The Business Court of Texas,
1st Division

| | | |
|---|---|---|
| XXXX, *Plaintiff*, | § | |
| | § | |
| v. | § | Cause No. XX-BC01B-XXX |
| | § | |
| YYYY, *Defendant* | § | |
| | § | |
| | § | |
| | § | |

---

**PROTOCOL FOR THE PRODUCTION OF ELECTRONICALLY STORED INFORMATION (ESI)**

---

Pursuant to Texas Rules of Civil Procedure (including particularly Tex. R. Civ. P. 196.4) and Texas Business Court Local Rules, this order governs the production of Electronically Stored Information (the "Protocol") by the above-captioned Parties as follows:

**A. Definitions and Scope**. The following protocol and definitions shall control the production of discoverable documents and electronically stored information when production commences. The following terms shall be defined:

1. "Document(s)" is defined to be synonymous in meaning and equal in scope to the usage of the terms "all responsive documents, electronically stored information, and tangible things" in Rule 194.1 of the Texas Rules of Civil Procedure and includes ESI existing in any medium from which information can be translated into reasonably usable form, including but

not limited to email and attachments, word processing documents, spreadsheets, graphics, presentations, images, text files, databases, instant messages, transaction logs, audio and video files, voicemail, internet data, computer logs, text messages, and backup materials. The term "Document(s)" shall include Hard Copy Documents, Electronic Documents, and Electronically Stored Information (ESI) as defined herein.

2. "Electronic Document(s) or Data" means Documents or Data existing in electronic form at the time of collection, including but not limited to: e-mail or other electronic communications, word processing files (e.g., Microsoft Word), computer presentations (e.g., PowerPoint slides), spreadsheets (e.g., Excel), and image files (e.g., PDF).

3. "Electronically stored information" or "ESI," is information that is stored electronically as files, documents, or other data on computers, servers, mobile devices, online repositories, disks, USB drives, tapes or other real or virtualized devices or digital media storage.

4. "Extracted Text" is a file or files containing text extracted from a Document, and includes all available information including the header, footer, and document body information.

5. "Hard Copy Document(s)" means Documents existing in paper form at the time of collection.

6. "Hash Value" is a numerical identifier that can be determined from a file, a group of files, or a portion of a file, based on a standard mathematical algorithm that calculates a value for a given set of data, serving as a digital fingerprint, and representing the binary content of the data to assist in subsequently ensuring that data has not been modified and to facilitate duplicate identification. Unless otherwise specified, hash values shall be calculated using the MD5 hash algorithm or latest SHA-1 algorithm.

7. "Instant message" is any real-time message, post or other communication transmitted over the internet or another computer

network via a social messenger application, messaging application or chat application (e.g., Discord, Teams, Slack, Facebook Messenger, WhatsApp).

8. "Load File(s)" are electronic files containing information identifying a set of paper scanned (static) images or processed ESI and indicating where individual pages or files belong together as documents, including attachments, and where each document begins and ends. Load Files also contain data relevant to individual Documents, including extracted and user-created Metadata, coded data, OCR, and Extracted Text. A load file linking corresponding images is used for productions of static images (e.g., TIFFs).

9. "Metadata" is the term used to describe the structural information of a file that contains data about the file, as opposed to describing the content of a file.

10. "Native Format" means the file format associated with the original creating application and as collected from custodians. For example, the native format of an Excel workbook is an .xls or .xlsx file, and the native format of an image can be a .png, .gif, or .jpg.

11. "Optical Character Recognition" or "OCR" means "optical character recognition," which is a technology process that captures text from an image for the purpose of creating an ancillary text file that can be associated with the image and searched in a database. OCR software evaluates scanned data for shapes it recognizes as letters or numerals.

**B. Production of ESI**.

1. Search methodology. Parties shall try to reach agreement on appropriate search terms before production. Upon request, the producing party shall disclose what search terms, if any, were used to locate ESI likely to contain discoverable information. If search terms were not used, a party shall disclose the search methodology used to locate ESI likely to contain discoverable information.

2. Processing.

a) Parties shall normalize times and dates to conform to United States Central Time (CT).

b) For archive files (zip, jar, rar, gzip, TAR, etc.), all contents should be extracted from the archive with source pathing and family relationships preserved and produced. The fully unpacked archive container file does not need to be included in the production.

3. Production Format.

a) *Load Files.*

- Productions will, as applicable, include Concordance delimited file (.dat) format data. All metadata will be produced in ASCII, UTF-8, or UTF-16.

- All native format files shall be produced in a folder named "NATIVE."

- All TIFF images shall be produced in a folder named "IMAGE," which shall contain sub-folders named "0001," "0002," etc. Each sub-folder shall contain no more than 10,000 images. Images from a single document shall not span multiple sub-folders.

- All Extracted Text and OCR files shall be produced in a folder named "TEXT."

- All load files shall be produced in a folder named "DATA" or at the root directory of the production media.

b) *Images.*

- The Parties will produce Electronic Documents, Data and ESI as single page Group IV TIFF images, 300 dpi quality or better, and Letter or Legal page size, unless identified for Native production only.

- Each TIFF image should have a unique file name corresponding to the Bates number of that page with a ".tif" file extension. The file name should not contain any blank spaces and should be zero-padded (e.g., DEF-000001), taking into consideration the estimated number of pages to be produced. If a Bates number or set of Bates numbers is skipped in a production, the producing party will so note in a cover letter accompanying the production. Bates numbers will be unique across the entire production and prefixes will be consistent across all documents produced.

- All images of documents which contain tracked changes such as comments, deletions, and revision marks (including the identity of the person making the deletion or revision and the date and time thereof), speaker notes, or other user-entered data that the source application can display to the user will be processed such that all that data is visible in the image and in color.

- Where .TIFF images are illegible due to color content (such as colored text on a colored background) or where color is material to the interpretation of a document, color JPG image files shall be provided.

c) *Slipsheets.*

- The Parties will provide TIFF image slipsheets with Bates number and Confidentiality designation for the following documents with the following slipsheet language:

- Documents withheld as privileged within a produced family should be indicated on the slipsheet as "THIS DOCUMENT WITHHELD AS PRIVILEGED."

- Documents produced in Native format should be indicated on the slipsheet as "THIS DOCUMENT PRODUCED NATIVELY."

- Documents that will not render due to corrupt file or other reason should be so indicated on the slipsheet as "DOCUMENT WOULD NOT RENDER."

- Each slipsheet shall contain only one Bates number per document, regardless of the number of pages in the privileged, native, or non-rendered document.

d) *Natives.*

- The Parties will produce the following forms of ESI in native formats:

    - Spreadsheets.

    - PowerPoint presentations.

    - Image files.

    - Access databases.

    - Audio and video files.

    - Documents of a type which cannot be reasonably converted to useful TIFF images.

- For all forms produced in Native format, the Parties will provide a corresponding slipsheet, text file, and all metadata.

- Each Native file should have a unique file name corresponding to the Bates number of the slipsheet or first TIFF image page for that document, appended with the Confidentiality designation. The file name should not contain any blank spaces and should be zero-padded (e.g., DEF-000001), taking into consideration the estimated number of pages to be produced.

e) *Text.* For each document, a single Unicode text file containing Extracted Text shall be provided along with the image files and metadata. The text file name shall be the same as the Bates number of the first page of the document. File names shall not have any special characters or embedded spaces. Electronic text must be extracted directly from the native electronic file to the extent reasonably feasible unless the document is an image file or contains redactions, in which case, a text file created using OCR should be produced in lieu of Extracted Text.

f) *Hard-Copy Documents.*

- Hard copy documents shall be scanned to single page Group IV TIFF format, 300 dpi quality or better with corresponding searchable OCR text. Production format of these documents shall be the same as Electronic Documents and Data.

- The Parties shall use reasonable efforts to ensure distinct documents are not merged into a single record and single documents are not split into multiple records (i.e., paper documents should be logically unitized). For example, Hard Copy Documents stored in a binder, folder, or similar container should be produced in the same order as they appear in the container. The front cover of the container should be produced immediately before the first document in the container. The back cover of the container should be produced immediately after the last document in the container. The Parties will undertake reasonable efforts to,

or have their vendors, logically unitize documents correctly, and will commit to address situations of improperly unitized documents.

- The Parties shall use reasonable efforts to index documents in order to provide Custodian and Source data as well as Filename or Title for each unitized document.

**C. Production Method**. Production media shall always be encrypted and will be sent via FTP or SFTP link, unless the parties agree otherwise. On the occasion in which a particular production is of a size that would make sending it via FTP link impractical, the parties may agree to send encrypted physical media such as a Hard Drive or USB. Production letters will always accompany productions including the name of the matter in which it was produced, the production date, the Bates number range of the material contained in the production, and a short description of its contents. Passwords for encrypted media will be sent separately from the media itself.

**D. Resolution of Production Issues**. Documents that cannot be read because of imaging or formatting problems (including encrypted, corrupt and/or protected files identified during the processing of ESI) shall be promptly identified by the receiving party. The producing party and the receiving party shall meet and confer to attempt to resolve any problems. The Parties will use reasonable efforts and standard industry practices to address and resolve the problem(s).

**E. Document Unitization**. When scanning paper documents into Document Images, they shall be unitized in a manner so as to maintain the document(s) and any attachments, as they existed in their original state.

**F. Document Families**.

1. The Parties agree that if any part of a Document or its attachments is responsive, the entire Document and attachments will be produced or a slipsheet provided consistent with the instructions in the Protocol.

2. For Electronic Documents, the relationship of documents in a document collection (e.g., cover letter and enclosures, e-mail and attachments, binder containing multiple documents, or other documents where a parent-child relationship exists between the documents) shall be maintained using the Begin Family and End Family fields of the load file, provided however that the parties shall only be required to present one level of parent-child relationship. Document Images generated from attachments to emails stored in Native Format shall be produced contemporaneously and sequentially immediately after the parent email in their Bates numbering.

**G. Duplicates**.

1. A producing party who has more than one identical copy of an Electronic Document (i.e., the documents are actual duplicates) need only produce a single copy of that document. A producing party need not produce the same electronically stored information in more than one form.

2. Deduplication. The producing party shall de-duplicate documents based on MD5 or SHA-1 hash values at the family level or by Message ID and other standard methodology for email within the collection of a custodian or a data source. Attachments to parent documents may not be deduplicated against a duplicate standalone version of the attachment, and standalone versions of documents may not be suppressed if a duplicate version exists as an attachment.

**H. Email Threading**. To reduce the volume of entirely duplicative content within email threads, the Parties may, but are not required to, use email threading. A party may use industry standard message threading technology to remove email messages where the content of those messages, and any attachments, are wholly contained within a later email message in the thread; provided, however, that the use of threading must not serve to obscure whether a recipient received an attachment or whether a sender sent an attachment.

**I. Bates Numbering**. Each producing party shall Bates number its production(s) as follows:

1. *Document Images*. Each page of a produced Document shall have a legible, unique page identifier ("Bates Number") electronically "burned" onto the image at a location that does not unreasonably obscure any information from the source document. The Bates Numbers shall be enumerated as defined above in Definitions. The producing party will use a consistent prefix throughout the matter. No legend or stamp will be placed on the Document Image other than a confidentiality legend (where applicable), redactions, and the Bates Number identified above. The confidentiality legend shall be "burned" onto each document's image at a location that does not unreasonably obscure any information from the source document.

2. *Sort Order.* For Bates numbering, documents will be sorted by their original file path in ascending order, preserving family ordering.

**J. Altering a Document Production**. In the event that a party alters an already-produced document production (e.g. producing an unredacted copy of a previously redacted document or clawing back a document previously produced), the producing part shall provide an overlay load file for each such altered document. If the producing party lacks the technical ability to produce an overlay load file, the producing party shall reimburse the receiving party for any additional cost associated with a manual alteration of already produced documents.

**K. Metadata**. Documents withheld or produced redacted for privilege may be produced with metadata scrubbed to the extent necessary to protect the information in the document over which privilege is asserted. Otherwise, the producing party shall produce the metadata information described below with each production and in the format described above. For each Document, the producing party shall produce a line in the index file with the following fields, where available. The field naming conventions shall be the following. Datetime metadata will be provided in Central Time in a consistent, reasonable, and clearly delimited format. For example, MM/DD/YYYY HH:mm. To the extent timezone information is provided for context, it will be provided in a separate loadfile field.

| Field Name | Description |
|---|---|
| BEGBATES | Beginning Bates number of first page of a document. |
| ENDBATES | Ending Bates number of last page of a document. |
| BEGATTACH | First Bates number of attachment range, including range of instant messaging conversation. |
| ENDATTACH | Last bates number of the last page of an attachment range. |
| ATTACHCOUNT | Number of attachments to an e-mail. |
| CUSTODIAN | Individual from whom the document originated/was collected. |
| DUPLICATE CUSTODIAN | Custodians whose file/message has been de-duplicated including the original Custodian; separated by semicolons. |
| CONFIDENTIAL | Export as field the coded designation for each document. |
| REDACTED | Export as field for all documents, whether redacted or not. |
| FAMILY DATE TIME | Date and time parent document was sent or last modified. |
| SENT DATE TIME | Date and time email or instant message was sent. |
| CREATE DATE TIME | Date and time document, including instant message, was created. |
| LAST MOD DATE TIME | Date and time document, including instant message, was last modified or edited. |
| LAST MOD BY | Last individual to modify the document, including instant message. |
| AUTHOR | Author of the document. |
| FROM | Email or instant message Sender. |
| TO | Email or instant message “To” recipient(s) or Recipient(s) of an instant message, separated by semi-colons. |
| CC | Email or instant message “Carbon Copy” recipient(s), separated by semi-colons. |
| BC | Email or instant message “Blind Carbon Copy” recipient(s), separated by semi-colons. |
| PARTICIPANTS | User name(s) or ID(s) for participants in an instant message thread. |
| SUBJECT | Subject line of the email. |

| | |
|---|---|
| FILENAME | Original name of file as appeared in location where collected. |
| CONVERSATION TOPIC | Topic or subject of an instant message conversation. |
| THREAD ID | Thread identifier for an instant message. |
| CONVERSATION ID | Conversation identifier for an instant message. |
| FILE EXTENSION | The file type extension of the document. |
| FILEPATH | Path where native file document, including email, was stored when collected, including original file name. |
| GROUP NAME | Name of channel, group or meeting of an instant message. |
| WORKSPACE NAME | Name of workspace or community of an instant message. |
| DEDUPED FILEPATH | Full filepath of deduplicated instances, separated by semicolons. |
| FILESIZE | Size of native file document/email in kilo/mega bytes. |
| PAGE COUNT | Number of pages. |
| HIDDEN CONTENT | Denotes presence of Tracked Changes / Hidden Content / Embedded Objects in item(s). |
| NATIVELINK | Relative path and filename for native file on production media. |
| TEXTLINK | Relative path and filename for text file on production media. |
| MSGID | E-mail or instant message Unique Message ID. |
| HASH | MD5 or SHA-1 # Hash value of the item. |

It is so ORDERED.

BILL WHITEHILL
Judge of the Texas Business Court,
First Division

DATE: [date]