



3. Documents originating from hard copy sources (“Hard-Copy Documents”) and attachments are to be produced in Group IV TIFF format (black and white, 300 dpi) with corresponding searchable OCR text, along with the below-listed metadata fields when available, to be provided in a standardized load file compatible with Concordance or Summation (load file specifications can be found in Appendix I *infra*), with a bates-number field included on the load file so that text and metadata can be matched with TIFF images.

- a. Beginning Document Bates Number
- b. Ending Document Bates Number
- c. Attachment Beginning Bates Number
- d. Attachment Ending Bates Number
- e. Custodian or Source
- f. Confidential (Y/N)
- g. Page Count
- h. Redaction (Y/N)
- i. Document Type – (described in Paragraphs 3 - ESI)

4. The unitization and scanning protocol to be utilized can be found in Appendix II *infra*. To the extent that either party decides to undertake objective coding in addition to the fields above, that party shall provide notice of such coding to the opposite party.

5. Perrigo may redact (1) information that is privileged or protected from discovery as work product or by reason of any other applicable privilege or immunity and (2) information subject to non-disclosure obligations imposed by governmental authorities, law, or regulation (*e.g.* protected personal information). Redacted documents will be produced in TIFF format with

corresponding searchable OCR text and the associated metadata for the document, ensuring the redacted content is fully protected from disclosure.

6. Perrigo maintains some documents in proprietary databases. Perrigo states that, because of the proprietary nature of these databases, documents cannot be produced in native form, because they would not be usable by any party that does not have access rights to the software used for these databases. Plaintiff states that she will not accept production of responsive documents from proprietary databases in image format. Perrigo and Plaintiff agree to meet and confer at a later date regarding the form of production for responsive documents in proprietary databases. If Perrigo and Plaintiff are unable to agree upon the form of production for responsive documents in proprietary databases, they shall submit the issue to the Court for determination.

7. Documents originating from electronically stored information (“ESI”), except those documents as noted in Paragraphs 5 and 6, are to be produced in TIFF format with extracted text directly from the electronic file, along with the below-listed metadata fields when available, to be provided in a standardized load file compatible with Concordance or Summation (load file specifications can be found in Appendix I infra), with a bates number field included on the load file so that text and metadata can be matched with TIFF images.

- a. Beginning Document Bates Number
- b. Ending Document Bates Number
- c. Attachment Beginning Bates Number
- d. Attachment Ending Bates Number
- e. Custodian or Source
- f. Confidential (Y/N)

- g. Page Count
- h. Redaction (Y/N)
- i. Document Date (the document's "date last modified" prior to collection, in the format YYYYMMDD, for email this will be the sent date)
- j. File Name (including extension)
- k. From
- l. To
- m. CC
- n. BCC
- o. Email Subject
- p. Title (when available in native metadata for non-email)
- q. Document Type - Descriptor for the type of document: "EFILE" for electronic documents not attached to emails; "EMAIL" for all emails; "ATTACHMENT" for files that were attachments to emails; and "HARDCOPY" for hard copy physical documents that have been scanned and converted to an electronic image.
- r. DocExt - The file extension of the document is defined as the substring of the file name which follows but does not include the last occurrence of the dot character.
- s. NativeFile – represents that this file is produced in its native file format (Flag – Y/N). Applicable for EFILES and ATTACHMENTS.
- t. Importance – Priority for E-mail – (Flag – Y/N for emails sent with a "High Importance" or similar designation); the parties understand that this

information was not captured in prior ESI sweeps, but it will be captured in prospective ESI sweeps.

- u. MD5Hash - Checksum for a file, a 128-bit value (SHA-1 value is also acceptable).
- v. Replacement - Descriptor for documents that are replacements for previously-produced documents. Will populate "Y" for replacement documents only.
- w. SourceFilePath - The directory structure of the original file(s). If a file is inside of a container, the container name is included in the path. This includes zip files and email message directories, such as: Mailbox – Smith, Joe\Inbox\ClientMaterials\CrivellaWest.
- x. Attachment Count - The total number of attachments to a parent document, including any attachments that were not processed and the contents of additional attached containers. A value of zero (0) should be returned for any files/documents without attachments.
- y. Duplicate Custodian – In the event that a document is removed during the deduplication process as an exact duplicate by MD5Hash, a duplicate custodian field will be populated with the identify of each custodian that possessed the file that was removed by deduplication (with separate values appropriately delimited by a semi-colon).
- z. Duplicate Custodian Filepath – In the event that a document is removed during the deduplication process as an exact duplicate by MD5Hash, the information for SourceFilePath (item “w” above) for each duplicate file

that was removed by deduplication shall be set forth in this field (with separate values appropriately delimited by a semicolon).

8. To the extent additional metadata fields have been captured in prior ESI sweeps or are captured prospectively, the parties shall meet and confer, and reasonably cooperate, concerning the production of any additional requested metadata. ESI files with compression file types (*i.e.*, .CAB, .GZ, .TAR, .Z, .ZIP) shall be decompressed in a reiterative manner to ensure that a zip within a zip is decompressed into the lowest possible compression resulting in individual folders and/or files. The parties may redact (1) information that is privileged or protected from discovery as work product or by reason of any other applicable privilege or immunity and (2) information subject to non-disclosure obligations imposed by governmental authorities; law or regulation (e.g. protected personal information). Redacted documents will be produced in TIFF format with corresponding searchable OCR text and the associated metadata for the document, ensuring the redacted content is fully protected from disclosure.

9. Microsoft Excel Spreadsheets will be produced as native files unless redactions are required. Spreadsheets that require redactions will be converted to TIFF images as follows: remove user-defined print areas; unhide and expand all columns, rows and sheets; expand/outline groupings; print to TIFF each sheet across (left to right) and then down; set for landscape orientation; and remove blank pages as possible. Microsoft PowerPoint presentations or slide shows will be produced as native files unless redactions are required. PowerPoint or slide shows that require redactions will be converted to color TIFF images as follows: the conversion to single page color TIFF image (compression set to high quality JPEG, with 300 dpi

resolution) will be done utilizing PowerPoint's print feature; background will be turned off so all text is visible; hidden slides will be revealed; and if any slides in the presentation contain speaker notes, then all slides will be printed in Slide view (one slide per page). Any autodate macros within any electronic documents will be indicated as "<autodate>".

10. For documents and ESI that do not convert well to TIFF (e.g. oversized drawings, picture files, audio and video files), the producing party will either produce the document in native format or will ask the receiving party to meet and confer regarding a reasonable alternative form of production. The file name for the documents produced in native form will consist of a Bates number and a confidentiality designation if applicable. For native files included in a production, the corresponding placeholder TIFF image will also be provided.

11. The parties shall meet and confer concerning the scope and production format for discoverable information contained in databases and other structured data sources. Such production shall be governed by separate stipulation of the parties or Order of the Court.<sup>1</sup>

12. De-duplication of ESI will be performed within and across custodians according to MD5 hash values. Upon request, an explanation of how these hash values are being calculated shall be provided. Only a single-copy of exact duplicate ESI need to be produced. Upon reasonable request by a receiving party, the producing party shall make a reasonable effort to produce copies of any document that was removed from a particular custodian's production during the de-duplication process.

13. The parties will meet and confer regarding appropriate key words that may be used to reduce the volume of documents to be reviewed for production. Productions may be

---

<sup>1</sup> This paragraph does not apply to documents maintained in proprietary databases. Documents maintained in proprietary databases are governed by paragraph 6, *supra*.

prioritized by particular custodians, with the party seeking the production having the opportunity to set the priority of certain custodians over others.

14. Parent-child relationships (the association between an attachment and its parent document) shall be preserved to the extent possible. Regarding produced attachments, a Beginning Attachment and Ending Attachment Bates number will be provided for each attachment and included in the data load file, as identified in paragraphs 1c, 1d, 3c, and 3d above.

15. Individual pages of TIFF documents and documents produced in native format shall be assigned a Bates number for ease of reference and to preserve the integrity of the documents and avoid modification of any documents or information. Bates-numbers shall be unique across the entire document production and sequential within a given document.

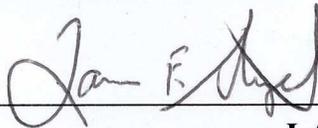
16. The parties understand that this protocol contemplates the production of large volumes of documents, and they acknowledge that nothing in this Order waives, restricts or eliminates the parties' "claw-back" rights pursuant to the Protective Order(s) in this case, or governing law, rules, orders, or agreements regarding inadvertently produced documents.

17. This order concerns the production of ESI from sources that are reasonably accessible. The production of ESI from sources that are not reasonably accessible has not been addressed among the parties. To the extent necessary or appropriate, the production of ESI from sources that are not reasonably accessible will be the subject of subsequent stipulation or Order.

18. The parties agree that documents produced according to this protocol are in a reasonably usable form.

19. The parties shall meet and confer and endeavor to resolve any disputes arising hereunder, before submitting such disputes to the Court for determination.

**IT IS SO ORDERED** this 17<sup>th</sup> day of June, 2013.



---

**LAWRENCE F. STENGEL, J.**

## **APPENDIX I: FILE FORMATS**

### **A. Image Load Files**

- Every document referenced in a production image load file shall have all corresponding images, text, and data logically grouped together.
- Documents shall be produced in only one image load file throughout the productions, unless that document is noted as being a replacement document in the Replacement field of the data load file.
- The name of the image load file shall mirror the name of the delivery volume, and should have an .lfp, .opt or .dii extension (*e.g.*, ABC001.lfp). (If .dii file is produced, the accompanying metadata load file shall be separate from the .dii file and not contained within the .dii file.)
- The volume names shall be consecutive (*i.e.*, ABC001, ABC002, *et. seq.*).
- The load file shall contain one row per Tiff image.
- Every image in the delivery volume shall be contained in the image load file.
- The image key shall be named the same as the Bates number of the page.
- Load files shall **not** span across media (*e.g.*, CDs, DVDs, Hard Drives, etc.), *i.e.*, a separate volume shall be created for each piece of media delivered.

### **B. Metadata Load Files**

- The metadata load file shall use the following delimiters:
  - Columns (ASCII: 20)
  - Quote ª (ASCII: 254)
- Data for documents shall be produced in only one data load file throughout the productions, unless that document is noted as being a replacement document in the Replacement field of the data load file.
- The first record shall contain the field names in the order of the data set forth in Appendix 2.

- All date and time fields are populated in the mm/dd/yyyy and hh:mm AM/PM formats without leading zeros. For example, the month of April would be a “4” rather than a “04”.
- A carriage-return line-feed shall be used to indicate the start of the next record.
- Load files shall *not* span across media (*e.g.*, CDs, DVDs, Hard Drives, etc.); a separate volume shall be created for each piece of media delivered.
- The name of the metadata load file shall mirror the name of the delivery volume, and shall have a .dat, .csv or .txt extension (*i.e.*, ABC001.dat).
- The volume names shall be consecutive (*i.e.*, ABC001, ABC002, et. seq.).

## **APPENDIX II: SCANNING AND UNITIZATION**

- Original document orientation shall be maintained (i.e., portrait to portrait and landscape to landscape). When subjecting physical documents to an OCR process, the settings of the OCR software shall maximize text quality over process speed. Settings for "Auto-Skewing", "Auto-Rotation" and "Zone Location" should be turned on when documents are run through the process.
- Physical document boundary determination will be based on the smallest physical binding element. The following binding elements will be used to determine document boundaries:
  - Staples, Clips, Rubber Bands, Binder Clips – These binding elements at the smallest level will be used to determine the document boundary.
  - Consecutive Single/Loose Pages – If a series of loose pages exist, each loose page will be captured as an individual document
  - Bound Items – All bound items will be captured as one document.
  - Binders with No Tabs – Contents inside the binder will be captured as one document.
  - Binders with Tabs – Each tabbed section, along with the tab, will be captured as a separate document.
  - Computer Forms – Non-perforated computer paper will be captured as one document.
  - Post-It Notes – Post-It Notes will be captured as a separate image and will be placed at the end of the respective document.

- Physical attachment associations are determined by common physical barriers. Where there are nested common physical barriers, the inner-most barrier will be captured as the attachment association. The following physical barriers will be used to determine the physical attachment association:
  - Manila File Folder
  - Parent Clip
  - Hanging File Folder
  - Redweld File
  - Binder
  - Parent Rubber Band