

Production Request Standard

- 1) Produce electronic documents in their complete native/original format incorporating names as the beginning Bates number and as multi-page searchable PDF endorsed with a Bates number on each page.
 - a. If images for a native can't be created, that native should have a corresponding PDF placeholder endorsed with the Bates number and the native file must be produced unless it or its children are redacted, or it is an archive, other container, virus, or similar exception.
- 2) Scan and process all paper into images endorsed with Bates numbering per page as multi-page searchable PDF files.
- 3) Produce fully searchable text for every produced document (extracted text for native documents is preferred; otherwise, OCR data from an image of the document). Provide a blank text file for documents without text.
- 4) Produce only a single copy of a document family across all productions using an accepted de-duplication method for emails and loose files. Provide a list of custodians, paths, and other information for documents which had a copy of a duplicate document. This information may need to be updated for previously produced documents with subsequent duplications. Further explanation and documentation of the de-duplication method may be necessary.
- 5) Produce DAT, OPT, and LST load files as described below.

DAT Load File: All produced documents, regardless of their original file format, must be produced with the metadata fields described below in a UTF-8 data file (DAT). The first line of the DAT file must be a header row identifying the field names.

- The DAT file must use the following default delimiters: Comma ASCII character (020); Quote ¸ ASCII character (254); Newline® ASCII character (174).
- Date, time, and datetime fields should be in the format mm/dd/yyyy and/or hh:mm:ss AM/PM.
- All child documents should sequentially follow the parent document/email.
- Paths in the DAT file for produced native documents should be pathed relative to location of the DAT file and not absolutely pathed in any way.

LST Load File: All produced documents should have a corresponding text file describing the text of the entire document.

- Text must be provided for every produced document. Extracted text for native documents is preferred. If extracted text is not available, not complete, or the document is redacted, provide OCR text from an image of the document. Provide a blank text file for documents without text. For redacted documents produce OCR text of the redacted image and do not provide its extracted text.
- Document text should be saved as UTF-8 text files with the files named the same as the beginning Bates number (begdoc field). Document text should NEVER be provided in the DAT load file.
- These text files should be placed in a separate folder from images or native documents with a cross-reference file (LST) to link the text to the documents in the DAT file. The LST is a comma-delimited file with two fields per line. It should be pathed relative to location of the LST file and not absolutely pathed in any way. There must be a line in an LST file for every document produced.

OPT/LOG Load File: All produced documents should have a single corresponding PDF file which is either an image of the document or a placeholder.

- Native files should be imaged at 300 dpi and reduced to optimal color palette, preserving color when necessary to interpret the document.
- PDF file names should be named by their beginning Bates number.
- Bates page identifiers should be endorsed on the lower right corner of all pages and designations should be in the footer where they do not obscure the Bates number.
- Endorsements, Bates numbers, and/or designations should be placed in headers/footers added to the image and not cover any content of the document image created from the native.
- For documents partially redacted on the basis of any privilege, ensure the redaction box is black and is clearly labeled with white REDACTED lettering and the redaction is "burned in" and not removable. In addition, the document metadata in the DAT should indicate the redaction.
- These PDF files should be placed in a separate folder from text or native documents with an image cross-reference file (OPT or LOG) to link the PDF files to the documents in the DAT file. It is a comma-delimited file consisting of seven fields per line. It should be pathed relative to location of the OPT/LOG file and not absolutely pathed in any way. There must be a line in an OPT/LOG file for every document produced.

Field Order	Standard Field Name	Description only on [Loadfile Key] can be chosen	Data Type	Max Size	Example
1	BegDoc	[Loadfile Key] Number of the first page of the document	Text	30	Prod10003442
2	EndDoc	Number of the last page of the document	Text	30	Prod10003448
3	BegAttach	Displays BegDoc of parent document	Text	30	Prod10003442
4	EndAttach	EndDoc of last attached document in family	Text	30	Prod10003449
5	Batesrng	Range of Bates numbers that make up the document.	Text	100	Prod10003442 - Prod10003448
15	AttachRange	Range of Bates numbers that make up the document family	Text	100	Prod10003442 - Prod10003449
6	ParentID	The BegDoc value of the document's parent	Text	30	Prod10003442
7	NumAttach	Total number of documents attached to the document	Text	5	4
8	CustodianList	List of Custodians which have the document and provide each custodian it deduped from	Text	Full	Doe, John; Smith, Joe;
9	ItemCustodian	Item Custodian	Text	300	Doe, Jane
10	NativeFile	[Loadfile Key] Relative path to the native file	Text	300	Natives\Prod10003442.ext
11	Filename	Filename of document or email from a container may have subject as filename	Text	300	sample.doc
12	Path	Filepath on the original media	Text	1000	\OriginalFilePath\Folder\sample.doc
13	FileDesc	Description of the document's file type	Text	100	MS Outlook Message, or MS Word for Windows Document (OLE), ...
14	DocExt	Documents file extension	Text	20	msg or doc or xls or etc.
15	AttachIDs	Semi-colons delimited list of BegDoc values of all attached documents	Text	Full	Prod10003442, Prod10003447, Prod10003449
16	AttachNames	Semi-colons delimited list of original filenames of all attached documents	Text	Full	Attach1.doc; Attach2.xls
17	From	Person indicated in the document to send the e-mail message	Text	300	someone@company.com
18	EMailSubject	Subject of the e-mail message	Text	Full	This is a sample subject
19	SentDateTime	Sent date and time of an e-mail message	DateTime		1/1/2001 12:01:02 PM
20	SentDate	Sent date of an e-mail message	Date		01/01/2001
21	SentTime	Sent time of an e-mail message	Text	20	12:01:02 PM
22	To	Person(s) indicated in the document to receive the e-mail message	Text	Full	someone@company.com
23	CC	Recipient(s) of "Carbon Copies" of the e-mail message	Text	Full	someone@company.com
24	BCC	Recipient(s) of "Blind Carbon Copies" of the e-mail message	Text	Full	someone@company.com
25	RecievedDateTime	Received date and time of an e-mail message	DateTime		1/1/2001 12:01:02 PM
26	ReceivedDate	Recieved date of an e-mail message	Date		01/01/2001
27	ReceivedTime	Recieved time of an e-mail message	Text	20	12:01:02 PM
28	Filesize	Size of native file, in bytes	Text	50	8987414
29	Author	Author value pulled from native metadata	Text	300	John Doe
30	Organization	Company extracted from native metadata	Text	300	Company X
31	Title	Title value extracted from the native metadata	Text	300	Title of document

Field Order	Standard Field Name	Description only on [Loadfile Key] can be chosen	Data Type	Max Size	Example
32	LastAuthor	Last Saved By value pulled from native metadata	Text	300	Jane Doe
33	Template	Template value pulled from native metadata	Text	300	Normal.dot
34	ModifiedDateTime	Date and time the native file was last modified	DateTime		1/1/2001 12:01:02 PM
35	ModifiedDate	Date the native file was last modified	Date		01/01/2001
36	ModifiedTime	Time the native file was last modified	Text	20	12:01:02 PM
37	PrintDateTime	Date and time the native file was last printed from native metadata	DateTime		1/1/2001 12:01:02 PM
38	PrintDate	Date the native file was last printed from native metadata	Date		01/01/2001
39	PrintTime	Time the native file was last printed from native metadata	Text	20	12:01:02 PM
40	AccessedDateTime	Date and time the native file was last accessed	DateTime		1/1/2001 12:01:02 PM
41	AccessedDate	Date the native file was last accessed	Date		01/01/2001
42	AccessedTime	Time the native file was last accessed	Text	20	12:01:02 PM
43	CreatedDateTime	Date and time the native file was last created	DateTime		1/1/2001 12:01:02 PM
44	CreatedDate	Date the native file was last created	Date		01/01/2001
45	CreatedTime	Time the native file was last created	Text	20	12:01:02 PM
46	ApptStartDateTime	Date and time for the Start of an Appointment	DateTime		1/1/2001 12:01:02 PM
47	ApptStartDate	Date for the Start of an Appointment	Date		01/01/2001
48	ApptStartTime	Time for the Start of an Appointment	Text	20	12:01:02 PM
49	ApptEndDateTime	Date and time for the End of an Appointment	DateTime		1/1/2001 12:01:02 PM
50	ApptEndDate	Date for the End of an Appointment	Date		01/01/2001
51	ApptEndTime	Time for the End of an Appointment	Text	20	12:01:02 PM
52	HasTrackChanges	Document has Track Changes	Text	5	Yes or No
53	HasComments	Document has Comments	Text	5	Yes or No
54	HasHiddenRowCol	Document has Hidden Row or Columns	Text	5	Yes or No
55	HasHiddenSheet	Document has Hidden Sheet	Text	5	Yes or No
56	HasNotes	Document has Notes	Text	5	Yes or No
57	SuspectDocExt	Document may not match actual data	Text	5	Yes or No
58	ConvIndex	E-mail thread identification	Text	Full	01C72C4C . . . E1493B522 (Various Length)
59	PageCount	Number of image pages produced for the document	Num	6	2
60	MD5File	MD5 Hash of the actual contents of the document	Text	50	C564668821D34200FF3E32C9BFBCGC80
61	MD5Dupe	MD5 Hash used for deduplication (different than MD5File for emails)	Text	50	FF3E32C9BFBCGC80C564668821D34200
62	Placeholder	Notation when slipsheet inserted	Text	50	Encrypted File, Archive File, Empty File, Produced as Native, Corrupt File, ...
63	Redacted	If Redacted	Text	5	Yes or No
64	TimeZoneProcessed	Time Zone Processed	Text	20	UTC
65	Designation	Designation	Text	300	Confidential; AEO; Highly Confidential
66	ProdRound	Production Round	Text	10	Prod001

Field Order	Standard Field Name	Description only on [Loadfile Key] can be chosen	Data Type	Max Size	Example
67	DeletedFile	True if the file was detected as deleted during processing	Text	5	True or False
68	DuplicateCount	Number of duplicates removed by de-duplication	Num	6	3
69	DuplicatePathList	List of paths where the document was found during de-duplication	Text	Full	\OriginalFilePath\Folder\sample.doc; \email.PST\Message1\sample.doc; \Users\User1\Desktop\OtherFile.doc;
70	EmailDomains	List of the email domains contained in an email	Text	Full	company.com; othercompany.com;
71	EmailAddresses	List of the email addresses contained in an email header	Text	Full	someone@company.com; someoneelse@othercompany.com;