



1200 5th Avenue, Suite 750  
Seattle, Washington 98101  
Phone: 206.624.1900  
Fax: 206.442.4396  
www.aokilaw.com

February 17, 2016

The Honorable Kathleen Cardone  
Chair, Ad Hoc Committee to Review Criminal Justice Act Program  
Thurgood Marshall Federal Judiciary Building  
One Columbus Circle N.E.  
Washington, D.C. 20544

**RE: Testimony of Russell M. Aoki, Coordinating Discovery Attorney on  
Technology, March 2, 2016**

Dear Judge Cardone:

Thank you for inviting me to address the Review Committee regarding technology. I look forward to the opportunity to discuss my work as a Coordinating Discovery Attorney (CDA) and how we use technology support tools for high discovery volume, multiple-defendant cases we address across the country. Because of my unique role, I will focus on the work of a CDA; how we are assigned cases; and our goals of using technology to help contain costs, maintain quality representation, and expose court-appointed counsel to technology support tools and services. Many pages can be written about each phase of our work. However, the intent of this letter, and my testimony, is to provide an overview.

I have been licensed to practice in Washington State for nearly thirty years. I have been a member of the Western District of Washington's Criminal Justice Act (CJA) Panel for approximately twenty-four years. In 2005, our jurisdiction, under the leadership of U.S. District Court Judge Marsha Pechman, created best practices policies for large document cases and wiretap surveillance evidence. To ensure parties were complying with the policies and to develop further case management efficiencies, the District Court began appointing me as a CDA, to help defense counsel develop strategies and implement ways to organize large volumes of discovery. In 2011, I contracted with Administrative Office of U.S. the Courts, Defender Services Office (DSO) to provide the same services across the country. I have been appointed by U.S. District Courts over 60 times.

## **The CDA's Role**

The pervasiveness of technology is evident in the massive discovery productions increasingly seen in federal criminal cases. Government discovery productions now comprise electronically stored information (ESI) from servers, desktop computers, laptops, and mobile devices. The rise of social media has also increased discovery with data from popular sites like Facebook, Instagram and Twitter. Government-generated evidence includes audio recordings from wiretaps, body wires, and GPS tracker data. Videos are common, including pole-camera recordings, business security videos, and concealed-camera footage. Managing large volumes of discovery is a daunting task for a defense attorney. Time that could otherwise be dedicated to case analysis can be lost trying to navigate through sometimes millions of discovery items.

To assist court-appointed counsel, DSO contracts three CDAs to consult and provide organizational strategies for defense counsel on cases with many defendants, voluminous discovery, or complicated electronic discovery issues. I am one of these three CDAs, along with Shazzie Naseem of Kansas City, Missouri, and Emma Greenwood of New York City, New York. We each have experience working on CJA cases and are knowledgeable about how technology can organize, search, review, and analyze large volumes of discovery, while maintaining evidentiary integrity and avoiding conflicts of interest. Our work typically includes:

- Managing and distributing discovery produced by the Government and relevant third-party information common to all defendants;
- Assessing the amount and type of case data to determine what types of technology should be evaluated and used, so duplicative costs are avoided and the most efficient and cost-effective methods are identified;
- Acting as a liaison with federal prosecutors to ensure the timely and effective exchange of discovery;
- Identifying, evaluating, and engaging third-party vendors and other litigation support services;
- Assessing the needs of individual parties and identifying any additional vendor support that may be required—including copying, scanning, forensic imaging, data processing, data hosting, trial presentation, and other technology depending on the case;
- Identifying any additional human resources that may be needed by the individual parties for the organization and substantive review of information; and

- Providing training and support services to the defense teams.

### **Obtaining the Services of a CDA**

CDAs are assigned by DSO's National Litigation Support Team (NLST). The criteria used to determine the assignment of a CDA are:

- Whether the number of co-defendants is so large as to create a risk of costly duplicative efforts which could otherwise be eliminated or reduced upon the appointment of a CDA, or whether there are other factors that create a likelihood that the CDA's participation would enable costs to be contained;
- Whether the volume of discovery is so large that addressing the organizational needs in the case would interfere with defense counsel's ability to address the legal and factual issues in a case;
- Whether unusual organizational or technological issues exist, not commonly found even in complex cases, that would interfere with defense counsel's ability to address the legal and factual issues in a case;
- Whether the case is prosecuted in a region that lacks experts who can provide necessary technology support and document management expertise in addressing the factors described above;
- Whether the timing of the request, which preferably should be made early in a case, is such that the CDA's participation is likely to be of assistance to defense counsel, promote efficiency, and contain costs; and
- The CDA's work load.

### **Technology Strategies to Contain Costs**

We look for cost savings at all levels of our work. Our goals are to avoid duplicate work of defense teams, create discovery management tools for more efficient use of attorney time, and assess third-party services to ensure the work contracted for is necessary and the cost is reasonable.

#### ***Avoiding duplicate work***

Discovery management efficiencies can be realized at all phases of our work, starting with when initial discovery is produced. Too often, a CDA gets appointed after the first production of discovery. Discovery has already been formatted and

distributed. It is understandable that defense counsel may not recognize the need for discovery management assistance until the first production of materials are received and examined. However, the optimum time for a CDA to become involved is prior to the first production. At an early pre-production "meet and confer" with the Government, a CDA can help avoid unnecessary costs by identifying appropriate file formats for the particular case. Cases involving a few hundred pages of discovery may be suitable to be formatted into one large single PDF. The single PDF can be keyword searched using simple PDF readers. But for cases involving millions of pages of documents, large multiple-document PDFs are unworkable. For search tools to locate individual documents within multi-document PDFs, either the Government must produce them again as individual documents, or the defense will be required to engage in time-consuming and expensive manual identification of page breaks to disassemble the PDFs into single documents. A CDA's early involvement can avoid this dilemma. The CDA can also look for cost sharing opportunities with the Government, such as when a collection of materials is important to the defense but of uncertain interest to the Government. The cost of collecting and producing the discovery could be shared.

As CDAs, we ask to be the central distribution point of all discovery to avoid each defense counsel performing the same initial examination. Our office looks at the discovery to ensure the materials are complete, without corrupted files, and, where applicable, that the documents are searchable. We convert unusually formatted discovery items into a form that can be reviewed on a Windows or Mac computer, and make scanned documents searchable through optical character recognition (OCR). Discovery production errors are immediately addressed with the Government so replacement materials can be obtained.

Something as simple as consolidating individual discs of discovery onto a hard drive can save substantial attorney time. We have seen several cases where the Government produces discovery on hundreds of single discs. To review the materials, defense counsel would have to shuffle through the many discs to find a particular production, then place it into their CD/DVD drive, wait for it to load, and, after examining the contents, remove the disc. Without including the time to review the materials, the actual process of finding the right discs, loading and unloading could take, for example, three minutes. If there are 500 discs, finding, loading and unloading could equal 1,500 minutes. If 30 lawyers are on the case, the attorney time quickly escalates to 45,000 minutes, or 750 hours. At CJA rates the cost could reach over \$90,000. Loading all 500 discs onto one hard drive no longer requires swapping in and out hundreds of discs. The hard drive is broken down by multiple folders. Moving from one production to another is just a matter of a few mouse clicks.

Our office also ensures materials can be opened and viewed in both Windows and Mac computers. Many proprietary programs used for surveillance work are Windows-compatible only. We reformat the discovery in more common MP3 format

for audio and MP4 format for video, which allows the materials to be played on either computer operating systems. Much of the work is performed in-house at a substantial savings, versus third-party vendors which typically charge based on the length of the recording. Sometimes, small collections cannot be converted in-house; in these cases, the materials are sent to one of several vendors with whom the CDA has worked and knows to be knowledgeable and reasonably priced. Some video formats cannot be converted into a common file format. Our office uses screen-capture software to re-record these videos. This is a time consuming task, as the pace of re-recording is in real time where one hour of video takes one hour to record. In a recent case, our office used screen capture software to re-record hundreds of hours of jail visits, at a savings of tens of thousands of dollars based on the cost had the work been performed by an outside vendor. Reformatting to common file types is also critical if the discovery is to be used on any discovery management program, such as a database. By reformatting, the proprietary player is no longer required. Any computer system, including database servers, can host and play these formats. The benefit of reformatting is not only for Mac users but for the greater need to organize all discovery types into a central database.

Depending on the volume and discovery types, defense counsel can expect to receive discovery from our office within a few days, unless audio or video reformatting requires outside assistance or must be re-recorded. If such is the case, it could take several weeks.

### ***Providing efficient discovery management tools***

As CDAs, we look for the most cost-effective strategies to manage discovery. We consider the volume of discovery, and how much of the materials are more than just documents. Inexpensive programs such as Excel are often initially considered. With an Excel spreadsheet, discovery can be individually itemized and hyperlinked. Reviewers add information from the discovery such as document dates, Bates numbers, defendant names, document type, event dates, and other names mentioned so defense counsel can sort through the materials to find the discovery that is most pertinent to their clients. Defense counsel can merely click on a hyperlink and the document will appear, the photograph will come up, or the audio or video file will play. This process is quicker than going through a drive and opening each file individually, but is still time-consuming to create because reviewers must read each document, and therefore best used for cases with small collections of documents.

For wiretap discovery, we have worked with several technology companies to develop methods for associating wiretap linesheets with the audio recordings. Using an automated process, the information from each linesheet is extracted and put into an Excel spreadsheet. Hyperlinks are created for the linesheets and for the wiretap audio. For a typical drug case that may have up to 30,000 wiretaps, it used to take our office two-to-three months to manually create such a spreadsheet using contract

reviewers, at a cost of \$10,000 to \$15,000. Now, the work can be done in one-to-two weeks at \$2,000 for the vendor's service. The key to this advance method is to obtain high quality discovery that can be accurately searched. With the spreadsheet, defense counsel can sort by categories such as the target number; the telephone numbers dialed or those of incoming calls; whether the agent identified the calls as pertinent; and the date, duration, and synopsis of the call. This is one of the immediate tools we try to produce within the first few weeks of any major case involving wiretap evidence.

Sortable spreadsheets were once all that were necessary for many drug cases, because discovery used to be comprised primarily of approximately 20,000 to 50,000 wiretaps and 1,000 pages of related Title III pleadings and law enforcement reports. Today, the discovery encompasses GPS data from both law enforcement tracking devices and personal cellphones; cell tower data; contact directory information; text messages; photos; social media captures; and surveillance video from pole cameras, hand-held devices, and business security systems. When the volume of discovery is exceptionally large, the cost for reviewers to code discovery grows prohibitive, and the features of a sortable spreadsheet become inadequate to target key documents. For large volumes of discovery, we look for more effective methods such as a web-based database.

A web-based document database provides sophisticated search, review, and annotation features. Users can log in from any location through a secure Internet portal and review a single source of discovery. They can target documents using complex keyword searches, eliminating the need to review and code each document and sparing the associated cost and time. When selecting a database company, we look for a robust security system, ease of use, and features that allow counsel and their clients to work collaboratively. We typically interview several companies, check references, and ask for bids. Based on the service, costs, references, and input from defense teams, we will make a recommendation to counsel and petition the court for funds.

### ***Assessing third-party vendor services***

The CDAs stay current with the latest technology trends applicable to criminal cases. However, most litigation support programs cater to civil litigation, which means focusing on the management of documents and emails. For us, it is necessary to find a discovery management tool for the mixed media discovery often found in drug and death-eligible gang cases. My office has been working with a small web-based database company to modify its database, at its own expense, so it can accommodate mixed media discovery. It has now evolved to where it is Mac compatible; audio and video quickly stream instead of waiting for time-consuming downloads; jail calls can be found by defendant and the number dialed; and video-recorded jail visits can be searched by defendant, date, and time. The sortable wiretap spreadsheet data can be entered into the database to allow the association

of the linesheets with the audio recordings. This allows keyword searches of linesheets within the database to call up recordings of intercepted conversations based on the categories set out in the linesheet.

For larger fraud cases, the choice of databases is far greater because discovery is primarily documents and emails. Typically, we obtain non-disclosure agreements from several prospective vendors. Upon execution of the agreement, they are asked to review the discovery and provide a cost proposal that includes making the database available through the scheduled trial date. Unless additional discovery is produced, the vendor cost is capped. If the actual expense is less, the Court will pay the lesser amount. If the cost goes beyond the capped amount, the vendor will only be compensated for the quoted price.

There are typically two costs to web-based databases. The first is for loading the discovery into the database. It is a one-time fee for each new production of discovery. The second is monthly hosting. Hosting is based upon the digital size of the materials on the database. As CDAs, we ensure rates are competitive.

Comparing rates is often not an easy task. Vendor proposals are often confusing in how they lay out processing and hosting costs. Recently, some vendors have revised their pricing model. One vendor does not charge separately for processing but, instead, increases the monthly hosting charge to capture the cost of their initial work. Another vendor is considering a pricing model of both processing and hosting wrapped into a single fee of one penny a document. The CDAs can use their experience with the overall cost of multiple database cases, and prior dealings with vendors, to evaluate which of these differently-calculated rates provides the best value.

Based upon the proposal for work to be performed and the projected costs, I select three vendors to demonstrate their database to the defense teams so they may provide their input. After the demonstrations, a vote is taken as to which platform to use. I then petition the Court for funds. If the circuit has a Case Budgeting Attorney, that person is alerted of my funding request and receives a copy of my petition.

Upon approval, the database company loads the materials onto its review platform, which could take several days to several weeks depending upon the volume and the file types. Meanwhile, my office schedules training sessions with the defense teams to take place several weeks out. Once the database is loaded, and as another cost savings measure to reduce defendant's work, my office performs initial organizational work. They conduct keyword searches of defendants' names as they appear in the indictment, including any nicknames, and either tag the discovery or put it into defendant-specific folders. If organizations or unindicted individuals are repeatedly mentioned in the indictment they are also tagged or folders are created.

These are all searches likely to be run by each defense team. By having our office perform the work, duplicate attorney and paralegal time is avoided.

### **Maintaining Quality Representation**

With each case, we look for opportunities to help defense counsel locate and manage materials. The work we do often results in cost savings, as well, since we increase review efficiency. Mega-cases are a concern to defense counsel because the discovery is so voluminous it is impossible to review every discovery item. Instead, defense counsel must adopt strategies to analyze materials, often by first breaking down the collection into smaller subsets using keyword searches, and then continuing to refine the searches until the material has been narrowed to a manageable amount for review. It is our job to assist them with the process by conducting the initial creation of subfolders and providing continuous training, so counsel develops the skills to effectively use keyword search terms.

The opportunities to help maintain quality representation differ with each case. An important and common factor for any mega-case is receiving discovery with time for sufficient review and analysis.

#### ***Prioritizing productions***

If a CDA is appointed early enough, we advocate having the investigative reports such as FBI 302s, DEA-6s, and search-warrant pleadings produced first. It is the most helpful discovery for counsel to understand the case in its early stages, and allows for an early determination of whether resources should be devoted towards preparing for trial or negotiating a suitable disposition. However, discovery productions sometimes roll out based upon what is first available, which are often the clients' forensic computer hard drives. Some jurisdictions will accommodate our request for the immediate release of investigative reports and pleadings. However, some jurisdictions still insist that such important documents should be produced later.

Courts and counsel sometimes ask for strategies to increase attorney efficiency, to reduce attorney time and costs. From my own experience, I know the early stage of discovery review is spent trying to understand how the Government plans to prove its case. To accelerate that process, I often recommend a case schedule requiring early disclosure by the Government, typically months prior to trial, a preliminary exhibit list and preliminary list of witnesses with a reciprocal requirement from the defense. Once preliminary exhibits and witnesses are known, defense teams can better strategize how to target relevant materials within the discovery. The early disclosures help develop strategies and avoid the temptation to linearly review discovery item by item to understand a case.

Aside from recommending when and how discovery is to be produced, maintaining quality representation differs from case to case. Below are examples of how we look for opportunities to help.

### ***Avoiding Hosting Unnecessary Productions***

Large discovery cases are sometimes the product of the Government's uncertainty about how it will use materials to prepare for trial. On several cases, we received hundreds of thousands of emails in native file form. Later, as the Government prepared for trial, the same emails were produced again, but as PDFs so Bates numbers could be added. This complicated discovery review by doubling the haystack. Key word searches that previously called up 500 documents now returned 1,000. Even if defense counsel had marked prior emails as "reviewed" after examining them on the database, the tag would not include the newly produced duplicate email. And though most databases can identify and screen out duplicate documents, that capability was ineffective because any database would consider the new version different discovery both because it was in an alternative format (PDF instead of native), and because it contained additional information (a Bates number). Review bogged down, and data hosting costs increased with the new discovery.

The native files cannot simply be discarded in favor of the PDFs because they contain potentially important metadata such as the To, From, Date, and Subject fields. These metadata fields can be used to screen and categorize discovery during review, and, due to their original electronic format, the accuracy of searches performed on native files is very high. In contrast, the PDF versions have no useable metadata, instead requiring a search of the actual text to find the same information. Searches can be less accurate depending upon the quality of the PDF. But the PDFs are also indispensable because they have Bates numbers, which are often referenced by the Government in the final identification of trial exhibits.

The solution is to load both versions into the database, match up the native files with the PDF using a computer process, extract the Bates numbers from the PDFs and add them to a field of the native files, and then remove the PDF version to avoid increased hosting costs. In one instance, we avoided adding a terabyte of otherwise duplicate PDF emails. However, this is not always possible—in another case, the PDFs file names were different enough they could not be matched and, consequently, an additional 300 gigabytes had to be added to the database.

### ***Making the non-searchable searchable***

Besides working with discovery produced by the Government, we look for materials that *should* have been produced. In several cases, the Government produced tens of thousands of wiretap recordings without the accompanying linesheet documents. Without this data, defense teams would have been hampered in their ability to locate key calls. The only information they might have would be the

telephone target number and date if contained in the name of the file. The Government was immediately notified, and linesheets were eventually produced. With the linesheets, defense counsel could also search for the incoming or outgoing telephone number, duration, time, words within the synopsis and whether the agent classified them as “pertinent.”

My office continuously quality control checks the materials. One thing we look for are materials that do not have searchable text. This may mean they are photographs or scanned documents that were not properly OCR'd. In the latter case, we attempt to OCR the materials again or, if that does not make the documents searchable, we will engage reviewers to code the documents by document type, date, defendant names, and other names mentioned so they can be retrieved with general keyword searches.

Similarly, many cases involve handwritten documents by inmates, like letters and journals. Since handwritten documents cannot be accurately OCR'd, we identify handwritten documents and code them, as we would with other documents that cannot be OCR'd, to make them searchable. Without the coding, defense counsel could not locate these key documents among the hundreds of thousands and sometimes millions of discovery items.

## **Training**

We are involved in training on many levels. We train ourselves on the new forms of discovery production and discovery formats. We also train vendors on how to handle the kinds of discovery typically seen in federal criminal prosecutions, and we train CJA panel lawyers on how to use technology tools.

We have created a resource network of technology support consultants who help us with technology issues. The NLST, in particular, has been a tremendous asset. We also routinely participate in demonstrations of new web-based database platforms, and other technology services such as predictive coding.

We have worked with several vendors to create the wiretap spreadsheets described above and also database companies on the need to modify their database program to accommodate the mixed media discovery such as wiretaps, surveillance video clips, cell phone reports with active hyperlinks to cell phone contents, jail calls with metadata identifying inmates, dates and numbers dialed, and similar video of jail visits.

Most importantly, we dedicate time to court-appointed counsel to learn how to use technology to manage and review materials. We provide step-by-step instructions on encryption, using Acrobat search capabilities on PDF documents, how to key word search and use the filtering features of Excel spreadsheets, and using web-based databases.

For sophisticated web-based databases, our office prefers to train counsel with the vendor or, for the most complex database we use, conduct the training ourselves. Web-based database vendors cater to civil litigation, and are unaccustomed to locating mixed media discovery or addressing the concerns particular to criminal cases. Our office is well versed on the needs of defense counsel to locate certain types of discovery.

We advocate for database training in two phases. The first is group training sessions where all defense counsel are asked to participate. They are shown the initial organizational work performed by my office. We ask them to consider other common organizational needs my office may provide. After the general training session, each defense team is invited to private training sessions where they can learn how to find particular materials pertinent to their respective clients, without fear of inadvertently revealing case strategy or highlighting damaging evidence to the other defense teams.

The work of a CDA is challenging. To be most effective, it is necessary to understand federal criminal prosecutions and how CJA panel attorneys operate their businesses. Few CJA attorneys have staff, and few have spare time to learn complex technology skills. Understanding CJA counsel is our starting point in strategizing efficient and cost effective discovery management tools.

Very truly yours,

Russell M. Aoki

RMA:nbc