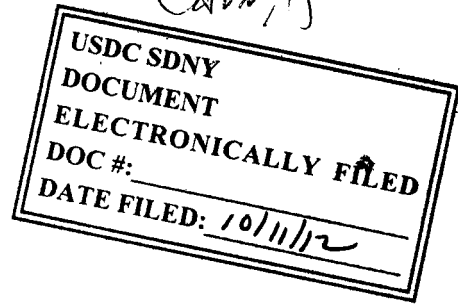


Chen, J



UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

-----X	
THE AMERICAN SOCIETY OF MEDIA	:
PHOTOGRAPHERS, INC., GRAPHIC ARTISTS	:
GUILD, PICTURE ARCHIVE COUNCIL OF	:
AMERICA, INC., NORTH AMERICAN NATURE	:
PHOTOGRAPHY ASSOCIATION,	:
PROFESSIONAL PHOTOGRAPHERS OF	:
AMERICA, LEIF SKOOGFORS, AL	:
SATTERWHITE, MORTON BEEBE, ED KASHI,	:
JOHN SCHMELZER, SIMMS TABACK,	:
LELAND BOBBE, JOHN FRANCIS FICARA,	:
and DAVID W. MOSER,	:
on behalf of themselves and all others	:
similarly situated,	:
	:
Plaintiffs,	:
	:
-against-	:
	:
GOOGLE, INC.,	:
	:
Defendant.	:
-----X	

Case No. 10-CV-02977 (DC)

This Stipulation Regarding the Format of Document Productions shall govern the parties in the above-captioned case (the "Litigation").

I. GENERAL PROVISIONS

A. The parties will make reasonable efforts to prepare responsive and nonprivileged data for production in accordance with the agreed-upon specifications set forth below. These specifications apply to hard copy documents or electronically stored information ("ESI") which are to be produced in the first instance in this Litigation.

B. SECURITY. Both parties will make reasonable efforts to ensure that any productions made are free from viruses and provided on encrypted media for submission.

C. CONFIDENTIALITY DESIGNATION. Responsive documents in TIFF format will be stamped with the appropriate confidentiality designations in accordance with the

Protective Order in this matter. Each responsive document produced in native format will have its confidentiality designation identified in the filename of the native file.

D. NON-STANDARD FILES. The format of production of non-standard electronic files, large oversized documents, etc. will be discussed before production to determine the optimal production format.

E. PRODUCTION MEDIA. Documents shall be produced on external hard drives, readily accessible computer(s) or electronic media ("Production Media"). Each piece of Production Media shall identify a production number corresponding to the production volume (e.g., "VOL001", "VOL002"), as well as the volume of the material in that production (e.g. "-001", "-002"). Each piece of production media shall also identify: (1) the producing party's name; (2) the production date; (3) the Bates Number range of the materials contained on the Production Media; and (4) the set(s) of requests for production for which the documents are being produced.

II. DATA PROCESSING

A. KEYWORD SEARCHING. The parties agree to use keyword searching to limit the universe of potentially responsive documents to be reviewed. The parties have exchanged search term lists and will apply those terms to the discoverable population and produce a hit report on a date to be agreed upon by the parties. The hit reports will include the number of documents that were responsive to each term as well as the total number of unique documents across all terms. The parties may then submit refinements to the search terms and re-run the searches and generate a hit report for a second pass on a date to be agreed upon by the parties. In the event either parties believes further refinement is necessary thereafter, the parties shall meet and confer in good faith to establish the date by which any proposed refinements and any correlating hit report must

be exchanged. During the course of the Litigation, the parties agree to securely preserve in such a manner so as to preserve without modification, alteration or addition, the results of the search performed and any documents determined by a producing party to be non-responsive to the searches performed or their review. Any search methodology employed must open compound and nested files and de-compression of archives and OCR any non-searchable imaged documents prior to searching. The search utilities employed must support the use of stemmed searches (e.g. using ~ to include variations on a keyword), bi-directional proximity searches (e.g. w/n where "n" is the number of words in between the term sets in either direction) and Boolean searches.

B. CULLING/FILTERING. Each party will use its best efforts to filter out common system files and application executable files by using a commercially reasonable hash identification process. Hash values that may be filtered out during this process are located in the National Software Reference Library ("NSRL") NIST hash set list. Additional culling of system file types based on file extension may include, but are not limited to: WINNT, LOGS, DRVS, MP3 (to the extent that they are only music files and not recordings of relevant conversations, voicemails or presentations, as discussed in Section VI (A)), C++ Program File (c), C++ Builder 6 (cpp), Channel Definition Format (cdf), Creatures Object Sources (cos), Dictionary file (dic), Executable (exe), Hypertext Cascading Style Sheet (css), JavaScript Source Code (js), Label Pro Data File (IPD), Office Data File (NICK), Office Profile Settings (ops), Outlook Rules Wizard File (rwz), Scrap Object, System File (dll), Temporary File (tmp), Windows Error Dump (dmp), Windows Media Player Skin Package (wmz), Windows NT/2000 Event View Log file (evt), Python Script files (.py, .pyc, .pud, .pyw), Program Installers.

C. The parties agree to use commercially accessible software to open password-protected or encrypted files.

D. DEDUPLICATION. A party is only required to produce a single copy of a responsive document and a party may de-duplicate responsive ESI (based on MD5 or SHA-1 hash values at the document level) across Custodians. For emails with attachments, the hash value is generated based on the parent/child document grouping. A party may also de-duplicate “near-duplicate” email threads as follows: In an email thread, only the final-in-time document need be produced, assuming that all previous emails in the thread are contained unaltered within the final message. The custodians of duplicates encountered and removed should be referenced in the coding field titled “Duplicate” or “DedupedCustodian.” Where a prior email contains an attachment, that email and attachment shall not be removed as a “near-duplicate.” To the extent that de-duplication through MD5 or SHA-1 hash values is not possible, the parties shall meet and confer to discuss any other proposed method of de-duplication.

III. PRODUCTION OF HARD COPY DOCUMENTS

A. TIFFs. Documents that exist in hard copy format only shall be scanned and produced as single page Group IV TIFFs, with at least 300 dots per inch (dpi). Each TIFF image shall be named according to the corresponding Bates number associated with the document. Each image shall be branded according to the Bates number and agreed upon confidentiality designation. TIFFs shall show all text and images that would be visible to a user of the hard copy documents.

B. OCR TEXT FILES. A commercially acceptable technology for optical character recognition “OCR” shall be used for all scanned, hard copy documents or non-searchable imaged ESI. OCR text shall be provided as a single text file for each document, and the

filename itself should match its respective TIFF filename. The text files will not contain the redacted portions of the documents.

C. DATABASE LOAD FILES/CROSS-REFERENCE FILES. Documents should be provided with (a) an ASCII delimited data file (.txt, .dat, or .csv), and (b) an image load file that can be loaded into commercially acceptable production software (e.g., Concordance, Summation). Each TIFF in a production must be referenced in the corresponding image load file. The total number of documents referenced in a production's data load file should match the total number of designated document breaks in the Image Load file(s) in the production.

D. CODING FIELDS. The following information shall be produced in the delimited data file accompanying hard copy documents: (a) BegBates, (b) EndBates, (c) Custodian, (d) Confidentiality, and (e) Redacted.

E. BATES NUMBERING. All images must be assigned a unique Bates number that is sequential within a given document and across the production sets.

F. UNITIZING OF DOCUMENTS. In scanning paper documents, distinct documents should not be merged into a single record, and single documents should not be split into multiple records (*i.e.*, paper documents should be logically unitized). The parties will use reasonable efforts to unitize documents correctly.

G. IDENTIFICATION OF HARD COPY DOCUMENTS. The parties will utilize best efforts to ensure that hard copy records for a particular Custodian that are included in a single production are produced in consecutive Bates stamp order. The parties will identify which documents in a production are scanned hard copy documents either in the cover letter accompanying the production or in a coding field titled "DocumentType."

IV. PRODUCTION OF ELECTRONICALLY STORED INFORMATION

A. METADATA FIELDS AND PROCESSING. Each of the metadata and coding fields set forth in Appendix 1 that can be extracted shall be produced for that document. The parties are not obligated to populate manually any of the fields in Appendix 1 if such fields cannot be extracted from a document, with the exception of the following: (a) BegBates, (b) EndBates, (c) BegAttach, (d) EndAttach; (e) Custodian, (f) Confidentiality, and (g) Redacted, which should be populated by the party or the party's vendor. The parties will make reasonable efforts to ensure that metadata fields automatically extracted from the documents are correct. Parties may request other native files be produced as described in Section IV.I. below.

B. TIFFs. Unless excepted below, single page Group IV TIFFs should be provided, at least 300 dots per inch (dpi). Each TIFF image file should be one page and named according to the unique Bates number, followed by the extension ".TIF". The parties will make reasonable efforts to ensure that original document orientation is maintained (i.e., portrait to portrait and landscape to landscape). During the course of the Litigation, an electronic copy of the originally data shall be securely preserved in such a manner so as to preserve without modification, alteration or addition the content of such data including any metadata therewith.

C. TEXT FILES. For each document, a text file should be provided along with the image files and metadata. When possible, the text of native files should be extracted directly from the native file.

D. DATABASE LOAD FILES/CROSS-REFERENCE FILES. (a) an ASCII delimited data file (.txt, .dat, or .csv), and (b) an image load file that can be loaded into commercially acceptable production software (e.g., Concordance, Summation). Each

TIFF in a production must be referenced in the corresponding image load file. The total number of documents referenced in a production's data load file should match the total number of designated document breaks in the Image Load file(s) in the production. The total number of documents in a production should match the total number of records in the database load file. Each TIFF in a production must be referenced in the corresponding image load file. The total number of pages referenced in a production's image load file should match the total number of TIFF files in the production.

E. BATES NUMBERING. All images must be assigned a unique Bates number that is sequential within a given document and across the production sets.

F. PRESENTATIONS. The parties shall take reasonable efforts to process presentations (MS PowerPoint, Google Presently) with hidden slides and speaker's notes unhidden, and to show both the slide and the speaker's notes on the TIFF image.

G. SPREADSHEETS. TIFF images of spreadsheets need not be produced unless redacted, in which instance, spreadsheets will be produced in TIFF with OCR. Native copies of spreadsheets should be produced with a link reference in the NativeLinkPath and NativeLinkName fields, along with extracted text and applicable metadata fields set forth in Appendix 1. A TIFF placeholder indicating the document was provided in native format should accompany the database record. If a spreadsheet has been redacted, TIFF images and OCR of the redacted document will suffice in lieu of a native file and extracted text. The parties will make reasonable efforts to ensure that any spreadsheets that are produced only as TIFF images are formatted so as to be readable.

H. HIGH RESOLUTION PHOTOGRAPHS/ARTWORK. To the extent a response to discovery would require the production of high resolution or raw photographic images

or artwork, those files may be converted to lower resolution images with the original aspect ratio maintained and produced in .jpg format.

I. MULTIMEDIA OR NON-PRINTABLE FILES. To the extent a response to discovery requires production of ESI that is not printable to TIFF format, those files should be produced in their original native file format.

J. PROPRIETARY FILES. To the extent a response to discovery requires production of ESI accessible only through proprietary software, the parties should continue to preserve each version of such information. The parties shall meet and confer to finalize the appropriate production format.

K. REQUEST(S) FOR ADDITIONAL NATIVE FILES. If good cause exists to request production of specified files, other than those specifically set forth above, in native format, the party may request such production and provide an explanation of the need for native file review, which request shall not unreasonably be denied. Any native files that are produced should be produced with a link reference in the NativeLinkPath and NativeLinkName fields, along with all extracted text and applicable metadata fields.

L. REDACTION OF INFORMATION. If documents are produced containing redacted information, the producing party shall supply a list of the documents for any such claim(s) of privilege, indicating the grounds for the redaction and the nature of the redacted material. During the course of the Litigation, an electronic copy of the originally, unredacted data shall be securely preserved in such a manner so as to preserve without modification, alteration or addition the content of such data including any metadata therewith. If scanning is necessary to create a redacted version of a document, the producing party shall produce a corresponding load file identifying the custodian

from whom the document was collected and such documents shall subject the TIFF files to Optical Character Recognition (“OCR”) process if the TIFF files do not already have extractable text, such that the processed files become searchable.

V. PROCESSING OF THIRD-PARTY DOCUMENTS

A. A party that issues a non-party subpoena (“Issuing Party”) shall include a copy of this Stipulation with the subpoena and state that the parties to the Litigation have requested that third-parties produce documents in accordance with the specifications set forth herein.

B. The Issuing Party shall produce any documents obtained pursuant to a non-party subpoena to the opposing party.

C. If the non-party production is not Bates-stamped, the Issuing Party will endorse the non-party production with unique prefixes and Bates numbers prior to producing them to the opposing party.

D. Nothing in this stipulation is intended to or should be interpreted as narrowing, expanding, or otherwise affecting the rights of the parties or third-parties to object to a subpoena.

VI. SEARCHING

A. Electronic Keyword Searching. The parties shall meet and confer to try to develop a mutually agreeable list of search terms and protocols prior to the production of documents. With respect to custodial documents, the parties will conduct electronic searches using such agreed search term lists. The parties agree that instant messages will be searched to the extent a custodian has enabled the logging of instant message or chat files. The parties agree that voicemails will be searched to the extent they have been transcribed and maintained in a custodian’s e-mail files or have been identified by a

custodian as containing potentially responsive information. The parties agree that they will utilize commercially reasonable efforts to open any encrypted files.

B. Locations That Will Be Searched for Responsive Documents. The parties shall meet and confer to try to develop a mutually agreeable list of custodians to be searched. The parties will search any electronic files or folders, or other parts of media, including any internal and external hard drives and other ESI venues (including, but not limited to, recordable optical media, media cards, thumb drives, non-volatile memory, floppy disks, work desktop and laptop computers, email servers, intranet servers, internet servers, network shares, instant messaging history, internal discussion forums, public data shares, externally hosted systems, cloud based storage/systems, and/or database servers) for each identified Custodian that the Custodian, information technology representative and/or legal counsel reasonably anticipates may contain non-duplicative Responsive Documents. The parties agree to search central repositories, including central databases, or relevant portions thereof to the extent that the party reasonably anticipates they contain non-duplicative Responsive Documents. The parties agree to meet and confer to limit the scope of production from central repositories if the search of central repositories (or relevant portions thereof) that the producing party anticipates contain Responsive Documents is unduly burdensome or is likely to be unreasonably inefficient in identifying relevant documents. The parties will then meet and confer to discuss the collection of Responsive Documents from such repositories, including potentially using other effective collection methodologies.

C. Locations That Will Not Be Searched for Responsive Documents. The following locations will not be searched under any circumstances, and as such need not be

preserved: personal digital assistants, mobile phones, automated disaster recovery backup systems and/or disaster recovery backup tapes. In addition, the parties agree that all custodial emails will be searched. Notwithstanding the foregoing, the parties agree that Responsive Documents that a Custodian indicates are stored on an archival storage medium that the Custodian can readily identify and locate, that cannot be located in any other repository of information, and that can reasonably be searched, will be searched. In addition, nothing in this paragraph shall limit a receiving party's right to request from a producing party more information about the nature of and burden associated with obtaining documents from a particular location. The parties further recognize their obligations to preserve any potentially relevant sources of data, whether live or in archival form, for purposes of this Litigation. The parties shall meet and confer with respect to Google's objection to producing documents resulting from automatic "autosave" functions and, if the need arises, may any dispute regarding the search of these documents with the Court if necessary at that time.

D. Locations That Require Clarification. The parties shall meet and confer to clarify any outstanding issues or other matters that arise in the course of the parties' discovery.

E. Source Code. To the extent relevant to the Litigation, source code will be made available for inspection pursuant to the terms of the Protective Order. The parties agree that the search terms will not be applied to source code.

VII. MISCELLANEOUS PROVISIONS

A. Objections Preserved. Nothing in this protocol shall be interpreted to require disclosure of irrelevant information or relevant information protected by the attorney-client privilege, work-product doctrine, or any other applicable privilege or immunity.

The parties do not waive any objections as to the production, culling methodologies, discoverability, admissibility, or confidentiality of documents and ESI.


B. No party shall be required to search for or identify on their respective privilege log any document or communication with outside litigation counsel dated on or after the filing of the Litigation or *The Authors Guild, Inc., et al. v. Google Inc.*, Case No. 05 CV 8136-DC (S.D.N.Y.), which absent this provision, the party may have been obligated to so identify on said privilege log.

C. Communications with Testifying Experts. Testifying experts shall not be subject to discovery of any draft of their reports in this case and such draft reports, notes, outlines, or any other writings leading up to an issued report(s) in this litigation are exempt from discovery. In addition, all communications between counsel for a party and that party's testifying expert, and all materials generated by a testifying expert with respect to that person's work, are exempt from discovery unless they relate to the expert's compensation or identify facts, data or assumptions relied upon by the expert in forming any opinions in this litigation and such information is not already disclosed in the expert's report.


D. No Effect on Cost Shifting. Nothing in this Agreement shall affect, in any way, a producing party's right to seek reimbursement for costs associated with collection, review, and/or production of documents or ESI.

E. Timing. Production of ESI shall commence on a rolling basis promptly after the parties finalize custodian and search term lists.

DO NOT WRITE IN THESE SPACES



U.S.D.J. LISG
10-11-12



Appendix 1: ESI Metadata and Coding Fields

A. Image Load File shall contain the following comma-delimited fields:
BEGBATES, VOLUME, IMAGE FILE PATH, DOCUMENT BREAK, FOLDER
BREAK, BOX BREAK, PAGE COUNT

B. Metadata Load File shall be delimited according to the following characters:

- o Delimiter = D(ASCII:0020)
- o Text-Qualifier = b (ASCII:0254)
- o NewLine Char in Text = ® (ASCII:0174)

C. The following fields can appear in the metadata load file in the order displayed below:

Field Name	Field Description
BegBates	Beginning Bates number as stamped on the production image
EndBates	Ending Bates number as stamped on the production image
BegAttach	First production Bates number of the first document in a family
EndAttach	Last production Bates number of the last document in a family
Custodian	Individual from whom the documents originated
Deduped Custodian	Individual(s) whose documents de-duplicated out; only relevant for global de-duplication.
DocumentType	Description of document (Email, Attachment, Edoc, Hard Copy)
FileType	Application used to create the file (e.g. Microsoft Word, Microsoft Excel, Adobe Acrobat)
FileSize	Size of file in Bytes
FileExt	Characters of the filename indicating the relevant program used to open the file (file extension)
Subject	Subject line of email
DateSent	Date email was sent (format: MM/DD/YYYY)
TimeSent	Time email was sent (format: HH:MM:SS)
To	All recipients that were included on the "To" line of the email
From	The name and email address of the sender of the email
Cc	All recipients that were included on the "CC" line of the email
Bcc	All recipients that were included on the "BCC" line of the email
FileName (Edoc only)	Filename of an attachment or loose file.
FilePath	The directory structure of the original file(s). Any container name is included in the path. (e.g. \Documents\jsmith\May2004.pst\Inbox\)
FileAuthor (Edoc only)	'Author' field embedded in extended properties of attachment or loose Edoc
FileTitle (Edoc only)	'Title' field embedded in extended properties of attachment or loose Edoc
FileCreationDate (Edoc only)	'DateCreated' field embedded in extended properties of

FileDatePrinted (Edoc only)	attachment or loose Edoc 'Last Printed' field embedded in extended properties of attachment or loose Edoc
DateModified (Edoc only)	Date an electronic document was last modified (format: MM/DD/YYYY)
DateCreated (Edoc only)	Date the document was created (format: MM/DD/YYYY)
HashValue	The MD5 or SHA-1 hash value.
NativeLinkPath	Path to the native file produced in the delivery
NativeLinkName	Name of native file produced in the delivery
TextPath	Path to the text file containing the extracted text or OCR produced in the delivery
TextName	Name of text file produced in the delivery
Confidentiality	Indication of the confidential status of the document
Redacted	Indication that the document being produced is redacted