The U.S. Supreme Court and Information Technology: From Opacity to Transparency in Three Easy Steps

Jerry Goldman

Follow this and additional works at: https://scholarship.kentlaw.iit.edu/cklawreview
Part of the Law and Politics Commons, and the Supreme Court of the United States Commons

Recommended Citation
Available at: https://scholarship.kentlaw.iit.edu/cklawreview/vol88/iss2/5

This Article is brought to you for free and open access by Scholarly Commons @ IIT Chicago-Kent College of Law. It has been accepted for inclusion in Chicago-Kent Law Review by an authorized editor of Scholarly Commons @ IIT Chicago-Kent College of Law. For more information, please contact dginsberg@kentlaw.iit.edu.
THE U.S. SUPREME COURT AND INFORMATION TECHNOLOGY: FROM OPACITY TO TRANSPARENCY IN THREE EASY STEPS

JERRY GOLDMAN*

“In a democracy, the law is the operating system.” Carl Malamud

Today we face a flood of information. Data from all sources—some created by ourselves, most created by others—is available to us at our fingertips. Of the three branches of the federal government, the courts lag behind in their willingness to share information with the public. This comment proposes three relatively simple steps that the Supreme Court of the United States could take to provide timely and useful information. It may serve as a beacon for others to follow.

I. BACKGROUND

The Supreme Court was the last of the three branches of the federal government to share information online. Its website went public in 2000. The Government Printing Office handled day-to-day management of the website. All or nearly all information posted there was in the form of portable document format (PDF) documents.

In 2010, the Court internalized management of its website after requesting an increase in its annual budget to support this change.1 As a result, the current website hosts recent slip and bound volumes of its opinions (in PDF); current docket sheet and calendaring information; oral argument transcripts released the same day as the argument; and oral argument audio at the end of every week when the Court hears cases.2 Though the Court requires parties to submit briefs electronically as PDFs, with only two exceptions, it does not host the briefs on its website. Rather, the Court website directs its visitors to the American Bar Association, which hosts the briefs.

* Research Professor of Law, IIT Chicago-Kent College of Law and Director, The Oyez Project at Chicago-Kent.  
II. PROBLEMS WITH THE SUPREME COURT'S CURRENT SYSTEM

The Court’s website format may serve the Court’s purposes, but it is my view that the public is ill-served by the current website policy. The Court would better serve the public if it took three simple steps to lower the barrier to its current information sharing by focusing on (1) accessibility, (2) structure, and (3) standards.

A. Accessibility

Accessibility means getting to the data with ease. Public information websites, like the Court’s, should cater to two key audiences: the human beings who view and consume the content, and the numerous computer programs and algorithms that seek to parse and index the content. The Court seems entirely focused on information access for humans. This approach means that a search for any piece of information—an opinion, a brief, a docket file—relies on humans to gather information piece by piece.

Suppose you were interested in information related to the Stolen Valor Act case, United States v. Alvarez. You would probably aim to locate the briefs, the argument transcript, the oral argument, the opinion, and perhaps the docket sheet. Let’s start from the naïve view that the Court has made a decision involving the Stolen Valor Act. A search of the Court website for “stolen valor act” will return 500+ documents. The first, and perhaps the most important, is the opinion of the Court, followed by a copy of the oral argument transcript. The remaining documents seem to capture any variation of the word “act” or “stolen.” But these results do not reveal the oral argument audio, any of the briefs, or the docket file. Locating the oral argument requires that you drill down through the website to the page where the Court hosts oral arguments. Armed with either the title or docket number, you can find a link to a page where the audio has been hosted. Then you can decide which of the several flavors of audio you prefer, either to listen to on the spot, or download. And, if you searched for the audio before you searched for the transcript, you would find a link to the transcript from this page.

Let’s turn to locating the briefs in this case. The Court website is not particularly helpful on this score. It directs users seeking online information to subscription-based services, an out-of-date service that
appears to have ended in the 2007 Term, and the location of ten depository libraries where you can contact the librarian to ask for access and look for a print copy of the briefs.

The Court’s website directs users to online merits briefs hosted by the American Bar Association’s “Preview of United States Supreme Court Cases,” where you will see an alphabetical listing of cases on the current docket. But our goal was to find the briefs in United States v. Alvarez decided in the 2011 Term. Entering that search term in the website search field returns 961 results. Not a single one of the first page results has anything to do with Alvarez. Thus, there is no effective search for and no browsing of Supreme Court briefs. Typically, this is the end of the line, as frustration has mounted and life goes on.

For those who are clever enough to understand the syntax used by the ABA to catalog the briefs, it is possible to locate the collection of merits and amicus briefs on the ABA website. Should you then browse this collection you will discover that the briefs have been encrypted and watermarked with the logo of the American Bar Association. A naive person might ask: What is the purpose of encrypting the document? And why would the ABA add its logo to a document it did not create? These questions remain unanswered.

Suppose you now ask, are these all the briefs? In order to answer this question, you would need to compare the briefs listed on the Court’s own docket file against the briefs inventoried by the ABA. A few years ago, I undertook such an audit for an entire Term’s worth of briefs only to discover that the ABA website did not host all of them, so it lacked comprehensiveness. I have not audited the ABA website again, but given the inability to search or browse for briefs beyond the current Term, I remain skeptical unless I compare the briefs against the Court’s docket sheets.

B. Structure

Structure means indexing all information within a document or set of documents. Structure provides the path to greater accessibility. The docket file is the official road map to a case. It would be a simple step to provide a link from the docket to each brief, order, argument, argument transcript, and opinion. In fact, a step to achieve this simple

goal has been undertaken, not by the Supreme Court, but by SCOTUSblog for the current Term of the Court. SCOTUSblog has regenerated the Court’s docket file into a linkable roadmap, duplicating the same information found on the Court’s docket, but with links to petitions as well as merits and amicus briefs. Structuring information through a roadmap like the docket file improves accessibility for humans. This structure makes the task of collecting large bodies of data a relatively simple matter by scripting computers to do the work.

We humans are constrained by our most precious resource: human attention. Computers have advanced our knowledge across myriad fields from mapping the human genome to deciphering mobile communications on a grand scale. We live now in the era of big data where every bit of information is the potential object of analysis. We rely on computers to undertake such analyses, provided the data are in a format that computers—not humans—can understand. This means that structuring information at the start makes it easier for machines to harness that data and transform it, with the help of human theorizing, into information and perhaps knowledge.

The requirement for structured data holds true for the law. But one big impediment is the lack of structure in the information created by the courts, including the Supreme Court. When we humans look at a Supreme Court slip opinion, we can separate the title, the docket number, the decision date, the identification of the court below, the majority or plurality opinion, the concurrence, and the dissent. Humans are good at parsing the briefs to identify the party, the kind of brief (merits, amicus), the side (petitioner or respondent), the attorneys of record, and perhaps their titles, their firms and clients, addresses, and contact information. But providing structure to these documents would enable machines to collect and analyze this data for us humans. Absent consistency in describing the data contained in these documents (we call this metadata or data about data), we remain hamstrung in the ability to employ big data analytic techniques to the law. For example, someone might be interested in comparing merits briefs and opinions in a particular area in order to search for commonalities and whether any string of text then appears in other briefs or opinions in any other court. Linking briefs and opinions would be a straightforward task, but for the fact that the data are not structured consistently. Looking beyond the human task of reading and understanding a body of text, big

data confronts an impossible terrain when it comes to the U.S. Supreme Court, due to the absence of structure.

C. Standards

Standards are accepted benchmarks and conventions for any media type. We rely on standards every day. Take measurement, for example. In the United States, we rely on the British imperial system of distance measurement (inches, feet, miles), though the British and the rest of the world have largely abandoned it for the metric system (millimeters, meters, kilometers). But we stick to the standard because everyone understands it and it works, more or less. So standards are good for us. Few in the United States would entertain the use of the rod or furlong for a measure of highway distance. It might be appropriate at a racetrack, but not on the Interstate.

The Supreme Court has ignored standards in its recording of public proceedings. From its inception of public recordings in October 1955 through June 2004, the Court recorded its sessions on reel-to-reel acetate tape using the accepted standard of the day of 7.5 ips (7.5 inches of acetate tape passing over the recording head every second). This was the accepted archival standard for audio recordings. For a period in the 1990s, the Court switched from 7.5 ips to 3.75 ips, in effect reducing the amount of audio signal recorded to tape by half. The results were apparent in the reduced quality of the public session recordings.

With the demise of analog (tape) recording around 2004, the Court switched to digital recording. But rather than adopt the accepted archival standard for digital recording, the Court chose a different path. With the implementation of its digital recording system, the Court has been capturing its public sessions as MP3. MP3 is a delivery format, not a recording or capturing format. Recording in MP3 requires that a portion of the audio signal be eliminated or tossed out so as to reduce the file size and storage requirements of the resulting session. But throwing out or ignoring a portion of the audio signal may distort and surely reduces the quality of the resulting event. With storage media ever declining in price and with improvements in storage reliability, it is difficult to fathom why the Court would choose a non-archival format to record its sessions. This is especially true when delivery formats change. Archivists can always return to the recording format and regenerate a new delivery format. However, relying on a delivery format to generate a new delivery format is a formula for degradation.
III. PROPOSED STEPS

So how can the Supreme Court correct the deficiencies I have identified? I would urge the Court to take the following steps:

On accessibility: Host all Supreme Court content on the Supreme Court website, assuring both the integrity and the comprehensiveness of the website holdings. The Court has a professional information technology team dedicated to supporting the Court’s managerial and public-facing functions. It is difficult to fathom why the Court would cleave off its responsibility for public access to its briefs to a non-public entity that is neither authoritative nor accessible to the general public.

One small and virtually costless step to enhance access involves syndication, i.e., publishing website alerts to subscribers that new information has been added. It is now a commonly accepted practice for websites to alert the public that new information has been posted. Users subscribe to a website by relying on its RSS feed. This RSS feed is the digital equivalent of readers subscribing to a magazine or newspaper. If the public is largely ignorant of the Court and its work, one small step to reduce that ignorance is to encourage the public to “follow” the Court’s work through RSS. More importantly, computers can subscribe to the RSS feed and pull in new data, without banging on the Court’s website by scraping it.

On structure: The Court should adopt the practice of SCOTUSblog and make the docket a true digital roadmap to every case. For documents such as opinions, argument transcripts, and briefs, the Court should establish standards for the creation of these documents (the metadata) that will assure machine-readable access to their contents. The future of big data has arrived and the nation’s courts ought not to ignore the opportunities such structuring will present for knowledge and study.

On standards: Audio archivists have agreed on the best way to capture and preserve recordings. The Court should follow these standards and record its proceedings accordingly. The Court can and does deliver its content by the accepted conventions of the day, but delivery standards are not the same as recording standards, and the former should not be treated as the latter.

Three easy steps. Low cost and low risk. Everyone benefits.