

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

KEWAZINGA CORP.,

Plaintiff,

v.

GOOGLE LLC,

Defendant.

Case No. 1:20-cv-1106-LGS

**STIPULATED DISCOVERY PLAN AND ~~PROPOSED~~ ORDER
FOR ELECTRONICALLY STORED INFORMATION**

The Parties to this action, by and through their undersigned attorneys, hereby agree and stipulate to the following:

1. This Stipulated Discovery Plan for Electronically Stored Information shall apply to this case in all respects.
2. Electronically stored information (“ESI”) will be part of the discoverable material in this case and the parties agree to cooperatively exchange discoverable material and use reasonable, good faith and proportional efforts to identify, preserve, collect, and produce information relevant to a party’s claim or defense and proportional to the needs of the case, considering the importance of the issues at stake in the action, the amount in controversy, the parties’ relative access to relevant information, the parties’ resources, the importance of the discovery in resolving the issues, and whether the burden or expense of the proposed discovery outweighs its likely benefit.

Sources of and Limitations on ESI

3. Parties are expected to use reasonable, good faith and proportional efforts to preserve, identify and produce relevant information consistent with [Fed. R. Civ. P. 26\(b\)\(1\)](#).¹ This includes identifying appropriate limits to discovery, including limits on custodians, identification of relevant subject matter, time periods for discovery and other parameters to limit and guide preservation and discovery issues. A party's meaningful compliance with this order and efforts to promote efficiency and reduce costs will be considered in cost-shifting determinations.

4. As in all cases, costs may be shifted for disproportionate ESI production requests pursuant to [Federal Rule of Civil Procedure 26](#). Likewise, a party's nonresponsive or dilatory discovery tactics are cost-shifting considerations.

5. The parties recognize the obligation to take reasonable and proportional steps to identify and preserve discoverable information within the party's possession, custody and control, and agree to identify and preserve ESI only from the following data sources:

- a. Electronic mail;
- b. PC or laptop hard drives and/or storage that are reasonably accessible;
- c. Network file servers and archives that are reasonably accessible.

6. These data sources are not reasonably accessible because of undue burden or cost pursuant to [Fed. R. Civ. P. 26\(b\)\(2\)\(B\)](#) and ESI from these sources will be preserved pursuant to normal business retention, but not searched, reviewed, or produced: -

- a. backup systems and/or tapes used for disaster recovery; and

¹ Information can originate in any form, including ESI and paper, and is not limited to information created or stored electronically.

- b. systems, server and network logs; and
- c. systems no longer in use that cannot be accessed.

7. The parties generally agree that discovery should be guided by Judge Schofield's Individual Rule II(A)(1)(b), but acknowledge that some of the claims and defenses raised in the litigation implicate facts and interactions from 2015 and earlier. Accordingly, the parties agree that Judge Schofield's Individual Rule II(A)(1)(b) should not limit discovery in this case, to the extent discovery still exists and is reasonably accessible.

8. Absent a showing of good cause by the requesting party, the parties shall not modify, on a going-forward basis, their internal procedures used by them in the ordinary course of business to backup, archive, store or manage information systems and ESI, except that the parties will preserve non-duplicative discoverable information currently in their possession, custody or control.

9. Each party shall identify the ten custodians most likely to have discoverable material in this case and collect ESI from same.

10. The parties generally agree that discovery should be guided by Judge Schofield's Individual Rule II(A)(1)(d), but agree that it should not limit discovery in this case given the complexity of the litigation and the number of claims and defenses at issue.

11. Absent a showing of good cause the following categories of ESI need not be searched, preserved or collected and are considered inaccessible due to undue burden and/or expense, or because the information is more reasonably available elsewhere:

- a. Unallocated, slack space, deleted data, file fragments or other data accessible by use of computer forensics;

- b. Random access memory (RAM), temporary files, or other ephemeral data that is difficult to preserve without disabling the operating system;
- c. Data relating to online access, such as temporary Internet files, browser history, file or memory caches and cookies;
- d. Data in metadata fields that are frequently updated automatically as part of the usual operation of a software application, operating system or network (e.g., date last opened or printed);
- e. Backup or archived data that is substantially duplicative of data that is more reasonably accessible elsewhere;
- f. Dynamic fields of databases or log files that are not retained in the usual course of business;
- g. Video and audio recordings;
- h. Voice messages;
- i. Automatically saved versions of documents and emails;
- j. Instant messaging and chat application data;
- k. Text messages sent to or from cell phones;
- l. Other data stored on handsets, mobile devices, cell phones, personal digital assistants, tablets, or blackberry devices, such as calendar, contact or notes, provided that copies of such information is routinely saved or stored elsewhere;
- m. Logs of calls made from mobile devices, cell phones or blackberries;

- n. Operating system files, executable files, network, server or system logs, other than accused instrumentalities;
- o. Data from systems that are no longer in use (“Legacy Data”) that is unreadable or unintelligible on current systems, other than accused instrumentalities.
- p. Electronic data temporarily stored by laboratory equipment or attached electronic equipment.
- q. Any other source of ESI not listed in paragraph 5, above.

Search

12. The parties agree that in responding to an initial [Fed. R. Civ. P. 34](#) request, or earlier if appropriate, they will meet and confer about methods to search ESI in order to identify ESI that is subject to production in discovery and filter out ESI that is not subject to discovery.

13. 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 file types based on file header information may include, but are not limited to: Application Package File, Backup Files, Batch Files, Binary Disc Image, C++ File Formats, Cascading Style Sheet, Configuration File, Database File, Dictionary Files, Dynamic Link Library, Event Log Files, Executable Files, Hypertext Cascading Stylesheet, Java Archive Files, JavaScript files, JavaScript Source Code and Class Files, Macintosh Resource Fork Files, MP3 Files, MP4 Files, Package Manager Files, Program Files, Program Installers, Python Script Files, Quicktime Files, Shell Script Files, System or Temporary Files, Thumbnail Cache Files, Troff Files, TrueType Font Files, Video

Media Files, Waveform Audio File Format, Windows Cabinet File, Windows Command Files, Windows File Shortcut, Windows Help Files, Windows Metafiles and Enhanced Metafiles, Windows Spool Files, Windows System File. Source code files will be provided according to the Protective Order and not included in custodial data productions.

General Production Format

14. The parties will produce ESI in TIFF image format provided that the documents do not become illegible or unusable when converted to TIFF image format. Certain documents that become illegible or unusable when converted to TIFF image format (Microsoft Excel files, other similar spreadsheet application files, Microsoft Access and Microsoft Project), must be produced in native format. The parties reserve their rights to reasonably seek additional electronic documents in their native format.

15. Documents originating in paper form will be scanned at 300 DPI and undergo Optical Character Recognition (OCR). These documents will be produced in single-page TIFF image format as indicated below together with document-level text files. Paper documents will be logically unitized to reflect correct document boundaries for each document.

16. All production documents will be produced with image load files, the data fields provided in Table 1, below, and either extracted text or text generated using OCR that render documents searchable. In those instances where redaction is used, OCR text will be provided in lieu of the extracted text to allow for removal of the redacted text from production. For documents produced in native format, in addition to producing extracted text and the data fields in Table 1 below, the producing party will provide slip sheets endorsed with the production number and level of confidentiality pursuant to any applicable protective orders in this case.

17. ESI will be processed and produced with all hidden text (e.g., track changes, speaker's notes, hidden rows or columns, comments, markups, notes, etc.) and formulas exposed, expanded and extracted and rendered in the TIFF image.

18. ESI and documents originally in paper form shall be produced as kept in the usual course of business. A producing party need not include in its production any indication of the document request(s) to which a document may be responsive.

Image Format

19. Documents that are converted to TIFF or JPEG image format will be produced in accordance with the following technical specifications:

- a. Single-page, 1-bit, group IV TIFF image files at 300 dpi;
- b. Image file names cannot contain embedded spaces;
- c. Bates numbers should be endorsed on the lower right corner of all TIFF or JPEG images and will be a unique, consistently formatted identifier, i.e. alpha prefix along with a fixed length number (e.g., ABC0000001). The number of digits in the numeric portion of the bates number format should not change in subsequent productions;
- d. Confidential designations, if any, will be endorsed on the lower left corner of all TIFF or JPEG images;
- e. Presentations, including PowerPoint slides, should be rendered to TIFF in full slide image format, with any speaker notes following the appropriate slide image;
- f. Drawings, photographs and any other graphical files, including 2-D and 3-D drawings, engineering drawings, AutoCAD and SolidWorks,

should be rendered to single-page TIFF image format. The format may be adjusted (e.g., drawing copied onto more than one page).

g. Excel spreadsheets will only be rendered to TIFF image format if (1) the file can be fully viewed, meaning that all hidden information in the Excel file (e.g., rows, columns and comments) will be rendered to TIFF; or (2) the Excel file has undergone redaction. If a native Excel file is to be produced in place of TIFF images, a slipsheet TIFF image placeholder containing the native file name will be produced.

h. Images will be delivered with an image load file in the Opticon (.OPT) format as follows:

REL00001,REL001,D:\IMAGES\001\REL00001.TIF,Y,,3

Column One (REL00001) –the page identifier

Column Two (REL001) –volume identifier; not required, but a space is required in each line of the load file for this field as illustrated below:

REL00001, ,D:\IMAGES\001\REL00001.TIF,Y,,3

REL00002, ,D:\IMAGES\001\REL00002.TIF,,,

Column Three (D:\IMAGES\001\REL00001.TIF) –path to the image to be loaded

Column Four (Y) – Document marker –indicates the start of a unique document.

Column Five (blank) –can be used to indicate box

Column Six (blank) –can be used to indicate folder

Column Seven (3) –often used to store page count, but unused in Relativity

20. The parties agree to respond to reasonable and specific requests for the production of higher resolution or color images. Nothing in this Stipulation shall preclude a producing party from objecting to such requests as unreasonable in number, timing or scope, provided that a producing party shall not object if the document as originally produced is illegible, difficult to read, or color provides substantive additional information in the document. The producing party

shall have the option of responding by producing a native-file version of the document. If a dispute arises with regard to requests for higher resolution or color images, the parties will meet and confer in good faith to try and resolve it.

Data Format

21. Extracted data, including the data fields listed in Table 1 below, will be produced in a delimited .DAT file in accordance with the following technical specifications:

a. The first line of the .DAT file must be a header row identifying the field names;

b. The .DAT file must use the following default delimiters:

- i. Comma ¶ ASCII character 020
- ii. Quote ¨ ASCII character 254
- iii. Newline ® ASCII character 174
- iv. Multi-value ; ASCII character 059

c. Date fields should be provided in the format: mm/dd/yyyy;

d. All attachments should sequentially follow the parent document/email;

e. The metadata of email, attachments and application files should be extracted and produced in a .DAT file using the fields and formatting in

Table 1 below:

TABLE 1: Metadata Field List

<i>Field Name</i>	<i>Description</i>
PRODBEG (or BEGBATES)	First Bates number of each document being produced
PRODEND (or ENDBATES)	Last Bates number of each document being produced
BEGATTACH	First Bates number of attachment range
ENDATTACH	Last Bates number of attachment range
CUSTODIAN	Email: User mailbox where the email resided Application File: Individual from whom the document originated Includes the Individual (Custodian) from whom the documents

	originated and all Individual(s) whose documents de-duplicated out (De-Duped Custodian) Alternatively, De-Duped Custodians can be included in the OTHER CUSTODIANS field
FROM	Email: Sender of an email Application File: (empty) **semi colon should be used to separate multiple entries
TO	Recipient(s) of an email **semi colon should be used to separate multiple entries
CC	Recipient(s) copied on an email **semi colon should be used to separate multiple entries
BCC	Recipients blind copied on an email **semi colon should be used to separate multiple entries
SUBJECT	Email: Subject line of the email
DATESENT	Email: Date the email was sent Application File: (empty)
TIMESENT	Email: Time the email was sent Application File: (empty) **This data must be a separate field and cannot be combined with the DATESENT field
AUTHOR	Email: (empty) Application File: Author of document
DATECREATED	Email: (empty) Application File: Date the document was created
DATELASTMOD (or DATEMOD)	Email: (empty) Application File: Date the document was last modified
FILESIZE	Size of application file document/email in KB
FILETYPE	Email, attachment or loose application file
FILEEXT	The file extension of the native file
FILENAME	The name of the application file, including file extension (if the information is easily extractable without undue burden on the parties)
PAGECOUNT	The number of pages of each individual document
DESIGNATION	The confidentiality designation assigned to the document pursuant to any confidentiality/protective order in the case
NATIVE_FILE	Link to the native email or application file ** The link must be named per the production Bates number
HASHVALUE	Hash value of each email or application file
OTHER CUSTODIANS (if used as an alternative to identifying De-Duped Custodians in the CUSTODIAN field)	Additional custodians who possessed a duplicate copy of an email or application file (if de-duping across custodians)
TEXT	The extracted text or OCR text of the application file or email

f. The .DAT file for scanned paper documents must contain, at a minimum, the following fields: PRODBEG, PRODEND and CUSTODIAN.

22. In addition to the metadata fields set forth in Table 1, any party may request, on a case-by-case basis, the producing party to provide additional metadata for specific files if the party reasonably believes the information may be relevant.

Searchable Text

23. Searchable text of entire documents will be produced either as extracted text for all documents that originate in electronic format, or, for paper documents and any document from which text cannot be extracted, as text generated using Optical Character Recognition (OCR) technology. For redacted documents, the full text of the redacted version of the document will be produced.

24. Searchable text will be produced as a document-level multi-page ASCII text file with the text file named the same as the PRODBEG field, placed in a separate folder. The full path of the text file must be provided in the .DAT file for the TEXT field.

Native Files

25. Native file documents may be included with the electronic production using the below criteria:

a. Native file documents must be named the same as the PRODBEG number;

b. The full path of the native file must be provided in the .DAT file for the NATIVE_FILE field;

c. When native files are produced because rendering the file to TIFF would not result in viewable images, a TIFF image slipsheet placeholder will

be produced endorsed with the file name and the legend “Document Produced in Native Format” (or something similar);

d. The confidentiality designation under the Protective Order to be entered in this action will be produced in the load file in the DESIGNATION field.

26. If documents produced in native format are printed for use in deposition, motion or hearing, the party printing the document must label the front page of the file that is printed with the corresponding production number and a sequencing page number and, if applicable, the confidentiality designation assigned by the producing party to that file under an applicable protective order to be entered in this action.

De-Duplication

27. A party is required to produce only a single copy of a responsive document, and a party may de-duplicate responsive ESI across Custodians. A party may also de-duplicate email threads and attachments as follows: In an email thread, only the most evolved responsive email in a thread will be produced. Where an earlier-in-thread email has a responsive attachment not contained within the most evolved responsive email, the most evolved earlier-in-thread email containing the attachment will also be produced along with its attachment.

28. Should a producing party de-duplicate any documents in accordance with the procedure outlined above, the producing party waives any objection as to the authenticity of the version of the document produced and the fields of data associated with such copy.

29. Should a producing party de-duplicate any documents in accordance with the procedure outlined above, the producing party agrees to produce in the CUSTODIAN (or

OTHER CUSTODIANS) data field listed in Table 1 the name of each custodian who possessed or controlled a duplicate copy of any such documents.

30. The parties agree that an email that includes content in the BCC or other blind copy field shall not be treated as a duplicate of an email that does not include content in the BCC field, even if all remaining content in the email is identical.

31. If applicable, no provision of this Order affects the inspection or production of source code which will be collected and made available consistent with the Protective Order governing this case.

32. For good cause shown, the receiving party shall have the right to request all duplicates of a produced document consistent with the limitations and procedures in [Federal Rule of Civil Procedure 26\(b\)\(2\)\(B\)](#).

Privileged Documents and Logs

33. The parties agree under Rule 502 of the Federal Rules of Evidence that any document containing privileged information or attorney work product that is inadvertently produced shall be returned to the producing party immediately and that production of such document or documents shall not constitute a waiver of privilege or protection in this or any other action. The procedures for addressing inadvertent production of privileged or otherwise protected material will be set forth in the Protective Order governing this case.

34. Nothing contained herein is intended to or shall serve to limit a party's right to conduct a review of documents, ESI or information (including metadata) for relevance, responsiveness and/or segregation of privileged and/or protected information before production.

35. Activities undertaken in compliance with the duty to preserve information are protected from discovery under [Fed. R. Civ. P. 26\(b\)\(3\)\(A\)](#) and (B).

36. With respect to privilege logs:
- a. The parties are not required to include in their privilege logs any information generated after the filing of the complaint.
 - b. For email threaded conversations including the same authors and recipients, the parties are required to log only the most inclusive and most recent email thread.

Miscellaneous

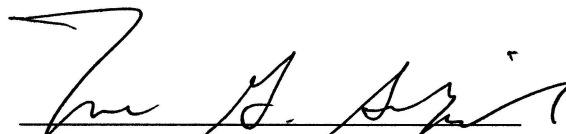
37. The parties agree to meet and confer in good faith as to any discovery dispute prior to bringing any such dispute to the attention of the court.

38. This stipulation relates solely to the protocol in this case for identifying, collecting, and producing ESI. Any party may bring a motion to modify or clarify the application of this Stipulation.

39. This stipulation will be approved and adopted as an order of the court in this action.

So Ordered.

Dated: May 5, 2020
New York, New York


LORNA G. SCHOFIELD
UNITED STATES DISTRICT JUDGE

SO STIPULATED AND AGREED

Dated: May 1, 2020

/s/ Ian G. DiBernardo

Ian G. DiBernardo
Timothy K. Gilman
Kenneth L. Stein
Saunak K. Desai
Gregory R. Springsted
STROOCK & STROOCK & LAVAN LLP

Dated: May 1, 2020

/s/ Elizabeth Weyl

John M. Desmarais
Ameet A. Modi
Steven M. Balcof
Elizabeth Weyl
DESMARAIS LLP
230 Park Avenue

180 Maiden Lane
New York, NY 10038
Tel: (212) 806-5400
Fax: (212) 806-6006
Email: idibernardo@stroock.com
Email: tgilman@stroock.com
Email: kstein@stroock.com
Email: sdesai@stroock.com
Email: gspringsted@stroock.com

Counsel for Plaintiff Kewazinga Corp.

New York, New York 10169
T: 212-351-3400
F: 212-351-3401
jdesmarais@desmaraisllp.com
amodi@desmaraisllp.com
sbalcof@desmaraisllp.com
eweyl@desmaraisllp.com

Emily H. Chen (*pro hac vice*)
DESMARAIS LLP
101 California Street
San Francisco, California 94111
T: (415) 573-1900
F: (415) 573-1901
echen@desmaraisllp.com

Counsel for Defendant Google LLC