment of an electronic citation system applicable to federal judicial opinions.<sup>4</sup> It was recognized then that although the courts of appeal published opinions in the form of slip opinions, the lower federal courts did not publish slip opinions. Thus, within the judiciary, there was no official selection process as to opinions. Lexis and Westlaw already by that time were collecting as many lower court opinions (published and unpublished) as they could - basically creating a privately controlled repository of federal law. It was recognized then, as well, that as opinions became available digitally, that some type of citation methodology would be needed to facilitate access, because computer files with meaningless names would only complicate matters. This explains the earlier use of the word "electronic" citation. Indeed much effort was then expended into compressing the citation into the 8 characters available in a DOS file name. With extended file name and the ability to include metadata<sup>5</sup> in files, limitations as to including document identification information within a digital file no longer exist.

Now, seventeen years later, structural issues having been resolved, it is time to take the minimal steps required to remove the few remaining barriers to effective access and dissemination of these opinions. The opinions need to be made accessible to Internet search engines, metadata information needs to be included in the opinions for effective searching, bulk downloading needs to be facilitated, and the quality in terms of completeness needs to be monitored and improved.

---

<sup>3</sup> Strictly speaking, citation information is any information that uniquely defines a document sought. This is to be distinguished by a citation format which may mean how one abbreviates and orders the citation information, or even a preference as to the location of the document.

<sup>4</sup> "Standard Citation to Electronic Opinions", Revised Draft Report, dated October 17, 1991, prepared by the Library Program Subcommittee of the United States Judicial Conference Committee on Automation and Technology. See also the earlier July 18, 1991, Draft Report. HyperLaw provided comments to the Administrative Office on April 9, 1992.

<sup>5</sup> In an Adobe PDF file (which is the file format used for all files in CM/ECF) the most significant item of metadata is the title field in document properties. Search engines such as Yahoo and Google will search first on this title field and, if one exists, the title field will be displayed first in the search results.

Although the district and bankruptcy courts do make opinions available on either separate opinion sites or through CM/ECF, many have significant limitations. Sites are not indexable, opinion files are hidden behind firewalls and have no file names or identifying information, and some sites make it difficult to download opinions in bulk. In some situations, documents are not marked as opinions on CM/ECF, but appear on a court's opinion page and vice versa. Opinions appear on Lexis and WestLaw, but are not marked as opinions on court sites.

Because the lack of uniformity and other access issues ultimately create barriers to collection of opinions, demonstrably limited public and free access to lower federal court case law exists today as described below. The problems as to the lower federal court opinions have persisted for nearly 15 years after the U.S. Supreme Court and courts of appeals have made their opinions available on dial up bulletin boards or web sites.

The Report Of The Proceedings Of The Judicial Conference Of The United States of September 23, 1997 included this statement from the Committee On Automation And Technology:<sup>6</sup>

*The Committee will explore studying the desirability, feasibility, and cost of establishing a centrally maintained, publicly accessible electronic database of all opinions submitted by federal courts for inclusion in the database.*

Since that report, another 11 years have elapsed. Now that all lower court opinions are saved and filed in the Adobe PDF format in CM/ECF, there would appear to be no known barriers to establishing "*a centrally maintained, publicly accessible electronic database of all opinions submitted by federal courts.*"

With the advent of Google and Yahoo type search engines, all that the courts need to do is to provide a central repository of opinion computer files, provided that each file has an appropriate file name and includes necessary identifying metadata as described below. The repository would also need to allow bulk downloading. The federal judiciary need not invest in creating or installing its own search software - rather, first, it should merely act as the

---

<sup>6</sup> The Report of the Subcommittee on Policy and Programs Concerning Standard Electronic Citations (1997) of the Judicial Conference Committee on Automation and Technology stated:

The primary assumption underlying the proposal is that judicial opinions are public documents and that it is in the best interest of the judiciary and the public for such opinions to be made available to judicial officers, litigants, and the public as quickly and inexpensively as possible. While official case reports (United States Reports, for example) have historically seldom been available to the public on a timely basis, the subcommittee sees no reason why this should be so if the opinions were posted electronically. Creation and maintenance of a central database of federal opinions would appear to be a matter undoubtedly within the authority of the judicial conference.

repository of the opinions by storing accessible files on a file server. This would be cost effective.

The Long Range Plan for Information Technology in The Federal Judiciary <sup>7</sup> provides for an IT budget of \$411 million for FY 2008: \$30.7 million alone is allocated to the Electronic Public Access Program portion of the overall CM/ECF budget. An Objective of the Long Range Plan is:

Provide the public and the bar with easy access  
to appropriate court and case information.

As to the district and bankruptcy court opinions, access is not easy.

The programming required to implement the changes indicated herein are, in relation to that size budget, insubstantial if not insignificant.

CM/ECF software needs to provide the accountability and reporting tools to ascertain how judicial opinions are being identified and made available to the public, to provide bulk download capabilities, to appropriately identify opinion files with proper file names and metadata, and to make the files available for indexing by search engines.

## **The E-Government Act of 2002**

Sec. 205(a) of the E-Government Act of 2002 (Public Law 107-347)<sup>8</sup> mandated all federal courts to maintain websites with "access to the substance of all written opinions issued by

---

<sup>7</sup> Long Range Plan for Information Technology In The Federal Judiciary, Fiscal Year 2008 Update, <http://www.uscourts.gov/itplan/2008/2008report.pdf>

<sup>8</sup> E-Government Act of 2002 ,Public Law 107-347, [http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107\\_cong\\_public\\_laws&docid=f:publ347.107.pdf](http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:publ347.107.pdf)

### SEC. 205. FEDERAL COURTS.

(a) INDIVIDUAL COURT WEBSITES.—The Chief Justice of the United States, the chief judge of each circuit and district and of the Court of Federal Claims, and the chief bankruptcy judge of each district shall cause to be established and maintained, for the court of which the judge is chief justice or judge, a website that contains the following ...

\* \* \*

(5) Access to the substance of all written opinions issued by the court, regardless of whether such opinions are to be published in the official court reporter, in a text searchable format.

the court, regardless of whether such opinions are to be published in the official court reporter, in a text searchable format."<sup>9</sup>

The E-Government Act of 2002 allowed individual courts to operate web sites for court opinions or to link to other sites with the opinions. Although it appears that the 2002 Act provided discretion to the clerks and judges of the separate district and federal courts as to how CM/ECF is implemented, there was no suggestion that there be 200 approaches to hosting these opinions. Some district and bankruptcy courts rely upon the facilities provide by the CM/ECF system. Because written opinions, whether published or unpublished, are supposed to be designated as such in the CM/ECF system, the better approach to supporting separate web sites, if that is believed necessary, is to synchronize sites with the same data in the CM/ECF system.

Minor enhancements to CM/ECF would permit all of the district and bankruptcy courts to provide uniform and complete access to their judicial opinions, which would have the result of simplifying public access.

The development of the software for CM/ECF is provided by the Administrative Office, and accordingly, the Administrative Office bears responsibility for software and systems to make uniform access of lower court opinions a reality. The Administrative Office should develop and provide a standard interface for individual court opinion web sites so that the opinion files are synchronized with the files designated as written opinions in CM/ECF using a uniform file naming convention. The Administrative Office should then maintain a central repository of all such files. No changes at all in the underlying CM/ECF system would be required - this is merely a task of creating a user interface.

By and large, the Administrative Office and individual courts have made good faith efforts to implement the 2002 Act requirement, but that does not mean the Act has been fully implemented by the federal judiciary. The federal judiciary has not been in compliance, nor, would it be in compliance were the reauthorization discussed below enacted.

## **The E-Government Act Reauthorization Act of 2007**

Senators Lieberman and Collins have introduced the E-Government Reauthorization Act of 2007 as S. 2321, reauthorizing and amending the E-Government Act of 2002. On December 11, 2007, the Senate Committee on Homeland Security and Governmental Affairs held a hearing titled "E-Government 2.0: Improving Innovation, Collaboration and Access" and representatives from the Office of Management and Budget, Google,

---

<sup>9</sup> Among the organizations that worked with Congress to assure that the judiciary was included in the Act were the American Association of Law Libraries, the American Library Association, and the Special Libraries Association.

Wikipedia, and the Center for Democracy and Technology provided testimony. The testimony focused on agencies and not the judicial branch.

The witnesses pointed to aspects of government databases and information that diminish the ability of citizens to access government information, including dynamic databases and inaccessible links. A Google witness stated that it was not possible for Google and other search engines to access such data for the public.

The deficiencies identified by Google and the other witnesses do exist as to the opinions of the lower federal courts.

The bill requires agencies to make their information accessible for searching on the Internet and provides a two year goal for compliance.<sup>10</sup>

The legislation has yet to be enacted - but, it is believed that these requirements are an expression of the public's expectations and that provisions of this type of requirements ultimately will be included in legislation.

Finally, even as to the E-Government Act of 2002, the requirement for the judiciary was to make the opinions "in a text searchable format." What would be the purpose of having the opinions "text searchable" if the opinions were then protected by a firewall from search engines and if bulk downloading was not possible?

## **Judicial Access to Opinions Prior to E-Government Act of 2002**

Even prior to enactment of the E-Government Act of 2002, the federal judiciary had recognized its responsibilities to the public to make judicial opinions available in electronic form to the government, starting with the Supreme Court's pioneering Hermes program in the early 1990s. Subsequently, in the early 1990s some U. S. courts of appeals made their

---

<sup>10</sup> The bill adds the following to the existing E-Government Act of 2002.

- A finding that " members of the public and governments commonly rely on commercial search engines to locate relevant information on the worldwide web, including information made available by government agencies."
- A finding that "some Federal agencies have not taken actions to make all of the information available through their websites readily accessible to commercial search engines."
- Guidelines : "Not later than 1 year after the date of enactment of the E -Government Reauthorization Act of 2007, the Director shall promulgate guidance and best practices to ensure that publicly available online Federal Government information and services are made more accessible to external search capabilities, including commercial and governmental search capabilities. The guidance and best practices shall include guidelines for each agency to test the accessibility of the websites of that agency to external search capabilities."



opinions available by dial-up bulletin boards. The courts of appeals, some in collaboration with law schools,<sup>11</sup> then made their opinions available on the Internet. By the time the 2002 Act was enacted, the courts of appeals were already making their judicial opinions available to the public. Not the same could be said for most district and bankruptcy courts - opinion availability was sporadic and inconsistent.

After the 2002 Act was passed, the lower courts were able to use some of the capabilities of the CM/ECF system to provide access to their written opinions or utilized separate web sites. As discussed below, the opinions are then presented in such an inconsistent manner as to make collection and access difficult.

### **Lower Federal Court Opinions Not Available on Public Access Sites**

As a result of the inconsistent, incomplete, and indirect access, non-judicial web sites devoted to free and public access to court opinions have been frustrated, and either are unable to provide opinions from these lower federal courts, or provide them inconsistently and incompletely.

Thus, sites that provides free public access such as Cornell Law School's Legal Information Institute, Columbia Law School's AltLaw, Justia, Public Resource and Precedent offer access to U.S. courts of appeals and U.S. Supreme Court opinions, but not in any comprehensive manner, if at all, to opinions of those lower federal courts. Moreover, indexing by Yahoo and Google is not possible for all opinions.

The inability of free public access sites to provide the complete district court and bankruptcy opinions, explains the following: both LexisOne and West's FindLaw provide free limited searching of the U.S. court of appeals and Supreme Court decisions for recent years, but *LexisOne and West's FindLaw do not provide free access to opinions of the U.S. district and bankruptcy courts.*

I would submit that this lack of access is directly attributable to the inconsistent attention provided to this issue by the Administrative Office and the federal judiciary.

---

<sup>11</sup> The law school programs hosting federal court opinions on law school web sites were intended to be demonstration programs only, as stated by Professor Robert Oakley of the American Association of Law Libraries in his 1998 presentation to the Senate Committee on Rules and Administration concerning S. 2288, copy at <http://www.aall.org/aallwash/tm0729a2.html> :

The purpose behind this voluntary project was to demonstrate to the courts that the use of electronic communications networks can facilitate the timely and low cost dissemination of court opinions. The project was not intended to relieve the courts of their own dissemination responsibilities, but rather to encourage them to follow the model of electronic public dissemination.

HyperLaw, Inc, 17 W. 70 St., New York, NY 10023, 212-873-1371 [www.hyperlaw.com](http://www.hyperlaw.com)

Alan Sugarman, [sugarman@sugarlaw.com](mailto:sugarman@sugarlaw.com)

## **Should the Federal Judiciary Provide Better Access to Lower Court Opinions?**

In 1997, the Subcommittee on Policy and Programs Concerning Standard Electronic Citations of the Judicial Conference's Committee on Automation and Technology in its report stated:

While official case reports (United States Reports, for example) have historically seldom been available to the public on a timely basis, the subcommittee sees no reason why this should be so if the opinions were posted electronically. Creation and maintenance of a central database of federal opinions would appear to be a matter undoubtedly within the authority of the judicial conference.<sup>12</sup>

Eleven years later, though, availability of lower federal court decisions is still not adequate - effectively, access to all published and unpublished opinions of all district and bankruptcy courts is available only through WestLaw and Lexis. This situation was not deemed appropriate in 1997 - that it remains so today in the face of nearly a half-billion dollar a year court technology budget cannot be rationalized, especially when the technical and financial requirements to remedy the situation are trivial.

### **Barriers To Access Easily Resolved Within CM/ECF**

With very minor enhancements to the CM/ECF program, the federal judiciary would be able very quickly to provide a substantial payback to the public which has invested heavily in the federal judiciary's technology.

Decisions of some courts are hidden by firewalls making them not available to search engines such as Google and Yahoo. Metadata in the CM/ECF database files providing information about the opinions are not necessarily included within the opinion files, making it difficult for search engines to locate particular opinions and for others to have associated information about the opinions. There also are issues of completeness.

This situation presents barriers to anyone who wishes to search among all district Court and bankruptcy decisions and presents barriers to any web site wishing to collect and compile all of the decisions in any methodical manner. Westlaw and Lexis, due to their size and financial resources, obviously are able to pay the costs to overcome these barriers, and do make these decisions available to their subscribers. Not even all lawyers are able to bear the costs of subscriptions to these services and the public has little access. Now that unpublished opinions are citable, the barrier of access to opinions not in print has become more significant.

---

<sup>12</sup> See footnote 6, *infra*.

The CM/ECF system includes PDF images of all documents filed on the docket and has been implemented in all United States district and bankruptcy courts.<sup>13</sup> Accordingly, all decisions of all judges in Adobe PDF format are available on CM/ECF and managed by a database system which stores the docketing information about each opinion. Thus, the data now exists electronically and exists in a methodical manner. This circumstance did not exist in 1997.

CM/ECF provides an standard report, the "Written Opinions" report, to make court opinions available to the public with a PACER ID and Password, for free. Although there is a charge to download other CM/ECF documents, there is no charge for CM/ECF opinions when accessed from the Written Opinions report.

On the surface, this sounds like a very good solution: all opinions are available for free from the CM/ECF system. But, if court documents are not completely and appropriately designated as "opinions," then access to all opinions cannot occur. If a case document is marked as a Written Opinion, then it will be listed in this report. That having been said, it is necessary for someone to consistently designate which documents are deemed "written opinions." Some district and bankruptcy courts appear to pay close attention to the designation of opinions for the Written Opinions report; others do not.

Because these opinions are behind a firewall and have no persistent name or location, they are not available to search engine indexing or bulk downloading.

The electronic opinion PDF files lack necessary metadata to absolutely and uniquely identify the opinion. All opinions on CM/ECF have a unique identifier within a case - the Docket Entry (DE)<sup>14</sup> number on the docket sheet. Together with the court, date, and docket number, the DE provides a ready citation for the opinion. This information should be included in PDF metadata.

The foregoing will now be explained in more detail.

## **Completeness and Suitability of Opinions Selected**

### ***Not All Judicial Opinions are Identified and Included in Written Opinions***

Not all decisions have been designated as a "written opinions" and thus do not appear in the CM/ECF Written Opinions report. It is not clear who is responsible

---

<sup>13</sup> An exception are in pro se cases, where all case documents frequently are not filed electronically, but, even there, the court's opinions and orders are filed in image format.

<sup>14</sup> Many district and bankruptcy courts routinely use the term "DE" number to refer to the Docket Entry of documents in a case.

for designating documents as written opinions (judge, clerk, assistant), but suffice it to say that in some courts, few if any documents are marked as written opinions (I know of one court with no opinions marked -- but have not advised the court, since I do not wish to chance upsetting a judge hearing a case in which I am involved.)

There are opinions, published and unpublished, available on Westlaw and Lexis but not marked as written opinions. Some district and bankruptcy courts have ignored the clear requirement of the E-Government Act and elected to include only so-called published opinions on their web sites and do not even mark so-called "unpublished" opinions in the CM/ECF system. The distinction of published versus unpublished opinions has never been honored by Westlaw and Lexis, as they collect everything.

### ***Routine Orders Inappropriately Identified as Written Opinions and Included in the Written Opinions Reports***

As opposed to not marking documents as opinions, some courts/judges/clerks go the other way and mark every single order and stipulation as a Written Opinion - including adjournment orders etc. This then makes it difficult to identify the actual real opinions buried in hundreds of one and two page orders. We assume that some courts do this to make all orders available for free to pro se litigants, or perhaps the courts do not wish to attempt to decide the status of a court order/memorandum/judgment. Judges need an alternative for these opinions, such as a Written Orders report.

### ***In General, No Focused Responsibility for Completeness of the Written Opinions Report***

There is no reporting methodology or administrator or clerk identified to whom to report documents missing from CM/ECF or other anomalies. I have found missing opinions, advised judges and clerks, but nothing is fixed. There should be a discrepancy reporting function in CM/ECF. The discrepancies should be monitored by the AO officials responsible for assuring and reporting on compliance with the E-Government Act - it is assumed that discrepancies will diminish once there is monitoring of any type.

### ***Not All Opinions are Searchable***

A small but important number of opinions are not searchable in that they are scanned image files without a text layer created by optical character recognition (OCR). Thus, the files are not searchable. Perhaps, the reason for the practice is that some courts wish to post files with the original clerk's stamp and the judge's signature. All image Adobe PDF files need to be processed by an OCR program. The courts that post image only Adobe PDF files without text are not complying with the E-Government requirement of "searchable."

## Accessibility by Search Engines

### *The Opinions Designated in the Written Opinions and Some Court Web Site Opinions Are Hidden Behind Firewalls*

The CM/ECF written opinions are not accessible by search engines such as Google and Yahoo. I.e., the opinions are hidden by a firewall. Thus, they are not indexable and not searchable.

### *No Public File Name Assigned To the Opinion Files*

CM/ECF is most frustrating in this regard: an opinion (or any other document) accessed to be downloaded has no file name associated with the file which will be used if the file is saved. The default file name would be something like:

"show\_case\_doc.pdf"

or

"http---ecf.akd.uscourts.gov-cgi-bin-show\_temp.pl?file=pdf39570003651490&type=application-pdf."

If saved with these names, there would be no uniformity as to the name of the accessed files. In order to make the files searchable by third party search engines, such as Google or Yahoo, there needs to be a consistent file name and a convention to include the basic metadata discussed elsewhere. The absence of a uniform file name for each opinion is extremely important not only for search engines, but to permit synchronization with other internal and external web sites offering access to the opinions. Bulk downloading would require unique file names. It bears repeating that this uniform file name must include the abbreviated name of the court to meet the purposes described.

### *Metadata in Header Not Used Uniformly*

Fortunately, CM/ECF permits courts as an option to include a header in all filings (including opinions) such as the following:

Case 1:01-cv-00400-T-DLM Document 94 Filed 09/14/2007 Page 1 of 1

Not all courts use the option to include this header on each page. Oddly, when opinions appear on a separate court web site frequently the header with the docket entry number is not included. Thus key metadata is lost.<sup>15</sup>

---

<sup>15</sup> Strictly speaking, one could in an Adobe Acrobat file create a special field or tag for each of the data items. Another approach is to include an XML type file within the Adobe file: Adobe describes this a XMP (Extensible Metadata Platform) The XMP information could be easily exported from the CM/ECF database and inserted into the Adobe file, and is arguably the more "proper" solution and could include all information on the docket sheet for the document. Others might argue for a separate XML file - but, it is simpler conceptually to encapsulate the metadata information.

### ***Header Metadata Does Not Include the Name of the Court***

An interesting observation is that frequently, federal, state, and local courts and agencies leave out the court name/jurisdiction in file names and html and Adobe PDF metadata, on the assumption that anyone searching the opinions knows that information already.<sup>16</sup> This suggests a distinct court-centric quality of many court operated web sites - and, it would appear that some sites are intended primarily to accommodate internal court users (there are 8000 federal judges, clerks, and related support professionals!) One possibly could identify the originating court by the web site address, but, once a document is copied from the web site, this information is no longer available in the file name or title metadata.

The CM/ECF document header does not include the name of the court issuing the opinion. Were the name of the court included, then the header would provide adequate metadata to identify uniquely the opinion for search engines. Thus, it would be recommended to include in the header the abbreviation for the court.

Case 1:01-cv-00400-T-DLM Document 94 Filed 09/14/2007 Page 1 of 1 NYSD Opinion

### ***The Header Does Not Identify the Documents As An Opinion***

Frequently, users of search engines such as Google and Yahoo wish to locate a specific opinion document and thus the search would need to know that the file being searched for is an opinion.

By including the word "Opinion" in the header and/or metadata, the document identifies itself as an opinion which can be used for effective retrieval by search engines if what is sought is the opinion identified by citation information.

### ***Failure to Use the Title Metadata Field in the Opinion Adobe PDF Files***

Most search engines will look to the Title field (under the File-Properties item) in a Adobe PDF file first in indexing a document and displaying a search result. The CM/ECF system has all such information in its database and can quite simply load the case name and the header information into the Adobe PDF title property. Since

---

<sup>16</sup> Resolution of Citation Issue - By including the name of the court, the header contains all information needed to provide a precise citation to the opinion. This is a natural citation requiring no effort at all by judges or clerks, that is indeed used routinely by the lower courts when referring to documents by the DE or DE number which together with court and docket number identify any hundred of thousands federal court opinion, and is locatable quickly in CM/ECF. Any other form of citation, such as a West Federal Supplement, Lexis Cite, WestLaw citation would have a one to one relationship with this unique citation for searching, retrieval, and authentication.

as currently presented, there is neither file name nor information in the title field, the opinions as presented are not adequate either for search engines or public access publishers.

### **Eliminate Requirement for a PACER/ECF ID to Access Free Opinions**

Access to CM/ECF is available only if one obtains a Pacer user name and password. If the court opinions are being made available in this manner so as to comply with 2002 Act, then access should not only be free, but should not require a Pacer ID.

## **CONCLUSION**

In 1991 when I first became involved in this issue, the technology and resources did not exist to make available to the public searchable digital versions of all district Court and bankruptcy opinions. Thanks to the hard work of the judiciary and the Administrative Office, and the successful implementation of CM/ECF, this long standing problem may now be resolved.

As Ari Schwartz of the Center for Democracy & Technology stated at the December 11, 2007 E-Government Reauthorization Act hearings:<sup>17</sup>

[C]ommercial search engines have simply become the most efficient and effective route to find information online. Government agencies must recognize that taxpayers will not find the information that is made available unless this information can be found on commercial search engines. Some agencies have public information resources that are not immediately accessible via search engines due to relatively minor technical problems that the agencies should quickly remedy.

It is my view that the issues of uniform and meaningful access to the opinions of the district and bankruptcy courts are due to "relatively minor technical problems that the agencies should quickly remedy."

I and public access legal publishers and other interested parties would ask to meet with you and your staff to discuss how these issues in CM/ECF could be so remedied at minimal cost to the judiciary.

---

<sup>17</sup> Statement of Ari Schwartz Deputy Director Center for Democracy & Technology before the Committee on Homeland Security and Governmental Affairs on E-Government December 11, 2007.

HyperLaw, Inc, 17 W. 70 St., New York, NY 10023, 212-873-1371 [www.hyperlaw.com](http://www.hyperlaw.com)

Alan Sugarman, [sugarman@sugarlaw.com](mailto:sugarman@sugarlaw.com)

James C. Duff  
May 7, 2008  
Page 16 of 18



Sincerely,

A handwritten signature in black ink that reads "Alan D. Sugarman".

Alan D. Sugarman

cc: Senator Joseph I. Lieberman  
Senator Susan M. Collins



## Specific Technical Suggestions

### (In No Particular Order and Not All Inclusive):

- A. Add a new report similar to the "Written Opinions" report named the "Written Orders" report.<sup>18</sup> The purpose would be to encourage judges and clerks not to load up the Written Opinions report with thousands of documents lacking substantive law content: one and two page orders and pre-trial conferences orders etc. If the judge or clerk wishes to mark these, they can mark these as "Written Orders." In the situation where a "judgment call" has to be made as to whether an order is in the nature of an opinion, then the judge or clerk has an option other than non-inclusion.
- B. "Encourage" the district courts and bankruptcy courts to be sure that all opinions in the separate opinion web sites are synchronized with the opinions selected in Written Opinions in the CM/ECF database.
- C. "Encourage" the district and bankruptcy courts to provide users an e-mail address to report the following discrepancies in the written reports and opinions / or include a reporting form within CM/ECF, or both. This function could be overseen by the E-Government Act official at the Administrative Office.
1. Opinions/Decisions on Westlaw or Lexis, but not included in the "Written Opinion" reports
  2. Any other opinions that should be included in "Written Opinions."
  3. Opinions that are not OCR'd (some courts use image on pages for the first and last pages of opinions and some judges include only image versions.)
- D. Include a "job" ticket approach to the resolution of discrepancies referred to above in C. A ticketing system provides a reporting mechanism used routinely for these types of issues.
- E. Modify the standard header form applied to Acrobat PDF Files to include the abbreviation of the court.
- F. Include in the "title" property of all Adobe PDF files for "Written Opinions" the following identification information: the court, docket number, docket entry number, and the date (i.e., all of the information in the "header" that most courts include PLUS the court name). Include the word "opinion" to assist search engines in distinguishing documents that are opinions. Embed information from the docket sheet for the document in the XML/XMP metadata for the Adobe PDF file.

---

<sup>18</sup> There are no charges for documents accessed from the Written Opinions report, assuming the user has a Pacer of CM/ECF Account.

G. Modify the show case/goDLS CM/ECF Perl procedure so that the default file name by which a document is saved will include this same information: court, docket number, docket entry number, and date. (This enhancement would help anyone saving any Adobe PDF document from CM/ECF/Pacer!)

H. "Encourage" the district and bankruptcy courts to use the same "header" and metadata and file name in their separate opinions web sites.

I. Include additional fields in new database implementations and modifications for the inclusion of parallel citation information from WestLaw, Lexis, and others. How to populate these fields would be a separate issue to be perhaps considered at some future time, but, at least major modifications would not be required.

J. Provide access to Written Reports and basic case information without requiring a PACER or CM/ECF account.