Privacy Impact Assessment
for the

Electronic Discovery (eDiscovery) Tools for Litigation Use

DHS/ALL/PIA-073

May 28, 2019

Contact Point
Michele Procter
Knowledge Manager
Office of General Counsel
Department of Homeland Security
(202) 447-4347

Reviewing Official
Jonathan R. Cantor
Acting Chief Privacy Officer
Department of Homeland Security
(202) 343-1717

Abstract

The Department of Homeland Security (DHS) Office of the General Counsel (OGC) and Component Offices of Chief Counsel (OCC) use commercial off-the-shelf electronic discovery software tools to facilitate the production of documents and disclosure of existing records during litigation or in response to a request for records. eDiscovery is a document processing method that supports the organization of paper and electronic documents for analysis, review, redaction, and production to meet litigation discovery requirements. DHS also uses eDiscovery tools to process agency records in response to subpoenas and *Touhy* requests (written requests for testimony or agency records or official information made in accordance with agency regulations in cases to which the United States is not a party).[1] DHS is conducting this Privacy Impact Assessment (PIA) because this process collects, maintains, stores, and shares personally identifiable information (PII).

Overview

The Department of Homeland Security (DHS) uses eDiscovery tools[2] to support the review, redaction, and production of agency records to comply with discovery requirements during litigation and to provide responsive documents.[3] eDiscovery tools replace the manual processes by which Office of the General Counsel and Offices of Chief Counsel (OGC/OCC)[4] gather, sort, review, and redact agency records that are potentially responsive to a discovery, subpoena, or *Touhy* request;[5] assesses its relevance and/or responsiveness; and applies appropriate privileges in litigation and other requests.

DHS increasingly relies upon electronically stored information (ESI) to conduct departmental business. Accordingly, DHS needs to use eDiscovery tools to review, redact, and produce Department records. In some cases, DHS may have an obligation to produce ESI in the same format in which it is originally compiled or maintained ("native format"), along with any associated metadata.[6] Additionally, OGC/OCC may scan paper documents into eDiscovery tools and review them, as it would an electronic document.

---

[1] *United States ex rel. Touhy v. Regan*, 340 U.S. 462 (1951).
[2] A full list of the eDiscovery tools currently at use within the Department can be found in Appendix A of this PIA.
[3] In this document, "responsive documents," means agency records that are potentially responsive to a discovery, subpoena, or Touhy requests.
[4] Throughout this PIA when the OCC acronym is used, it includes the Office of the Chief Counsel of U.S. Citizenship and Immigration Services, Cybersecurity and Infrastructure Security Agency, U.S. Customs and Border Protection, the Federal Emergency Management Agency, the Federal Law Enforcement Training Centers, the U.S. Secret Service, and the Transportation Security Administration as well as the Office of the Principal Legal Advisor for Immigration and Customs Enforcement and the Office of the Judge Advocate General of the Coast Guard.
[5] *See* 6 CFR §§ 5.41-5.49; 44 CFR §§ 5.80-5.89.
[6] Metadata are the data attributes that may or may not be hidden that may reveal sensitive information about a document or file's history, such as its creator, last editor, or version history, among others.

The range of capabilities varies across the Department, but eDiscovery tools generally streamline and automate the document review process. First, the tools load and analyze information in various data formats (*e.g.*, Microsoft Word, Microsoft Excel), allowing document analysis in bulk within a single data file and using a single integrated viewer that does not require use of the original application that created the file. Second, eDiscovery tools allow operators to view metadata within files stored in these varying file formats. Third, they identify and eliminate duplicate documents from the review process. Fourth, the tools automate the identification of protected information by searching for names, phrases, and terms (collectively, "keywords") that the reviewing operator inputs. The eDiscovery tools are able to use that information to automatically flag files that contain those keywords for the operator, who then reviews and determines whether the files or information therein are applicable for their purposes of the search. Finally, these tools allow operators to electronically redact protected portions of documents in the system.

eDiscovery tools have additional features that speed up the process by which OGC/OCC personnel designate or redact the same protected information from multiple records. For example, eDiscovery tools allow for the bulk redaction of any words or terms, including names or identifiers, for the attorney identify all occurrences of a keyword associated with a particular lawsuit so that the attorney can decide whether it is appropriate to redact each instance of that keyword as it appears in multiple records.

Background

DHS uses eDiscovery tools to facilitate the efficient compliance with federal requirements to preserve and produce ESI in civil and criminal litigation matters according to the Federal Rules of Civil and Criminal Procedure. These tools significantly improve the efficiency of OGC/OCC's processing of records during discovery in litigation. OGC/OCC's discovery productions can require the preservation, collection, and analysis of tens of thousands of emails, word processing documents, Portable Document Format (PDF) files, spreadsheets, presentations, database entries, and other documents in a variety of electronic file formats, as well as paper records. The current manual process of preserving, collecting, and analyzing those records can be burdensome and inefficient. For example, given this volume of records, OGC/OCC cannot possibly review all documents for metadata and privileged information under this current process, and needs the automation eDiscovery tools enable. Many of the discoverable documents are duplicates, and because it is difficult to manually identify duplicates among voluminous records, OGC/OCC may be burdened reviewing multiple identical documents. The automation of this process using these tools dramatically reduces the time OGC/OCC spends on administrative tasks related to document management and improves the quality and efficiency of overall document review and production within OGC/OCC. Only a small number of OGC/OCC personnel will receive access to the system: those involved in litigation, and those responding to other document requests.

Document Collection Process

In lawsuits, parties typically begin the document review and production process after filing litigation against the agency or in a case in which the agency may have an interest (such as when a civil lawsuit is initiated against DHS). DHS also may start gathering documents in anticipation of litigation based on particular circumstances. Once OGC/OCC becomes aware of the need to preserve records, it issues a litigation hold notice describing the information and records that may be discoverable in the context of that litigation.

OGC/OCC may also use eDiscovery tools to respond to other forms of document requests. For responsive document requests in which DHS is not a litigant, OGC/OCC gathers, reviews, and marks material as exempt, then produces the records to the requester, if appropriate. For document requests linked to litigation, OGC/OCC issues a litigation hold for relevant documents. The notice informs employees who may be custodians of such data that they are to preserve and/or produce it to OGC/OCC for review. Individual DHS employees and technical support personnel then take action to preserve the evidence described by the litigation hold notice. These actions may include identifying additional keywords that allow employees and technical support personnel to identify other relevant documents. OGC/OCC emails individual employees who may be custodians of requested information with separate litigation hold notices prohibiting them from deleting or destroying evidence, whether in paper or electronic form. The individual employees are not generally required to identify, harvest, and produce evidence until the case is in litigation and discovery commences. The requirements are dictated by the litigation itself.

When discovery commences, OGC/OCC notifies employees of their obligation to identify, harvest, and produce evidence to OGC/OCC. The role of technical support personnel varies, depending on the stage of the litigation and the media on which data are likely to be stored. OGC/OCC may require these technical support personnel to search all locations where responsive ESI might be stored, including central agency databases, agency file servers (*e.g.*, shared drives), and centrally stored agency electronic mail for records described in the litigation hold. In some cases, based on the needs of the attorney and requirements from the court, technical support personnel may initially set aside any back-up tapes and files containing relevant information and physically preserve them in their original form. In other cases, technical support personnel may have to search electronic storage systems for relevant agency records, which will be downloaded to portable storage media or drives maintained on servers. In extreme cases, when an employee likely possesses a significant amount of ESI on an individual work station or storage medium, technical support personnel may make images or copies of the entire storage medium for preservation purposes.

Once relevant information has been identified and litigation ensues, data and documentation are transferred to the respective OGC/OCC office. The collection may include bulk scanning of paper documents into an electronic format, such as PDF or Tagged Image File Formats

(TIFF), and preferably into a machine-readable format. Once all of the data is stored in a secure format, the data is then imported into an eDiscovery tool. The data stored in the secure format is then deleted once it is confirmed that the data has been successfully uploaded into the tool. The attorney with the case file will maintain the hardcopy records in their original form. If litigation does not ensue, OGC/OCC lifts the litigation hold when the statute of limitations expires, or when OGC/OCC otherwise concludes that litigation is not reasonably likely. DHS will maintain or delete the records that were covered by the litigation hold, but never uploaded into eDiscovery tools, in accordance with normal agency retention policy, as set forth in any applicable records disposition schedules.

Document Review Process[7]

eDiscovery tools support hundreds of different file formats for review in a native viewer, avoiding the need to install the application used to create the document onto the reviewing attorneys' computers. This also eliminates the necessity of converting documents into formats that OGC/OCC can view and redact using their current computer configurations, thereby reducing the risk that others will alter the documents and potentially violate federal evidentiary and discovery rules. eDiscovery tools can collect and process various formats such as TIFF, PDF files, JPEG images, and Microsoft Office documents. The documents loaded into these tools are exact duplicates of existing data that are already stored in other DHS paper or electronic recordkeeping systems. DHS maintains the documents it loads into eDiscovery tools in their original, unmodified form. The tools do not create new information that is associated with those records and do not alter the integrity of the original records themselves. The data in the eDiscovery tools consists of redactions;[8] tags; privilege logs[9] created by the software; search and filter reports; and an audit trail, which the tools may automatically create and maintain as an historical record of all actions users take in each case. eDiscovery tools may also assign a unique key to the original unaltered file, which helps establish the chain of custody as proof that no one altered the content of the produced file or the original version of the document.

eDiscovery tools receive uploaded records in a new "case" that they create for a particular litigation. The OGC/OCC supervisory attorney will work with administrators of the tool to create user accounts and grant rights to the privileges within the tools to the attorneys and paralegals assigned to that litigation, allowing them to access and review the documents. During the review process, attorneys may narrow the scope of documents reviewed by executing searches and filtering data, generally using search terms agreed between the litigants. As part of this initial

[7] While the specific document review process may differ from Component to Component, this section generally outlines the process.
[8] Redactions are not considered alterations of the documents. They merely hide information that should not be produced to opposing parties.
[9] A privilege log lists the location and basis of each redaction or withholding, cross-referencing each redaction to a specific page number within the production.

review, eDiscovery tools may automatically identify and remove duplicate documents from the collection of data in the tool's repository.[10] Attorneys and paralegals then review the subset of documents resulting from the search and filter and place tags on specific documents to classify and categorize those documents. In addition, they redact protected or privileged information. For each document, OGC/OCC may enter free-form text describing the reason for the redaction.

Using this information, a privilege log is generated to document the redactions or withholding of records on the basis of privilege. OGC/OCC typically shares this privilege log with the U.S. Department of Justice (DOJ), with other parties in the litigation, and sometimes with the court. eDiscovery tools can also generate reports based on the search terms and filters that were used to withhold records. OGC/OCC produces the search report to opposing parties to demonstrate a defensible process for gathering the totality of relevant data, as agreed upon between the parties. The presiding judge may also order OGC/OCC to produce search reports. Once the attorneys and paralegals complete the initial document review process, other attorneys (including supervisory attorneys) may conduct a quality review assessment to verify the accuracy and appropriateness of redacted and un-redacted information, as warranted by the case.

Once the review is complete, DHS attorneys place the reviewed records in a production folder in the eDiscovery tool indicating that they are ready to be produced to DOJ, and eventually, to the court and opposing counsel. System administrators of the tool place the reviewed records in the appropriate file format for production, which varies and depends on the agreement among the parties, or an order from the court. Production file formats are typically image files, such as PDFs and TIFFs. While redactions in eDiscovery tools are temporary (*i.e.*, the users can view the content underneath the redaction as needed), once those records are saved into a production format the redactions are permanent. eDiscovery tools may still retain a copy of the original records with only temporary redactions, which allows OGC/OCC to change the redactions if needed. The originals are also retained, should DHS need to produce the same records again in cases where the parties' agreement or a court order requires DHS to produce previously redacted information from those records or change the redactions. Files in any format included in the production file may contain PII data, unless DHS attorneys redact it prior to production.

Before production, the system administrator extracts the production file from the tool, encrypts it, and then writes it to an external portable storage device, such as a CD-ROM,[11] for transfer to the recipient. Once documents are sent to DOJ, and eventually to the court and opposing

[10] De-duplication is based on the use of hash values. Records are assigned a hash value based on a combination of the content of the file and the metadata associated with the record. Only one copy of records that have identical hash values are maintained.

[11] The CD-ROM is generally the most-used external portable storage medium for the transfer of large files. There is no standardized process for deciding to use a CD-ROM over other external portable storage devices. Other transfer mechanisms include DVD, encrypted drive, or USB Iron Key.

counsel, the OGC/OCC attorney will "close" the case, and OGC/OCC and system administrators will delete the matter from the eDiscovery tool and save an archive file of the matter.

Section 1.0 Authorities and Other Requirements

1.1 What specific legal authorities and/or agreements permit and define the collection of information by the project in question?

In civil cases, Federal Rules of Civil Procedure 16, 26, 34, and 37 govern most electronic discovery requirements, which the federal courts may enforce. In criminal cases, courts can compel full and open discovery of agency records via the Fifth and Sixth Amendments to the U.S. Constitution and the Federal Rules of Criminal Procedure, particularly Rule 16. For subpoenas and *Touhy* requests, Rule 45 of the Federal Rules of Civil Procedure and the Administrative Procedure Act[12] mandate the disclosure requirements.

1.2 What Privacy Act System of Records Notice(s) (SORN(s)) apply to the information?

In the context of litigation, the DHS General Legal Records SORN[13] applies to eDiscovery data gathered in the context of litigation. Specifically, the DHS General Legal Records SORN covers all agency records that are potentially responsive to a discovery, subpoena, or *Touhy* requests. In addition, the DHS General Information Technology Access Account Records System SORN[14] covers access to these types of tools and resources by authorized individuals.

1.3 Has a system security plan been completed for the information system(s) supporting the project?

Yes. eDiscovery tools are required to undergo a security authorization process. Each tool must be granted an Authority to Operate (ATO) prior to being deployed, or be covered by a General Support System Authority to Operate, based upon whether the tool is a Major Application or part of a General Support System respectively.

1.4 Does a records retention schedule approved by the National Archives and Records Administration (NARA) exist?

Litigation case files are retained under National Archives and Records Administration (NARA)-approved retention schedules, generally specific to each Component.

---

[12] 5 U.S.C. § 500 *et seq.* Components may have additional regulations that specifically apply to that Component.
[13] DHS/ALL-017 Department of Homeland Security General Legal Records, 76 FR 72428 (November 23, 2011).
[14] DHS/ALL-004 General Information Technology Access Account Records System, 77 FR 70792 (November 27, 2012).

1.5 If the information is covered by the Paperwork Reduction Act (PRA), provide the OMB Control number and the agency number for the collection. If there are multiple forms, include a list in an appendix.

Because the records are already in DHS's possession and not collected from sources outside of DHS, eDiscovery tools use are not subject to the provisions of the Paperwork Reduction Act (PRA).

Section 2.0 Characterization of the Information

2.1 Identify the information the project collects, uses, disseminates, or maintains.

eDiscovery tools store and process agency records, as necessary, to satisfy litigation discovery requirements and to respond to document requests. Information in eDiscovery tools could consist of any ESI or other information in any DHS formal or informal recordkeeping system or any paper documents scanned into an electronic format for review.

Because the tools are document processing systems, the ESI and other records that may be stored and processed could pertain to any matter in the scope of DHS's mission and may contain PII or sensitive personally identifiable information (SPII) of any nature captured and stored in such records. For civil litigation that is reasonably likely or pending, eDiscovery tools may collect and maintain any information that is potentially relevant to the matter for discovery purposes. To determine whether a document is "potentially relevant," OGC/OCC attorneys generally review the case's history to gain an understanding of the litigation itself, review search terms, and read the documents to determine if they may be responsive to the litigation. To the extent it is applied in the processing of non-litigation related document requests, eDiscovery tools may store and process any agency records that are potentially responsive to those requests. The actual information stored and processed will always vary and depend on the nature of the particular litigation or document request.

The types of individuals about whom information could be collected varies on a case-by-case basis, but may include any of the following: anyone involved in litigation with DHS, applicants for DHS benefits, persons who file responsive documents requests asking for DHS records, persons who correspond with DHS, employees and contractors of DHS and other federal agencies, witnesses and other sources of information, attorneys and authorized representatives, subjects of investigations, and others whose information is contained in the records collected during the course of an investigation, enforcement matter, or other matter of any kind handled by DHS.

Listed below are examples of general types of records that eDiscovery tools may store or process:

- **Electronic mail:** messages among DHS employees, or among DHS employees and personnel of other federal agencies or outside entities, sometimes with other documents attached;
- **Presentations:** documents such as PowerPoint presentations;
- **Spreadsheets**: typically data collections or tracking of broad information such as aggregate expenditures during a disaster;
- **Database entries:** information collected or compiled from program databases which could contain PII; and
- **Miscellaneous:** letters, memoranda, drafts, receipts, photographs, images, video recordings, etc.

eDiscovery tools store and maintain may also contain metadata, which may contain PII (*e.g.*, the name of the author of a particular electronic file[15]), which itself may be discoverable in litigation or by a document request. As described in the Overview, these tools support hundreds of different file formats that can be reviewed in a native viewer, such as TIFF, PDF files, JPEG images, Microsoft PowerPoint documents, and Microsoft Word documents. The specific PII collected in these records will vary based on the nature of the records themselves, the breadth of the request, and the nature of the request.

DHS records may contain the following information from:[16]

System Users

- Identity Credential/Access Management certificate and Personal Identity Verification code
- User name and password

Members of the Public

- Contact information (*e.g.*, name, address, phone number, email address)
- Social Security numbers

[15] For example, Bates Numbers may be created. These numbers are used in the legal, medical, and business fields to place identifying numbers and date/time-marks on documents as they are scanned or processed during the discovery stage of preparations for trial or identifying business receipts.

[16] The data elements listed in the document are provided for transparency, as they are generally the information involved in litigation incidents. However, users of the tools cannot completely control what goes in eDiscovery tools because parts of files cannot be excluded; generally, anything can go into an eDiscovery tool.

- Correspondence
- Benefit applications and associated files
- Law enforcement investigation information
- Criminal history
- Travel records
- Medical information
- Other financial records
- Family information

**it is possible that an array of PII/SPII may be contained within the above documents*

DHS Employees/Contractors

- Contact information (*e.g.*, name, address, phone number, email address)
- Emails
- Memos
- Correspondence
- Personnel files of those involved in litigation-related incidents
- Records generated that are related to active DHS litigation including but not limited to reports, analysis, guidance, training materials, charts, or photography

**it is possible that an array of PII/SPII may be contained within the above documents*

Employees of other Federal Agencies

- Contact information (*e.g.*, name, address, phone number, email address)
- Emails
- Memos
- Correspondence
- Records or documents related to active DHS litigation including but not limited to reports analysis, guidance, training materials, charts, or photography

**it is possible that an array of PII/SPII may be contained within the above*

documents

2.2 What are the sources of the information and how is the information collected for the project?

eDiscovery tools may store and process data from any DHS record keeping system. The nature of those records varies and may include benefit application files, law enforcement files, trade, maritime safety, other requests, and associated project worksheets with relevant information, personnel and employment records, and financial records. The source of such records varies depending on the type of activity the record is created to support. Any of these records may also contain metadata, which is typically generated by the source system. Because eDiscovery tools can contain any records that DHS receives, creates, or maintains, it is not possible to list all of the possible sources of information for those records.

DHS personnel originally gather the potentially relevant documents pursuant to direction from OGC/OCC to produce material in anticipation of litigation. Technical support personnel may also search and retrieve electronic data from all locations where responsive ESI might be stored including, central agency databases, agency file servers (*e.g.*, shared drives), and centrally stored agency electronic mail. Once relevant information has been identified and litigation ensues, the data is securely transferred to the tool.

2.3 Does the project use information from commercial sources or publicly available data? If so, explain why and how this information is used.

eDiscovery tools may contain commercial or publicly available data only to the extent that it is already contained in the records loaded into the tools' repository for litigation and responsive document requests. Commercial or publicly available information typically appears in contract documents, benefit documents, and other similar documents. The commercial and publicly available data is merely a category of data that could be included in the records input into these tools for review. Because DHS does use commercially-sourced data and publicly available data in executing its mission, it is possible that such data may be included.

2.4 Discuss how accuracy of the data is ensured.

eDiscovery tools operate under the principle of full and open discovery of whatever information exists in DHS recordkeeping systems. DHS may not alter, withhold, redact, or delete existing documents in the course of litigation discovery except as permitted by the Federal Rules of Civil or Criminal Procedure, federal statutes, and as authorized by the court. Federal discovery rules require the preservation and production of records in DHS systems, notwithstanding the accuracy of those records. The accuracy of the information in the documents themselves depends on their nature and source.

2.5 Privacy Impact Analysis: Related to Characterization of the Information

Privacy Risk: eDiscovery tools may present a risk of the over-collection of PII or the aggregation of disparate PII from separate agency recordkeeping systems.

Mitigation: DHS cannot fully mitigate this risk. DHS only collects and aggregates information in these tools when it is under a legal mandate to respond to discovery and other responsive document requests. The agency does not have discretion to limit the scope of the collection. OGC/OCC mitigates this risk by only using the system to support the review and production of records in litigation and other responsive document requests. OGC/OCC limits information based on role-based access in the specific tool, generally to those DHS attorneys and paralegals assigned to those matters. Additionally, the eDiscovery tools have the capability to generate a robust audit trail of all user activity, including the viewing of records in the system.

Privacy Risk: eDiscovery tools may present a risk that the PII in the system is not accurate, complete, and current.

Mitigation: Due to the nature and use of the system, DHS cannot fully mitigate this risk. DHS does not use the information in eDiscovery tools to make decisions about individuals. The system contains only copies of records from other agency recordkeeping systems, and DHS will not use the tools as an internal source of agency records about individuals. The purpose of eDiscovery tools is to support the mandatory production of agency records in pending litigation. Federal litigation rules require production of documents in their original form, even if they contain erroneous, incomplete, or outdated information. Incorrect information can be corrected in the source system. However, active litigation may prevent the correction of the information because federal litigation rules require production of documents in their original form, which must be maintained throughout the litigation.

Section 3.0 Uses of the Information

3.1 Describe how and why the project uses the information.

DHS uses the information loaded into eDiscovery tools to support the mandatory production of agency records in pending civil or criminal litigation and in response to responsive document requests.

DHS uses the production file generated by the tools to produce releasable portions of records in electronic and searchable form to DOJ to allow it to represent the United States' interests in litigation, to other parties in litigation as required or agreed to in discovery, and to the court. DHS may also use the production file to respond to responsive document requests.

3.2 Does the project use technology to conduct electronic searches, queries, or analyses in an electronic database to discover or locate a predictive pattern or an anomaly? If so, state how DHS plans to use such results.

These eDiscovery tools are not used to conduct predictive pattern or anomaly analysis.

3.3 Are there other components with assigned roles and responsibilities within the system?

Each Component may have its own instance of these tools. However, each tool should only allow those with a need-to-know to have access.

3.4 Privacy Impact Analysis: Related to the Uses of Information

Privacy Risk: There is a privacy risk of unauthorized access to and use of the information maintained in eDiscovery tools.

Mitigation: To mitigate this risk, eDiscovery tools employs appropriate role-based access controls so only authorized OGC/OCC personnel have access to the system and to the individual cases in the system, based on their work assignments. OGC/OCC supervisors decide which OGC/OCC personnel are granted access to records stored under a particular eDiscovery case, what functions those personnel will be able to perform in the system, and which records individual users may view or review. Users will only have access to the cases assigned to them. Additionally, all users receive training regarding the proper use of each specific tool prior to being granted access to the system. All users also complete annual mandatory privacy and security training, which stresses the importance of appropriate and authorized use of personal data in government systems and the penalties for violations.

Privacy Risk: There is a privacy risk that information in eDiscovery tools may be used for purposes beyond litigation.

Mitigation: This risk is mitigated. No agency records are uploaded into eDiscovery tools unless it is in anticipation of litigation. Moreover, these tools are able to maintain detailed audit logs that would capture any user's inappropriate access or viewing of information contained within the tool. These logs may be automatically generated, and are generally reviewed by OGC/OCC and IT security when there is evidence or reason to believe that the integrity of the data has been compromised.

Section 4.0 Notice

4.1 How does the project provide individuals notice prior to the collection of information? If notice is not provided, explain why not.

Because eDiscovery tools are not a primary information collection system, DHS does not provide notice to individuals prior to the tools' collection of information. Litigants in civil cases are aware that courts may compel DHS to search for and produce agency records pertaining to them and their claims during the litigation process. This PIA serves as notice to the general public as to the collection and use of information in these types of tools for the purposes described in this PIA. The DHS General Legal Records SORN[17] provides notice of the records that may be collected by OGC/OCC in anticipation of or related to litigation. DHS's other PIAs and SORNs also provide general notice to the public of the type of records and information DHS collects and maintains generally, which helps provide transparency as to the nature of the agency records which may be collected and loaded into eDiscovery tools.[18]

DHS provides notice at the point of original collection wherever possible; however, in cases in which the data collection supports a law enforcement activity, opportunities for the individual to be notified of the collection of information may be limited or nonexistent. The purpose and context of the original collection of information determines whether notice is provided.

4.2 What opportunities are available for individuals to consent to uses, decline to provide information, or opt out of the project?

Because eDiscovery tools are not a primary information collection system, any right or opportunity to consent or decline to provide information occurs at the point of original collection from the individual and is described in the relevant PIA and SORN for that recordkeeping system, program, or activity.

4.3 Privacy Impact Analysis: Related to Notice

Privacy Risk: There is a risk that individuals are not aware of the existence of eDiscovery tools and the data those tools collect and maintain.

Mitigation: This risk cannot be fully mitigated. This PIA serves as public notice of the existence of eDiscovery tools, the data they collect and maintain, and the limited purposes for which DHS will use the data. However, because these tools support a secondary collection of

[17] DHS/ALL-017 Department of Homeland Security General Legal Records, 76 FR 72428 (November 23, 2011).
[18] A list of DHS PIAs and SORNs may be found here: https://www.dhs.gov/privacy.

information from records already compiled in existing agency recordkeeping systems, individualized notice is not possible or practical.

Section 5.0 Data Retention by the project

5.1 Explain how long and for what reason the information is retained.

DHS retains these records until the final resolution of the case, claim, or action causing the collection of the documents for processing within eDiscovery tools. Litigation case files are retained in accordance with NARA-approved retention schedule for each Component.[19] A general overview of this retention period is outlined below.

For civil litigation, final resolution means an administrative settlement of the claim or case, a dismissal with prejudice of all claims arising from the same subject matter, a final judgment on the case or claim, or the exhaustion of appeals, whichever comes last. In the event litigation was anticipated but never filed, but OCC collected and uploaded records to an eDiscovery tool, DHS will store those records only until 6 months after the expiration of the appropriate statute of limitations,[20] currently two (2) years from the date of injury for common law tort claims, or as long as the state statute of limitations mandates for constitutional claims.

For common law tort claims, the statute of limitations requires an administrative claim to be submitted within two (2) years of the date of loss, after which the agency has six (6) months to adjudicate the claim. For constitutional tort claims, statutes of limitations vary from as little as two (2) years to as long as eight (8) years; the limitations are set by state law, and the laws of the fifty states vary between two (2) and six (6) years from the date of the injury.

For criminal litigation, DHS will retain data until final resolution. Final resolution means the dismissal with prejudice of all related charges, an acquittal of all related charges, or the exhaustion of appeals on all related charges.

All records will be maintained pursuant to the records retention schedules for the source systems' documents, and as dictated by the courts. Due to the storage limits within some eDiscovery tools, the records that are uploaded to and processed will be deleted from the tool after processing and production.

Retention of records gathered for input into eDiscovery tools: (1) Records stored in a secure format prior to being input into these tools will be deleted once it is confirmed that they have been properly uploaded; (2) copies of any hardcopy records received will be scanned and saved as a PDF document. The PDF document will then be saved and uploaded in a secure format. Upon

[19] For example, FEMA follows N1-311-86-1, Item 1F2, Official Litigation Case Files.
[20] A party can file a complaint on the date the statute of limitations expires, but not serve it for several months. Therefore, DHS must maintain files after the statute of limitations expires. This timeline may differ based on Component needs, but generally follows the timeline outlined above.

successful upload, the DHS attorney with the case file will maintain the hardcopy records until final resolution.

System administrators destroy all data in eDiscovery tools in accordance with DHS guidance on the secure destruction of electronic information.

5.2 Privacy Impact Analysis: Related to Retention

Privacy Risk: There is a privacy risk that information will be retained for longer than necessary to accomplish the purpose for which the information was originally collected.

Mitigation: DHS maintains the records in eDiscovery tools according to the records schedules and policies discussed in Section 5.1, and disposes of them accordingly. OGC/OCC certifies when records can be destroyed. However, due to some storage limits within eDiscovery tools, the records that are uploaded to and processed may be deleted from the tools after processing and production. Occasionally, there will be times when records are kept in the system due to ongoing litigation. To ensure that storage limits are not reached, users will receive periodic housekeeping notices to ensure users review any maintained files to delete those no longer needed.

Section 6.0 Information Sharing

6.1 Is information shared outside of DHS as part of the normal agency operations? If so, identify the organization(s) and how the information is accessed and how it is to be used.

Yes. DHS shares information stored and processed in eDiscovery tools with DOJ and any outside federal agency asked to consult on or concur with the disclosure of responsive information during civil or criminal proceedings. DHS does so in accordance with routine uses in the DHS General Legal Records SORN.[21] Moreover, DHS must share any information that is subject to discovery in litigation with DOJ, the court, and opposing counsel to fulfill DHS's obligations and ensure that all parties to the litigation have fair and equal access to the evidence. DHS generally discloses the information in encrypted form via secure email or delivery of the data on portable storage media.

For responsive document requests, DHS may ultimately share the information stored and processed in these tools with the requester to the extent the information is not subject to withholding under an exemption or exception. DHS may share the information with other agencies that own or originated the records or data contained therein, or otherwise have equities in the records or information, to determine whether the records are releasable or exempt. DHS may also

[21] DHS/ALL-017 Department of Homeland Security General Legal Records, 76 FR 72428 (November 23, 2011).

share the information with DOJ in the event that the requester files a lawsuit challenging the adequacy of the agency's response to the request.

For subpoenas and *Touhy* requests in matters to which DHS is not a party, DHS shares the information stored and processed in eDiscovery tools with the requesting party, and any other parties as required by the court under pursuant to 5 U.S.C. § 552a(b)(11). DHS does not share this information without a privacy waiver or an order from a court of competent jurisdiction.

6.2 Describe how the external sharing noted in 6.1 is compatible with the SORN noted in 1.2.

The DHS General Legal Records SORN supports the mission of the DHS Office of the General Counsel and DHS component legal offices to provide the agency with legal services, including supporting the agency during litigation. The external sharing of the records in eDiscovery tools for the litigation-related uses, as well as subpoena and *Touhy* request uses, described above is compatible with Routine Uses A and H in of that SORN.

6.3 Does the project place limitations on re-dissemination?

No. eDiscovery tools are document storage and processing tools only. DHS expects to disclose any records input into these tools during litigation or during the processing of responsive document requests. The re-dissemination of records processed through these tools may not be discretionary for DHS and may be mandated by law. In civil and criminal discovery, DHS discloses information through the DOJ and the courts. Limitations on the re-dissemination of information will generally be those described in the exemptions under open records statutes, the civil discovery privileges, court rules and orders, and agency policies limiting the re-dissemination of law enforcement sensitive information.

6.4 Describe how the project maintains a record of any disclosures outside of the Department.

Audit trails in these tools capture actions associated with the creation of a production file. The audit trail maintains the date and time when a production file was created, as well as the user performing the action. In the event case information is provided to a third party, DHS will save the production file in a format that identifies the third party recipient, case name, and the date the file was created.

6.5 Privacy Impact Analysis: Related to Information Sharing

Privacy Risk: There is a risk that disclosure of information collected in the eDiscovery tools will be incompatible with the original purposes for which the information was collected.

Mitigation: DHS only uses eDiscovery tools to facilitate the Department's production of records as mandated by statute or federal court rules, and responsive document requests. Disclosures of records in litigation to which they are relevant, or as mandated by open records statutes, support the underlying democratic principles of fairness, transparency, and accountability. Any information sharing is done pursuant to Routine Uses A and H in the DHS General Legal Records SORN.

Section 7.0 Redress

7.1 What are the procedures that allow individuals to access their information?

Individuals may request access to records about themselves in eDiscovery tools. All or some of the requested information may be exempt from access, pursuant to the Privacy Act, in order to prevent harm to law enforcement investigations or interests, or if DHS compiled the information in reasonable anticipation of litigation. Providing individual access to records contained in eDiscovery tools could inform the subject of an actual or potential investigation or reveal an investigative interest on the part of DHS. Access to the records could also permit the individual who is the subject of a record to impede the investigation, tamper with witnesses or evidence, and avoid detection or apprehension.

Individuals seeking access to any record contained in this system of records may submit a Privacy Act (for U.S. citizens and Lawful Permanent Residents) or Freedom of Information Act (FOIA) (for all individuals) request to the respective component FOIA Office which can be found under "Contact Information" at https://www.dhs.gov/freedom-information-act-foia.

7.2 What procedures are in place to allow the subject individual to correct inaccurate or erroneous information?

Because of the nature of eDiscovery tools as a repository for records gathered from other DHS recordkeeping systems pursuant to discovery obligations or open records laws, the tools are not designed to allow the individual to correct inaccurate or erroneous information about him or herself. Federal discovery rules require the preservation and production of records in DHS recordkeeping systems, notwithstanding the accuracy of those records. DHS is not permitted to modify those records even if they contain inaccurate or outdated information. Permitting amendment of eDiscovery tool records could interfere with ongoing litigation, investigations, and law enforcement activities.

Because information in these tools is obtained from other DHS recordkeeping systems, individuals are able to request correction of any inaccurate or erroneous information in the source systems themselves, subject to any Privacy Act exemptions intended to prevent harm to law enforcement investigations or interests. Individuals seeking to contest the content of a record may

submit a Privacy Act request in writing to the component's FOIA Officer, whose contact information can be found at http://www.dhs.gov/foia under "Contact Information." If an individual believes more than one component maintains Privacy Act records concerning him or her the individual may submit the request to the Chief Privacy Officer, Department of Homeland Security, 245 Murray Drive, S.W., Building 410, STOP-0550, Washington, D.C. 20528.

7.3 How does the project notify individuals about the procedures for correcting their information?

Because information in these tools is obtained from other DHS recordkeeping systems, individuals are able to request correction of any inaccurate or erroneous information in the source systems themselves. These source systems have different processes and procedures for correcting information, which are outlined in their respective PIAs and SORNs.[22]

7.4 Privacy Impact Analysis: Related to Redress

Privacy Risk: There is a risk that individuals may not have access or the ability to correct their information in eDiscovery tools used by the Department.

Mitigation: This risk cannot be mitigated. While individuals can request to access and correct (if necessary) information about themselves in the system from which their information was originally collected, the nature of eDiscovery tools is such that the ability of individuals to access or correct their information is non-existent.

Section 8.0 Auditing and Accountability

8.1 How does the project ensure that the information is used in accordance with stated practices in this PIA?

DHS can use the audit trail ability of these tools to monitor and track document review functions, as needed. This audit trail can assist investigating officials in identifying unauthorized use of the system so that DHS may take any appropriate follow-up actions. The audit trails track the following types of information:

- Information on when users logged in to and logged out of the system;
- Search terms and the date and time they were executed;
- Exports of metadata, native, and production files;
- Printing of all files;
- Tagging (for example, PII);

[22] A list of DHS PIAs and SORNs can be found here: https://www/dhs.gov/privacy.

- Redactions;
- The user performing the action;
- The identity of DHS employees authorized to access a particular case;
- Any changes to or redactions of data within the tool;
- Any determination that a document is privileged; and
- Any information that is exported.

Designated users, such as system administrators or OGC/OCC supervisors, can access the audit trail. If an OGC/OCC employee were to disclose information from a tool inappropriately, DHS would be able to review this audit trail to determine the potential sources of the unauthorized disclosure and take appropriate corrective action. For litigation matters, the federal courts would be an additional control regarding the unauthorized disclosure, dissemination, or re-dissemination of PII or privileged information. These audit logs are typically generated automatically.

Authorized DHS personnel access eDiscovery tools from their DHS computers, which are encrypted and password-protected and have other security features such as automatic locking of the desktop after 15 minutes of inactivity. In all cases, OGC/OCC managers control who may access data within these tools.

Each Component may have different methods/infrastructure to store data used with eDiscovery tools. Some data may be hosted in the cloud. Each cloud instance would have to complete the same security requirements of DHS policy and the DHS Offices of the Chief Information Security and Chief Information Officer, regardless of eDiscovery tool being used.

8.2 Describe what privacy training is provided to users either generally or specifically relevant to the project.

All DHS employees and contractors complete annual mandatory privacy and security training. Additionally, all users receive on-the-job training regarding the proper use of these tools, which will include privacy information. DHS will also provide instruction manuals for general use and best practices of using the system.

8.3 What procedures are in place to determine which users may access the information and how does the project determine who has access?

OGC/OCC supervisory attorneys in charge of teams and divisions responsible for litigation support assign attorneys and paralegals to individual cases. The assigned attorneys and paralegals are responsible for assuring a complete and diligent discovery search, preservation, collection, and production of relevant records. They will have access to records gathered and inputted into a tool

for a particular litigation matter, along with the OGC/OCC and Office of the Chief Information Officer personnel who serve as system administrators for the tool.

There is no standardized process for revoking access. Those who have received access will not have their access revoked after a case is completed because of the likelihood of future litigation matters. Users will have access revoked in rare circumstances, such as when an employee transfers to a different department or leaves DHS. While users retain access to the tools, they will only be able to access the cases assigned to them.

There are typically three user roles within these types of tools: system administrator, super user, and end user.

(1) *System administrators* have full privileges to perform all functions in the tool. System administrators can create user groups and grant customized levels of access and privileges to these groups and the users within them. Certain functions are reserved for the system administrator, such as the ability to load and process ESI into an existing case, and to generate the production version of records in the system. System administrators can also assign users to any user role or group.

(2) *Super users*, by default, have more restricted privileges than system administrators. Super users can assign other super users and end users to user groups or assign them levels of access and privileges for particular cases. Super users may also access system audit trails. Super users may not perform certain functions that are reserved for system administrators, such as loading records into the tool. Supervisors will often be assigned the role of super user.

(3) *End users* have the most limited privileges in the tools. End users may only access and take actions on those cases and/or records based on the levels of access and privileges they are granted and groups to which they are assigned by a system administrator or a super user. Attorneys and paralegals will usually be assigned the role of end user.

8.4 How does the project review and approve information sharing agreements, MOUs, new uses of the information, new access to the system by organizations within DHS and outside?

Typically, these tools do not share information directly from the tool, nor is it envisioned that the expansion of the users of the tools or the intended uses of the information collected and maintained in the tools would require an information sharing agreement. In the event it considers such changes, OGC/OCC would engage the DHS Privacy Office and the respective Component Privacy Office to discuss the intended expanded users and/or uses of this information, and update the relevant privacy compliance documentation (including this PIA) as appropriate.

Responsible Officials

David Palmer
Chief of Staff
Office of the General Counsel
Department of Homeland Security

Approval Signature

Original, signed copy on file with the DHS Privacy Office.

_____

Jonathan R. Cantor
Acting Chief Privacy Officer
Department of Homeland Security

Appendix A

eDiscovery Tools Deployed at DHS

1. Relativity
2. Clearwell
3. EnCase
4. EverLaw