IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF PENNSYLVANIA

IN RE WAWA, INC.	:	
DATA SECURITY LITIGATION	:	**CIVIL ACTION**
	:	
	:	**This document applies**
	:	**to all tracks.**
	:	
	:	
	:	
	:	**No. 19-6019**
	:	**and all related cases**

CASE MANAGEMENT ORDER NO. 8 (STIPULATION AND ORDER FOR THE PRESERVATION, DISCOVERY, AND PRODUCTION OF DOCUMENTS)

1. PURPOSE

This Stipulated Protocol and Order for the Discovery of Electronically Stored Information ("Protocol") will govern the preservation, discovery, and production of electronically stored information ("ESI") and hard copy documents as a supplement to the Federal Rules of Civil Procedure in the Consumer, Financial Institution, and Employee Tracks established by Case Management Order No. 2. Where this Protocol discusses "Parties," it means, with respect to the Consumer Track, the Consumer Track Plaintiffs and Defendants; with respect to the Financial Institution Track, the Financial Institution Track Plaintiffs and Defendants; and with respect to the Employee Track, the Employee Track Plaintiff(s) and Defendants. This Protocol does not alter the parties' discovery obligations or protections under the Federal Rules of Civil Procedure except as otherwise specifically provided herein.

Where this Protocol specifies that the parties are to meet and confer, if the parties are unable to reach agreement, they may seek resolution by the Court.

2. DEFINITIONS

For purposes of this Protocol and these Actions, aside from the following definitions, all technical terms used herein shall be interpreted according to their usual and customary meaning, and understanding, of those terms in the eDiscovery field.

a) "Custodial data source" means any document source used by a custodian to store documents or information and includes, but is not limited to, personal computers, laptop computers, work stations, email (whether stored locally or centrally), calendars, mobile devices (such as cellphones, pagers, blackberries, or tablets), shared network servers, shared storage systems, shared or individual network folders (excluding network folders used by departments, business units, or divisions included under non-custodial data sources) that are controlled or managed by a single individual, structured data systems if maintained or used by a single individual, cloud-based personal storage locations, and social media if maintained or used by a single individual.

b) "Documents" shall have the same definition as set forth in Federal Rule of Civil Procedure 34, and when used herein, refer to both ESI (including, but not limited to, loose documents/unstructured data, email, structured data, data compilations) and hard copy documents. As used herein, "documents" and "data" are synonymous.

c) "Document family" means a collection of pages or files maintained together and constituting a logical single communication of information but consisting of more than a single stand-alone record. Examples include but are not limited to a fax cover, the faxed document, and an attachment to the faxed document (the fax cover

being the "parent," and the document and attachment being the "children"); an email and all associated attachments and any files embedded in those attachments (the email being the "parent" and the attachments and embedded files being the "children"); and a PowerPoint presentation with embedded files (the PowerPoint being the "parent" and the embedded files being the "children").

d) "Non-custodial data source" means any data source that is not kept, maintained, or used exclusively by any particular agreed or potential custodian but which may contain potentially relevant documents. These sources will typically include those sources used by multiple individuals, by any department, business unit, or division of a Producing Party, or that are kept or maintained by such departments, business units, or divisions. For purposes of clarity, a non-custodial data source includes, but is not limited to, shared storage systems or other shared information sources (such as collaboration software, cloud-based sources, shared network folders or electronic mail folders used by, or assigned or allocated to, any relevant department, business unit, or division of a party); structured data sources; backup systems, backup media (other than those associated exclusively with a single custodian), hard copy repositories (e.g., file cabinets and storage units); or archival systems or archival storage (other than those associated exclusively with a single custodian).

e) "Search Term" means a word or a string of words or phrases joined by proximity and Boolean connectors and other syntax designed to capture potentially relevant ESI.

f) "Structured Data" means data that resides in a fixed field within a record or file, or stored in a structured format, such as databases (such as Oracle, SQL, Access)

or data sets, according to specific form and content rules as defined by each field of the database.

g) "TAR" (Technology-Assisted Review) means a process for prioritizing or coding a collection of ESI using a computerized system, using algorithms or systematic rules, that harnesses human judgments of subject matter expert(s) on a smaller set of Documents and then extrapolates those judgments to the remaining Documents in the collection. TAR systems generally incorporate statistical models and/or sampling techniques to guide the process and to measure overall system effectiveness.

h) "Unstructured Data" refers to free-form data which either does not have a data structure or has a data structure not easily readable by a computer without the use of a specific program designed to interpret the data, including but not limited to, word processing Documents, PowerPoint and comparable presentations, email, PDFs, Excel spreadsheets and comparable spreadsheets, webpages, blogs, image files, audio and video files, instant messages, text messages, and others of similar variable format.

3. LIMITATIONS AND NO WAIVER

This Protocol provides a general framework for the search and production of documents on a going-forward basis. The parties and their attorneys do not waive any attorney-client privilege or work-product protection by agreeing to make the disclosures required herein. Any such waiver shall be strictly and narrowly construed and shall not extend to other matters or information not specifically described herein. Further, all parties preserve their attorney-client privileges and other legally recognized privileges (to the extent applicable to the Parties) and there is no intent by this

Protocol, or the production of documents pursuant to this Protocol, to in any way waive or weaken these privileges. All documents produced according to the requirements of this Protocol are also governed by Case Management Order No. 6 (Protective Order) entered by the Court on July 21, 2020 (ECF No. 130). Further, the parties do not waive any objections to the production, discoverability, or confidentiality of documents or the scope of discovery, arguments in opposition to such objections to the foregoing, or their rights to compel responses to, or seek a protective order from, discovery requests.

4. E-DISCOVERY LIAISONS

Prior to commencement of the Parties' negotiations over electronic sources and systems, custodians, and search methodology, each Track shall identify to the other parties E-Discovery Liaisons who will be primarily responsible for meeting and conferring concerning ESI. Each E-discovery Liaison will, at a minimum:

a) be knowledgeable about the party's e-discovery efforts;

b) be, or have reasonable access to those who are, familiar with the party's electronic systems and capabilities, including the location, nature, accessibility, format, collection, search methodologies, and production of ESI in this matter;

c) be prepared to participate in e-discovery dispute resolutions; and

d) be, or have reasonable access to those who are, knowledgeable about the technical aspects of e-discovery, including electronic document storage, organization, and format issues, and relevant information retrieval technology, including search methodology;

e) be responsible for organizing the Party's e-discovery efforts to ensure consistency and thoroughness and to facilitate the e-discovery process.

The parties will rely on the E-Discovery Liaisons, as needed, to confer about ESI and to help resolve disputes without court intervention. The E-Discovery Liaisons for Plaintiff Tracks must report to and be authorized by the Interim Co-Lead Counsel and Interim Liaison Counsel on behalf of the respective Tracks, and Liaisons for Defendants must report to and be authorized by counsel of record for Defendants, to negotiate the resolution of disputes and to bind such parties to any such resolution. Each party will notify the other Party of its designated E-discovery Liaison.

5. PRESERVATION

A. Preservation Obligations: The parties will take reasonable steps to preserve potentially relevant documents according to the parties' obligations under the common law and the Federal Rules of Civil Procedure. Parties are required to act in good faith and may not transfer electronic Preserved Records (as defined in Section 5(B) below) to paper form or otherwise downgrade such ESI in a manner that degrades its searchability or usability. By preserving documents for the purpose of this litigation, the Parties are not conceding that such material is discoverable.

B. Preserved Records: Documents required to be preserved under this Protocol are "preserved records." Preserved records must be preserved in their native format including all available metadata identified in Appendix A to this Order, formulaic metadata associated with spreadsheets, any comments and/or track changes associated with word processing documents, and any other hidden text or content preserved.

A party need only preserve potentially relevant evidence in the Party's possession, custody or control and one copy among two or more exact duplicate electronic preserved records (with duplicates identified only by the method set forth in Section 3 of Appendix B of this Protocol) so

long as the "Duplicate Custodians" and "DupeLocation" metadata, as defined in Appendix A of this Protocol for all custodians of the deduplicated record are preserved.

Preserved records include documents created or maintained during the period from March 2017 through 30 days following entry of this Protocol (the "preservation period") and any software, applications, systems, or other tools required to render ESI from legacy systems or storage locations reasonably accessible.

6. SEARCH AND COLLECTION

A. De-NISTing. The Parties agree that each party will use its best efforts to exclude from collection and production those files with hash values identical to those included in the National Software Reference Library reference data set (RDS Hash), a sub-project of the National Institute of Standards and Technology ("NIST"), of known traceable system and application files by using a commercially reasonable hash identification process. This process is commonly referred to as "De-NISTing." Any other system files that a Producing Party proposes to exclude from collection and/or search shall be disclosed to the requesting Party for its consideration.

B. Custodians and Data Sources.

The parties shall meet and confer concerning custodial and non-custodial data sources most likely to contain relevant information and attempt to reach agreement on custodians as well as custodial and non-custodial data sources that will be preserved and searched for responsive documents consistent with Case Management Order No. 2, Sec. 7, page 6. Within 20 days of commencement of the Discovery Process Meet and Confer Period set forth in CMO No. 2, the parties will exchange the information listed in items (i) through (iv) below, as applicable in preparation for or during meet and confers regarding Agreed Custodians and Agreed Data Sources. Each party shall amend or supplement its responses to items (i) through (iv) if necessary to make

the disclosures complete. The Parties agree and understand that exchange of this information may occur iteratively as the Parties meet and confer and that such information can and should be shared on a rolling basis.

(i) The Producing Party's most likely custodians of relevant electronic materials, including each person's job title, a brief description of such person's responsibilities (including dates employed by the Producing Party) as well as additional information or documents necessary for the Receiving Party to meaningfully evaluate the sufficiency of other party's list of custodians. By way of example only, such information may be organizational charts, and employee directories, or their equivalents;

(ii) A list of the Producing Party's custodial and non-custodial devices and electronic communications and/or ESI storage system(s), regardless of whether such systems are still in use, that may house any potentially relevant data and a general description of each system, including the nature, scope, character, organization, and formats employed in each system, and information sufficient to identify substantial changes in these systems during the relevant time period;

(iii) A copy of a written rendition of the party's document retention policies and practices, including those pertaining to any electronic communications and/or ESI storage system(s) that may house any potentially relevant data that were operative as of January 1, 2019 to the present, or, at the party's option, a general description of the policies and practices; and

(iv) A description of any problems reasonably anticipated to arise in connection with production of ESI.

The parties shall meet and confer to reach agreement on Custodians whose documents shall be collected, searched, and produced ("Agreed Custodians") and the Custodial and Non-Custodial Document Sources from which documents will be collected and searched ("Agreed Data Sources"). The Parties shall memorialize their agreements in writing.

The parties acknowledge that new claims or defenses, rulings on dispositive motions, further inquiry, and newly acquired information may impact the custodians and data sources to be searched. If that occurs, the parties will meet and confer concerning whether to add or eliminate key data sources or custodians from their agreed lists.

C. Inaccessible ESI.

The parties will meet and confer and use best efforts to agree on data sources that are not reasonably accessible because of undue burden or cost pursuant to Fed. R. Civ. P. 26(b)(2)(B). The parties will meet and confer concerning the following information:

(i) The general nature of the information the party contends is not reasonably accessible (e.g., correspondence, financial records, etc.);

(ii) The reason(s) why the information is considered inaccessible or has only limited accessibility;

(iii) Information sufficient to identify the type of backup and disaster recovery media used and the number of backup tapes involved, the identity and version of the applicable legacy software or systems, and when such software or systems achieved legacy status by the party;

(iv) Proposed capture and retrieval protocol necessary for identification and recovery of the information deemed not reasonably accessible.

If the Receiving Party intends to seek discovery of ESI that the Producing Party contends is not reasonably accessible, the parties shall promptly meet and confer in an attempt to reach an agreement as to how they will proceed with the retrieval and production.

D. Search Methodology for Structured Data.

If ESI responsive to discovery requests is located in Structured Data sources (such as ESI contained in, but not limited to, enterprise databases or database management systems such as Oracle, SQL server, DB2, or comprehensive electronic accounting systems), the parties shall meet and confer regarding the production format, content, and scope of structured data to be produced. The Producing Party shall make such disclosures necessary to ensure the meet and confer discussions are meaningful, including the nature of the data contained in structured data sources the types of information stored in the database(s), the types of reports that can be or were generated from or for the data and how and whether specialized queries may be run, whether there are existing and reasonably available reports that include the information, if the nature of the data stored in a field is not clear from the name of the field, a description of all potentially responsive fields, and whether the Receiving Party will need any additional information. To the extent information is not available in its native form, Parties agree to meet and confer regarding what format would be reasonably usable. The Producing Party shall provide a description of each field produced, as well as database schema, codes, values, and abbreviations used to populate a data field, and other information necessary for the requesting party to understand the nature and function of the structured data source and the meaning of structured data produced.

E. Search Methodology for Unstructured Data.

Upon agreement on the Agreed Custodians and Agreed Data Sources during the Discovery Process meet and confers, a Producing Party will disclose to the Requesting Party the electronic

search methodology(ies) it intends to employ, whether and, if so, how those methodologies will differ by Custodian, document source, document type, or document request, and whether, and if so, how, it intends to filter out certain documents from those collected prior to searching (the "search technology or process"). If the parties bring disputes to the Court regarding some but not all custodians or data sources, the foregoing search methodology disclosures will nevertheless be made related to the undisputed custodians or data sources despite, and before rulings on, any disputes pending before the Court. The parties will cooperate in good faith regarding the disclosure and formulation of appropriate search terms, search methodologies, protocols, search software capabilities in advance of any electronic search for responsive documents as well as methods to validate the effectiveness of the search methodologies.

Following the disclosure of such search technology or process by Producing Party to the Receiving Party, the parties shall attempt to resolve any objections to the use of such technology or process before seeking relief from the Court. Nothing in this Protocol shall require a party to use certain search, culling or prioritization methodologies.

Although this Protocol envisions the use of electronic search methodologies to locate potentially responsive documents, where a party knows or reasonably should know that information is relevant and responsive to a document request, it shall not withhold that document from production on the basis that it was not identified or captured by the electronic search methodology.

1. Search Terms. If the Producing Party elects to use search terms to identify responsive documents, the parties agree to collaborate and cooperate in good faith and to attempt to exhaust all efforts to reach an agreement on a set or sets of search terms that will be used to search for responsive documents and validation methodologies of the same, as prescribed below:

a) After inquiry of custodians and others with knowledge regarding terms likely to capture responsive documents and within a reasonable period of time following disclosure of a Party's proposed search technology or process (to be agreed upon by the Parties), the Producing Party shall develop and share a list of proposed search terms. At the same time, if the Producing Party is using a tool other than Relativity, the Producing Party shall also provide information regarding the syntax and search capabilities and limitations of its ESI tool with respect to search terms (e.g., ordering of search terms and parentheses, what wildcard options are available, stemming capabilities, noise words, etc.) sufficient for the Receiving Party to understand the search terms proposed and to offer counterproposals to them.

b) The Receiving Party shall provide any revisions or additions to the Producing Party's proposed search terms ("counterproposals"), if any, within a reasonable time (to be agreed upon between the parties) after receipt of proposed search terms from the Producing Party.

c) After receipt of the first set of revisions or additions from the Receiving Party (as set forth above), and within a reasonable period after receipt of the counterproposals (to be agreed upon by the Parties) or receipt of any revisions in subsequent comments rounds, if applicable, the Producing Party will identify which of the Receiving Party's proposed search terms the Producing Party agrees to use, and which terms it rejects, and shall propose alternatives to counterproposals to which it objects. If a Producing Party objects to a term on grounds of burden, the Receiving Party may request, and the Producing Party shall provide, a Search Terms Hit Report ("STHR Report") as described in subsection (f) below. Upon

request, the Producing Party shall confer with the requesting party concerning the Producing Party's reasoning behind why particular search terms are not acceptable and possible alternative search terms.

d) The parties shall reasonably repeat the process in subsection (b) and (c) in good faith to come to a final agreement regarding search terms. The parties will meet and confer regarding disputed search terms and STHR Reports as needed. The Parties shall meet and confer to attempt to develop a mutually agreeable means of evaluating the effectiveness of disputed search terms in identifying responsive documents, such as a responsiveness review of a statistically significant random sample of the documents hit by a disputed term.

e) **Search Terms Hit Report ("STHR Report"):** STHR Reports will be provided in Excel (or other format as agreed to by the producing and receiving parties). An STHR Report will: (a) exclude from the number of hits any duplicates (as defined herein) and logo image files if the Producing Party's processing tool does not, as a matter of course, generally prevent the extraction of them; (b) provide the number of documents hit by the search term, excluding family members that were not hit, with a separate count for all documents hit, plus family members that were not hit and (c) the number of unique documents hit by each search term (i.e., the number of documents hit by that search string and no others, and not previously produced in this litigation).

f) **Validation:** Once the parties have agreed to a list of search terms subject to the process outlined above, the Parties shall meet and confer to attempt to develop a mutually agreed, statistically sound methodology to determine the Recall Rate and

other measures of the effectiveness of search terms in identifying responsive documents. The parties shall meet and confer regarding whether further modification to the final set of search terms is necessary.

g) The mere fact that a document is hit or captured by the application of any agreed upon search terms does not mean that such document is necessarily responsive to any propounded discovery request or is otherwise relevant to this litigation.

2. Technology Assisted Review ("TAR") and Other Advanced Analytics.

The Parties also recognize the availability of a variety of search tools and methodologies, including but not limited to TAR. If the Producing Party intends to use TAR or similar advanced analytics, the parties agree to meet and confer in good faith to attempt to reach agreement about the technology and process that a Producing Party proposes to use to identify responsive ESI and a statistically sound methodology to determine the Recall Rate and other measures of the effectiveness of the tool and processes in identifying responsive documents. The Producing Party shall make disclosures regarding its tool and processes necessary to make the meet and confers meaningful and for the Receiving Party to negotiate on an informed basis.

If, prior to commencement of negotiations over search terms, a Producing Party intends, or is likely, to use both search terms and TAR (or similar advanced analytics), it shall notify the Receiving Party prior to commencement of search term negotiations. If a Producing Party decides to employ TAR or similar advanced analytics during, or after the conclusion of, negotiations over search terms, it shall notify the Receiving Party when that decision has been made. Under either of these circumstances, the parties shall meet and confer regarding whether use of search terms prior to application of TAR is appropriate, and if so, how, to what extent, and under what conditions search terms may be applied.

7. PRODUCTION FORMATS & PROCESSING REQUIREMENTS

The parties agree that attention to the form of production at the outset of discovery facilitates the efficient and cost-effective conduct of discovery. Appendices A-C, which are expressly incorporated into this Protocol, set forth technical and other specifications for the processing and production of documents in this litigation, absent other agreement by the parties. To the extent a Producing Party encounters problems complying with the requirements in Appendices B & C, the parties agree to meet and confer. For documents that cannot be produced in a useable format in either TIFF or native format, the parties will meet and confer on the appropriate format.

If reasonably available and accessible, each of the metadata and coding fields set forth in Appendix A that can reasonably be extracted from a document by the Producing Party's processing software shall be preserved and produced in accord with the requirements of Appendix B.

8. DEDUPLICATION

The parties agree that documents must be globally deduplicated at the family-group level as required in Appendix B, and the Producing Party will identify the additional custodians of the deduplicated document and the file paths/directories in which those duplicate documents were found in the "Duplicate Custodians" and DupeLocations metadata fields identified in Appendix A.

9. PHASING OF PRODUCTIONS

When a party propounds discovery requests pursuant to Fed. R. Civ. P. 34, the parties agree to meet and confer to discuss potential phasing of the production of documents, if appropriate, by expediting productions based on selected sources, custodians, or other criteria.

10. DOCUMENTS PROTECTED FROM DISCOVERY

Unless otherwise provided in this Protocol, any document falling within the scope of any request for production or subpoena that is withheld on the basis of a claim of attorney-client privilege, work product, or any other legally cognizable claim of privilege or immunity from discovery will be identified by the Producing Party on a privilege log, which the Producing Party shall produce in native Excel format.

11. BEST EFFORTS COMPLIANCE AND DISPUTES

The parties agree to use their best efforts to comply with and resolve any differences concerning compliance with any provisions of this Protocol. If a Producing Party cannot comply in a particular circumstance with this Protocol, such party shall promptly inform the Receiving Party in writing why compliance with the Protocol is not reasonable or feasible. No party may seek relief from the Court concerning compliance or non-compliance with the Protocol until it has met and conferred with the other party in a good faith effort to resolve or narrow the area of disagreement.

12. MODIFICATION

This Stipulated Order and incorporated Appendices A-C may be modified by a Stipulated Order of the parties or by the Court for good cause shown.

IT IS SO STIPULATED, through Counsel of Record.

/s/ Gary F. Lynch
Gary F. Lynch
CARLSON LYNCH
1133 Penn Avenue, 5th Floor
Pittsburgh, PA 15222

/s/ Jeannine M. Kenney
Jeannine M. Kenney
HAUSFELD LLP
325 Chestnut St #900
Philadelphia, PA 19106

/s/ Christian Levis
Christian Levis
LOWEY DANNENBERG, P.C.
One Tower Bridge
100 Front Street, Suite 520
West Conshohocken, PA 19428

Interim Co-Lead Class Counsel for Financial Institution Plaintiffs

/s/ Mindee J. Reuben
Mindee J. Reuben
LITE DEPALMA GREENBERG, LLC
1835 Market Street, Suite 2700
Philadelphia, PA 19103

Interim Liaison Counsel for Financial Institution Plaintiffs

/s/ Roberta D. Liebenberg
Roberta D. Liebenberg
FINE, KAPLAN AND BLACK, R.P.C.
One South Broad St., 23rd Floor
Philadelphia, PA 19107

/s/ Sherrie R. Savett
Sherrie R. Savett
BERGER MONTAGUE, PC
1818 Market Street, Suite 3600
Philadelphia, PA 19103

/s/ Benjamin F. Johns
Benjamin F. Johns
CHIMICLES SCHWARTZ KRINER
& DONALDSON-SMITH LLP
361 Lancaster Avenue
Haverford, PA 19041

/s/ Linda P. Nussbaum
Linda P. Nussbaum
NUSSBAUM LAW GROUP, P.C.
1211 Avenue of the Americas, 40th Fl.
New York, NY 10036-8718

Interim Co-Lead Class Counsel in the Consumer Track

/s/ Donald E. Haviland, Jr.
Donald E. Haviland, Jr.
HAVILAND HUGHES
201 South Maple Way, Suite 110
Ambler, PA 19002

Lead Counsel for Employee Track Plaintiffs

/s/ Gregory T. Parks
Gregory T. Parks
Ezra D. Church
Kristin M. Hadgis
Terese Schireson
MORGAN, LEWIS & BOCKIUS LLP
1701 Market Street
Philadelphia, PA 19103

Attorney for Defendants Wawa, Inc. and Wild Goose Holding, Co., Inc.

PURSUANT TO STIPULATION, IT IS SO ORDERED.

DATED: August 20, 2020

Hon. Gene E.K. Pratter

APPENDIX A:
ESI METADATA AND CODING FIELDS

Field Name	Field Description	Populated For	Example Values
Subject/E-Subject	Subject line extracted from an e-mail or calendar item.	E-mails Calendar	Text of the subject line
To	All recipients that were included on the "To" field extracted from the e-mail.	E-mails	John.Doe@e-mail.com
From	From or sender field extracted from the e-mail.	E-mails	Jane.Doe@e-mail.com
CC	All recipients that were included on the "CC" line extracted from the e-mail.	E-mails	Bill.Black@email.com
BCC	All recipients that were included on the "BCC" line extracted from the e-mail.	E-mails	ceo-gs@email.com
BegDoc	Bates number of the first page of the document.	All	Prefix-0000000001
EndDoc	Bates number of the last page of the document.	All	Prefix-0000000002
BegAttach	Bates number of the first page of the first document of the document family.	All	Prefix-0000000001
EndAttach	Bates number of the last page of the last document of the document family.	All	Prefix-0000000004
PageCount	Number of printed pages in the document.	All	Integer
Confidential	Confidentiality designation, if any, of the document assigned under any Protective Order in the Action. Field should be populated with specific designation; if NO designation, field should be left blank.	All	Confidential Highly Confidential
Custodian	Name of the custodian who possessed the document, in format LastName, FirstName. In the case of deduplicated documents,	All	Doe, John

	the natural person or other source that had possession of the production copy is the Custodian.		
Duplicate Custodian	Names of all other custodians who possessed a deduplicated document, in format LastName, FirstName.	All	Doe, Jane; Public, John Q.
DupeLocations	The absolute file path/directory path to the unproduced duplicate copies of files as they existed in their original environment, including the root container or directory name (e.g., folder, drive) and the files names of unproduced duplicate copies of files. For email, this includes the email folder path. The name the custodian should precede the file path	All	Doe, Jane, C:/Jane Doe/ /Documents/Exec Meetings/2020/Janes_Meeting.docx; Public, John Q., J:/Public/Board of Directors/Tech Comm/2020-06-30 Mtg Agenda.doc
DateSent	Date an e-mail was sent.	E-mails	01/01/2015
TimeSent	Time an e-mail was sent.	E-mails	12:30:00
DateModified	Date the document was last modified.	E- attachments; Electronic documents	01/01/2015
TimeModified	Time the document was last modified.	E- attachments; Electronic documents	12:30:00
DateCreated	Data extracted from the Date Created metadata field of an E-doc.	E- attachments; Electronic documents	01/01/2015
TimeCreated	Data extracted from the Time Created metadata field of an E-doc	E- attachments; Electronic documents	12:30:00
DateReceived	Date email was received.	E-mails	01/01/2015
TimeReceived	Time email was received.	E-mails	12:30:00
Importance	Indicator of priority level	E-mails Calendar	High

	set for message		
Title/E-Title	Title of document or title field extracted from the Metadata of non-email ESI	E-mails; E-attachments; Electronic documents	Smith Summary
Redaction	User-generated field that will indicate redactions or the basis for redactions if the basis is not provided on the face of the document.	ALL	Y: [reason]
FileName	File name of original native document	Electronic documents; E Attachments	2020-06-30 Mtg Agenda.doc
File Type	Application type	ALL	Word
NativeLink	Relative file path to each native file in the Natives Directory on the Production Media.	All documents produced in native format	\Natives\Document 12345.doc
Author	Data extracted from the author metadata field of an E-doc	E- attachments; Electronic documents	John Doe
HASH	MD5 or SHA-1, SHA-256 Hash Value	All	
Prod Volume	Production Volume Name	All	Defendant X Volume 1
File Path	The absolute file path to the native file as it existed in its original environment, including the root container or directory name (e.g., folder, drive). This field should be populated for both e-mail and e-files.	All	
TextPath	Relative file path to each extracted text/OCR text file on the Production Media.	All	\Text\Document_12345.txt
Source	The Producing Party.	All	

AttachCount	The number of attachments that are indicated in the document metadata.	All	integer
AttachIDs	The IDs of the documents that are attached to the produced document.	All	
Attachment Name	File name of the document attachment or embedded document; associated with parent/email which lists any and all filenames of the children/attachments delimited by a semicolon. The field would be populated for parent documents (not for the children).	E-mails; e-docs	
Parent ID	Indicates the Parent ID for an attachment or embedded document. The parent document ID field should be set for all attachments (including attachments that are emails) but should not be set for parents.	E-mail attachments; embedded documents	
ConversationIndex	Unique alphanumeric identifier for an email thread, which may be populated by the email client for each outgoing message.	E-mails	alphanumeric code generated by the email client
MessageHeader	Technical details about an email, such as email addresses, message transfer, the software used to compose it, thread index, and the email servers that it passed through on its way to the recipient and related time stamps.	E-mails	
Mail_Folder	The name of the mail folder inside the mailbox.	E-mails	
MessageID	The message ID of an email or other type of electronic message.	E-mails	Microsoft Outlook PST EntryID, the UniqueID (UNID) for Lotus Notes, equivalent value for other proprietary formats
ApptStartDate	The start date of a calendar event.	Calendar Events	01/01/2015

ApptStartTime	The start time of a calendar event.	Calendar Events	12:30:00
ApptEndDate	The end date of a calendar event.	Calendar Events	01/01/2015
ApptEndTime	The end time of a calendar event.	Calendar Events	12:30:00
FileExt	The extension of the file.	All	.xlsx
FileSize	The size of the file.	Electronic documents; E-Attachments	146.6 MB
ModifiedBy	The name of the last person who modified the document.	Electronic documents; E Attachments	Doe, John
Comments	Comments that metadata extracted from the document.	Electronic documents; E-Attachments	
ExceptionReasons	To the extent extracted in processing, populate with exception reasoning from processing software.	All	Encrypted
AttachRange	Bates Number of the first page of the parent item to the Bates Number of the last page of the last attachment "child" item.	E-mails	e.g., ABC000000001-ABC000000010
ExtractedText	Relative path to and filename of the extracted text for the produced ESI in the TEXT directory	All	
DateLastPrinted	The date the document was last printed	Electronic documents	01/01/2015
DateLastAccessed	The date the document was last accessed	Electronic documents	01/01/2015
Record Type	To indicate if a Hard Copy Document, e-doc, email, calendar appointment		Hard copy
Time Zone	The time zone the document was processed in,	Electronic documents	EST, CST, PST, etc.

** As it relates to the Custodian and Duplicate Custodian metadata fields above, the Producing Party reserves the right to produce in a single field (e.g., ALLCUSTODIAN) both the custodian of the production copy of the de-duplicated document as well as the custodians that held the duplicates.

** DATE and TIME Fields can be provided in separate fields or be combined into a single field as long as the required information is produced in the load file.

APPENDIX B

1. PRODUCTION FORMAT AND PROCESSING SPECIFICATIONS

a) **Standard Format for Documents Originating as ESI**. Unless otherwise specified in this Appendix B, the Parties shall produce the following types of responsive, non-privileged ESI in native format: spreadsheets and presentations created by readily available office suite software (e.g. Microsoft Office, G Suite, WordPerfect Office); comma or tab delimited text files; Microsoft Access files; and video and audio files.

All other responsive non-privileged documents, excluding structured data per Section 7(d) of this Protocol, shall be produced in tagged image file format ("TIFF") unless the document is not reasonably usable in TIFF format. If the document is not reasonably usable as a TIFF, the parties shall meet and confer on the production format. In connection with the Parties' agreement regarding TIFF production (except as otherwise provided herein regarding documents for which native format is required), the Parties stipulate and agree that costs of converting native format documents to TIFF format and any costs associated with processing TIFF files for production shall not be considered costs that are taxable under Federal Rule of Civil Procedure 54(d)(1) and 28 U.S.C. § 1920 and further agree that the Parties will not seek reimbursement of such costs under Rule 54(d)(1) or 28 U.S.C. § 1920, or any other provision of law.

TIFFs shall convey the same information and image as the original document, including all commenting, versioning, and formatting that is visible or can be made visible in any view of the document in its native application. All hidden text, columns, rows, and worksheets will be expanded, extracted, and rendered in the TIFF file. The Producing Party will instruct its vendor to force off Auto Date so that any ESI converted to TIFF maintains any date/time shown in a document as it was last saved by the custodian or end user, not the date of collection or processing. Any TIFFs produced shall be single-page, 300 dots per inch (DPI), Group IV TIFF files.

For each document produced in native format, the Producing Party shall also produce a corresponding placeholder slip sheet in TIFF image format ("native placeholder"), specifying that the document has been "produced in native format" and endorsed with the Bates Number and Confidentiality Designation, if applicable, which will be inserted into the image population in place of the native file. When the native file is produced, the Producing Party shall preserve the integrity of the electronic document's contents, i.e., its original formatting and metadata, and shall expand or reveal all otherwise hidden text, rows, cells, or columns. A "NativeLink" metadata entry for each document being produced in native format should be included in the data load file indicating the relative file path to each native file on the Production Media. Native Files should be produced with extracted text and applicable metadata fields as set forth in Sections 1(c) and 1(l), and Appendix A.

The file name assigned to ESI for production purposes shall be: (i) for documents produced in native format, the Bates number that is assigned to that specific record in the production per Section (b) below; (ii) for TIFF images, the Bates number of the first image page of the document (i.e., BegDoc). The original file name will be produced in the file metadata per Appendix A.

b) **Bates Numbering:** All images must be assigned a Bates number that shall always: (1) be unique across the entire document production, (2) maintain a constant length (zero/0-padded) across the entire production, (3) contain no embedded spaces or special characters other than underscores and/or hyphens, (4) be sequential across a family, and, in the case of ESI that may be produced as TIFFs, sequential within a given document (5) be of a consistent font type and size. If one or more Bates numbers is skipped in a production, the Producing Party will so note in a cover letter accompanying the production or in a privilege log.

c) **Extracted Text/OCR Text Files:** For each electronic document, whether produced in native format or as a TIFF image, an extracted text file (.txt) should be provided at the document level and in the Text directory. For TIFFs, the file name of each extracted text file should be identical to that of the Bates number of the first image page of its corresponding document, followed by .txt. For native documents, the file name of each extracted text file should be identical to that of the Bates number assigned to the corresponding native document. File names should not contain any special characters or embedded spaces. Regardless of whether the document is produced in native or TIFF format, the text should be extracted directly from the native file. However, if a document is a hard copy document or a native file that is TIFF'd and redacted (per Section 1(m) below), Optical Character Recognition may be applied to the document (following scanning from hard copy or conversion to tiff) and the resulting OCR text will suffice in place of extracted text. Extracted Text/OCR Text files should be provided in a "Text" folder on the provided media.

d) **Parent-Child Relationships:** Family documents are to be produced contemporaneously and sequentially, with children immediately after the parent document. If any member of a family is responsive, the entire family shall be produced. Hardcopy documents attached or appended to a hardcopy document, are to be produced contemporaneously and sequentially immediately after the parent document. Parent-child relationships within a document family (the association between an attachment and its parent document) shall be preserved. Each document shall be produced with the Bates number for the first and last page of that document in the "BegDoc" and "EndDoc" fields of the data load file and with the "BegAttach" and "EndAttach" fields listing the Bates number for the first and last page in the document family. The

parties agree to meet and confer to address situations in which a party believes that parent-child relationships have not been adequately preserved.

e) **Track Changes and Comments & Hidden Content.** To the extent that a document contains tracked changes or comments, speaker notes, or hidden text, and may be produced as a TIFF under this protocol, the document should be imaged showing tracked changes and comments, speaker notes, and any other hidden text. The extracted text shall include all non-privileged or protected text from tracked changes and comments, speaker notes, and any other hidden text.

f) **Encrypted Data.** To the extent documents were encrypted, the Producing Party shall make reasonable efforts to decrypt the documents prior to electronic search, review, and production. If there are potentially responsive documents that the Producing Party has been unable, after reasonable efforts, to decrypt, the parties shall meet and confer.

g) **Password Protected Files.** To the extent documents were password protected at the time of collection, prior to electronic search, review, and production, the Producing Party shall take reasonable efforts to remove the password protection. If there are potentially responsive password protected documents, the password for which the Producing Party has been unable, after reasonable efforts, to locate, the parties shall meet and confer. If a party believes that removing password protection is unduly burdensome, the Parties shall meet and confer.

h) **Embedded Objects:** Objects embedded in documents shall, when possible, be extracted and produced as separate documents. The Producing Party will make reasonable and good faith efforts to treat such extracted embedded documents like attachments to the document (with the exception of image logo files in emails which shall not be extracted or produced). Embedded documents will be produced following the parent with a continuous Bates range.

i) **Compressed and Container Files.** Compressed and container file types (e.g., .ZIP, .RAR, etc.) shall be decompressed to ensure that a compressed file within a compressed file are decompressed into the lowest possible compression to ensure all files are searched and produced as separate records. Emails and their attachments shall be extracted separately and shall be produced with the family relationship intact via the BegAttach and EndAttach fields identified in Appendix A.

j) **Hyperlinked Files**. Hyperlinked files must be produced as child documents to the parent containing the hyperlink.

k) **Load Files.** Each production of ESI and Documents shall be accompanied by Concordance or comma delimited load files (.dat and .opt). The .dat file shall contain all fielded data for the production, including the beginning bates number for each record, the relative path to the text files for each record, and, for native files, the relative path to the native file for each record. The .dat file shall also provide all metadata fields identified in Appendix A. The .opt files provide image cross reference files for all TIFF images (including native placeholders and placeholder docs for family members withheld for privilege or work-product protection) and shall identify document breaks for an image. The parties shall adhere to the production delivery requirements for these load files set forth in Appendix C.

l) **Metadata Fields:** If reasonably available and accessible, each of the metadata and coding fields set forth in Appendix A that can be extracted from a document shall be produced for that document. The parties are not obligated to populate manually any of the fields in Appendix A if such fields cannot be extracted from a document, with the exception of the following: (1) BegDoc, (2) EndDoc, (3) BegAttach (if applicable), (4) EndAttach (if applicable), (5) Custodian, (6) NativeLink fields, (7) TextPath, (8) Redaction, (9) Confidentiality, and (10)

Prod Volume (which must be populated by the party or the party's vendor). With respect to the Custodian metadata field, if the document was found in the possession of a natural person (or on data sources used by that natural person), that person should be identified in the Custodian field (whether or not that person was an agreed custodian); if the document was found in the possession of a department, group, entity, or other common facility (e.g., office, file room, archive, network storage, file share, back-up, hard drive, etc.), the Custodian field should identify that source. A Producing Party shall use a uniform description of a particular custodian across productions.

m) **Redacted or Withheld Documents:** The Producing Party can redact documents for privilege, work-product, or other legally recognized privilege, and the following sensitive personally identifiable information: social security numbers or tax id numbers and non-relevant financial account numbers. If, during the course of discovery, a Party identifies a specific document or specific type of information that a Party has a reasonable basis for redacting, before such redactions are made, the Parties will meet and confer regarding the nature of the information sought to be redacted, the reason such redaction is desired, and whether the parties agree the redaction(s) are appropriate. Documents produced with redactions shall identify in some manner (such as through highlighting in black or through the use of redaction boxes) the location and extent of redacted information. A metadata field shall indicate that the document contains redactions and the basis for redaction (e.g. "A/C Privilege") shall be provided on the face of the document or in a metadata field. The redacted text shall not be included in a redacted document's text file. Where a responsive document contains both redacted and nonredacted content, the Producing Party shall produce the remainder of the non-redacted portions of the document unless otherwise agreed to by the Parties. Redacted versions of documents, whether required to be produced in native format or TIFF format, may be produced as TIFF image files and OCR text in

lieu of extracted text. If the document would otherwise have been produced in native format under this Protocol, all other requirements for TIFF images in Sections 1(a), (c) and 1(e) must be met. If a native document that requires redaction is difficult to convert to TIFF or is not reasonably usable when converted to TIFF format, the Producing Party can redact in native format.

When one or more but fewer than all document family members are withheld, a Bates-stamped, TIFF-image placeholder will be produced in lieu of the withheld family document and shall be in the same position in the family as if the document had been produced.

n) **Hard Copy Documents:** Responsive paper documents shall be converted to single-page TIFF files and produced following the specifications set forth for TIFFs in Subsections 1(a), (c), and (e) above, including the production of OCR text generated to make such documents searchable. Each page of a hardcopy document shall be scanned into an image. In scanning all Hard-Copy Documents, Hard-Copy Documents shall be physically unitized and can be scanned as they existed in the ordinary course, provided however, if producing in such manner renders the document not reasonably usable, the parties shall meet and confer regarding an alternative unitization. In the case of an organized compilation of separate documents (for example, a binder containing several separate documents behind numbered tabs), each of the Hard-Copy Documents should be separately scanned, but the relationship among the documents in the compilation should be reflected in the proper coding of the beginning and ending documents and attachment fields. The Parties will make their best efforts to unitize the documents correctly. Producing Hard-Copy Documents as provided herein does not change their character from Hard-Copy Documents into ESI. For Hard-Copy Documents, the Parties need only populate the following metadata fields: “BegDoc,” “EndDoc,” “Prod Volume,” “Custodian,” “Source,”

"Confidential," and "Redaction," "TextPath" as well as "BegAttach" and "EndAttach" fields where applicable.

o) **Non-Standard Files:** If a response to discovery requests requires the production of files in non-standard formats or unusually large files, the parties will meet and confer to determine the optimal production format.

p) **Producing Color Documents:** If an original document that may be produced as a TIFF under this Protocol contains color, then a party may request by Bates number the document be produced in color. The Producing Party will honor reasonable requests for a color image of a document, if the original document contains color necessary to understand the meaning or content of the Document, however categorical or wholesale requests, without good cause, are deemed invalid.

q) **Prior Productions/Productions from other proceedings:** Notwithstanding anything to the contrary herein, production of documents made by a Producing Party in other civil investigations, litigations, and/or administrative actions by federal (including Congressional), state, or local government entities, if relevant to the litigation may be made in the format in which they were previously produced, including any previously produced metadata, load files and accompanying text files, and same Bates numbers and confidentiality designations. This paragraph addresses the format of production only and does not impose any obligation to produce documents produced in other proceedings. The parties will meet and confer regarding any requests for production in an alternative form or with additional information (e.g., if a prior production did not include any metadata).

r) **Confidentiality Designation**. If any document or ESI (or portion thereof) produced as a TIFF image in accordance with this Protocol is designated as confidential or highly

confidential under the CMO No. 6 (ECF No. 130), the Producing Party will brand the required Confidentiality Designation in a corner of any TIFF images representing the produced item and in a consistent font type and size that does not obscure any part of the underlying image or Bates number, if possible. Native documents may include any Confidentiality designations on the native placeholder TIFF image following the Bates number.

s) **Time Zone.** Each Party's ESI should be processed using a consistent Time Zone for all data. The Party shall share the Time Zone selected for processing of its data with the other Party and populate the TimeZone metadata field.

t) **Viruses/Malware:** All Production Media should be scanned for, and free of, viruses or malware.

2. EMAIL THREAD SUPPRESSION

A Producing Party, in the first instance, may not use email thread suppression to exclude email from its production. However, the Producing Party may use email threading only for purposes of efficiency in reviewing documents and reserves their right seek a meet and confer to revisit this issue upon good cause shown, and the Parties shall thereafter meet and confer.

3. DEDUPLICATION

a) If exact duplicate documents reside within a Party's ESI data set, the Party shall produce only a single copy of a responsive document. "Exact duplicate" shall mean bit-for-bit identity of the document content and metadata with exact Hash Value matches determined by MD5, SHA-1, or SHA26 algorithms. Hard copy documents shall not be eliminated as duplicates of responsive ESI nor of other responsive hard copy documents.

b) The parties shall de-duplicate exact duplicates of ESI horizontally (i.e., across the entire collection and across custodians). Deduplication must be done only at the family

level (i.e., if there are identical child documents that have unique parents, they are not duplicates; an attachment to a document is not a duplicate of the same document maintained as a standalone document). 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 or other blind copy field, even if all remaining content in the email is identical. Exact duplicates of emails retrieved from different storage locations (i.e. desktop vs. laptop vs. smart phone) shall not be considered unique based solely on the originating storage location.

c) The custodian that produced the production copy of the duplicate document, all custodians who were in possession of the deduplicated document, and the file paths for each duplicate custodian's copy shall be identified in the "Custodian," "Duplicate Custodian," and "DupeLocations" fields of the copy of the single record that is produced.

d) Multiple custodians in the "Duplicate Custodians" field and multiple file paths in the DupeLocations field shall be separated by semicolons. If processing and production is done on a rolling basis, an updated Duplicate Custodians and DupeLocation fields with additional values shall be provided in an overlay. The Producing Party shall identify whether the overlay replaces previously produced fields for a file or supplement them.

APPENDIX C
PRODUCTION DELIVERY REQUIREMENTS

1. **Means of Production**. The Producing Party shall produce documents via secure FTP or secure file share. In the alternative, and subject to prior agreement with the Receiving Party, documents may be produced via CD, DVD, flash drive, or hard drive if (a) the size of the production exceeds the size limitations applicable to the Producing Party's secure FTP or secure file share; (b) producing the documents via secure FTP or secure file share would be unduly burdensome; or (c) if the interest of preserving the confidentiality of the information produced outweighs the speed and efficiency of producing documents via secure FTP or secure file share. To the extent possible, physical media should be write-protected before it is produced.

2. **Contents of Production.** Productions shall have the following four directories: (1) IMAGES for the document images (; (2) DATA for the .dat metadata load file and .opt image load files; (3) TEXT containing the extracted text/OCR files; and (4) NATIVES containing the original native files.

3. **Naming Convention for Production Media.** Whether produced via secure FTP, file share, or physical media, the files produced should be combined into a compressed file such as .zip, .rar, etc. The compressed file should be named so as to indicate the Producing Party, the date of the production, and the sequence of the production (e.g., "Wawa ProdVol20200401-001."). Production volume names should be consecutive (i.e., "WawaProdVol20200401-001, Wawa ProdVol20200401-002, Wawa ProdVol20200401-003"). If the production is made using physical media, the media should be labeled to include (a) text referencing that it was produced in *In re Wawa, Inc. Data Security Litigation*, Civil Action No. 19-6019; (b) the Bates Number range of the materials contained on the media; (c) the filename(s) of the compressed file(s) contained on the media such as the example included above; (d) and the production volume

number.

4. **Replacement Productions.** Any replacement production will be transmitted with a cover letter or email to identify the production as a replacement and cross-reference the BegDoc and EndDoc of the documents being replaced. Replacement productions shall include load files necessary to link the replacement file to other previously produced document family members. Where the replacement document occupies more pages than an original placeholder image, the Producing Party shall apply an appropriate numerical suffix to the Bates number of the file (e.g., ABC000123.001 to ABC000123.005. If the replacement production is being transmitted by physical media, the media shall include the phrase "Replacement Production."

5. **Specifications for Load Files.**

<u>Image Load Files</u>

- Image load (cross-reference) files should be produced in Concordance Image (Opticon) format.
- The name of the image load file should mirror the name of the delivery volume, and should have the appropriate extension (*e.g.*, ABC001.OPT).
- The volume names should be consecutive (*e.g.* ABC001, ABC002, et seq.).
- There should be one row in the load file for every TIFF image in the production.
- Every image in the delivery volume should be cross-referenced in the image load file.
- The imageID key should be named the same as the Bates number of the page.
- Load files should not span across media (*e.g.*, CDs, DVDs, hard drives, etc.), *i.e.*, a separate volume should be created for each piece of media delivered.
- Files that are the first page of a logical Document should include a "Y".
- Subsequent pages of all Documents (regular document, email, or attachment) should include a blank in the appropriate position.

- The file will be comma delimited file and consist of six fields per document as follows (FolderBreak and BoxBreak should be left empty):

  ImageID,VolumeLabel,ImageFilePath,DocumentBreak,FolderBreak,BoxBreak,PageCount

Sample Concordance Image (Opticon) Load File:

MSC000001,MSC001,D:\IMAGES\001\MSC000001.TIF,Y,,,3
MSC000002,MSC001,D:\IMAGES\001\MSC000002.TIF,,,,,
MSC000003,MSC001,D:\IMAGES\001\MSC000003.TIF,,,,,
MSC000004,MSC001,D:\IMAGES\001\MSC000004.TIF,Y,,,2
MSC000005,MSC001,D:\IMAGES\001\MSC000005.TIF,,,,,

Concordance Data Load Files:

- Data load files should be produced in Concordance .DAT format.
- The data load file should use standard Concordance delimiters:
- Comma - ¶ (ASCII 20);
- Quote – þ (ASCII 254);
- Newline - ® (ASCII174).
- The first line of the .DAT file should contain the field names arranged in the same order as the data is arranged in subsequent lines.
- All date fields should be produced in mm/dd/yyyy format, if possible.
- All attachments should sequentially follow the parent Document/email.
- Use carriage-return to indicate the start of the next record.
- The name of the data load file should mirror the name of the delivery volume, and should have a .DAT extension (*e.g.*, ABC001.DAT)
- The volume names should be consecutive (*e.g.*, ABC001, ABC002, et seq.).

  Sample Concordance .DAT Load File:

  þBegDocþþEndDocþþBegAttachþþEndAttachþþPgCountþþCustodianþ

6. **Specifications for** OCR/Extracted Text Files

- OCR or Extracted Text files shall be provided in a separate directory containing Document level text files (i.e., a single extracted text file per document).

- Document identifiers shall have alphanumeric prefixes. Document identifiers with only numbers are improper.

- Documents shall not be named with alphabetical suffixes. The text files must be named the same as the Bates number assigned to the corresponding file.