Information Request

1. For the period beginning January 1, 2013 through June 30, 2014, provide:

   a. plan level subscriber data by zip code for the period beginning June 1, 2013 through June 30, 2014, as requested in the Billing Plan Data Table (attached);

   b. the residential locations for which the Company's Internet Access Service is available and provide the number of residential subscribers by census block, technology and bandwidth as requested by the Internet Access Table (attached); and

   c. separately, for each person from whom the Company purchases Transit Service or provides Peering, provide the data requested in the Purchases of Transit Service Table, Sales of Paid Peering Table, and Settlement-Free Peering Traffic Table (attached).

2. Explain or provide documents discussing:

   a. the Company's ability to compete on price, bandwidth speed, features and promotions in the provision of Internet Access Service to consumers using DSL;

   b. whether and to what extent the competitiveness of the Internet Access Service offered by the Company is affected by increases or decreases in the bandwidth speed of that service;

   c. how the Company's Internet Access Service compares to the Internet Access Service offered by competing providers, including comparisons of the relative attractiveness to subscribers of Internet Access Services offered by competing cable providers and competing DSL providers;

   d. the downlink and uplink speeds the Company believes are required to support the average individual's and a household's internet usage and video consumption, both at present and in the next three year period; and

   e. churn, subscriber acquisition and retention costs, subscriber costs incurred in switching to another provider of Internet Access Service, reasons why subscribers disconnect or switch including but not limited to pricing, quality of service and disputes between the Company and Edge Providers, CDNs or transit service providers.

3. Explain or provide documents sufficient to show the Company's policies or procedures with respect to decisions to establish or augment interconnection capacity with any CDNs, Internet backbone services, edge providers, Internet Access Service providers, and all other persons with whom the Company may engage in Internet Traffic Exchange.

4. Provide the most recent regularly prepared network planning documents including but not limited to budgets and financial projections regarding the expansion, development, deployment, and improvement of the Company's Internet Access Service.

5. Provide examples of or documents sufficient to demonstrate examples of negotiations between the Company and another person that did not result in an agreement for the Company to acquire any rights (linear or non-linear) to distribute broadcast or non-broadcast, video programming from

that other person and, if possible, explain the Company's view as to whether such negotiations failed because of: (i) any economic or non-economic Most-Favored-Nation clause in a contract with another person; (ii) grants to distribute programming exclusively to another person; or (iii) any other limits by another person on the distribution of the programming being negotiated.

6. Explain the Company's use of data caps or usage allowances, including future plans for such data caps or usage allowances, and how such plans are or would be affected by the decisions of competing cable providers and competing DSL providers to impose data caps or usage allowances.

Definitions

The following definitions apply only to this Information and Data Request. They are not intended to set or modify precedent outside the context of this document. In this Information and Data Request, the following terms shall have the following meanings (such meanings to be equally applicable to both the singular and plural forms of the terms defined):

1. The term "Company" means RCN Telecom Services, and its Subsidiaries.

2. The term "2010 Census Block Identifier" means a concatenation of Census 2010 state FIPS code, Census 2010 county FIPS code, Census 2010 census tract code and Census 2010 tabulation block number; see the 2010 TIGER/Line Shapefiles Technical Documentation, Chapter 5, Part 5.2 at http://www.census.gov/geo/maps-data/data/pdfs/tiger/tgrshp2010/ TGRSHP10SF1CH5.pdf for more information.

3. The term "Affiliate" means, as to any Person, any other Person that, directly or indirectly, controls, is controlled by or is under common control with such Person.

4. The terms "and" and "or" have both conjunctive and disjunctive meanings.

5. The word "any" shall be construed to include the word "all," and the word "all" shall be construed to include the word "any." The word "each" shall be construed to include the word "every," and the word "every" shall be construed to include the word "each." All words used in the singular should be construed to include the plural, and all words used in the plural should be construed to include the singular.

6. The term "Autonomous System" or "AS" means a connected group of one or more IP prefixes run by one or more network operators under a single and clearly defined routing policy.

7. The term "Autonomous System Number" or "AS Number" or "ASN" means a globally unique number assigned to autonomous systems, used in both the exchange of exterior routing information (between neighboring ASes) and as an identifier of the AS itself.

8. The term "Average Recurring Service Plan Revenue per Subscriber" or "ARSPR" means the average monthly revenue received from subscribers to a Service Plan, where the revenue only includes monthly recurring subscription revenue to MVPD Service, Internet Access Service and Telephone Services tiers. It should not include fees and charges associated with taxes, facilities

fees, installation fees, activation charges, customer equipment fees, high definition fees, premium channels or any other add-on services. Credits such as gift cards, retention credits and other discounts related to one time payments should be amortized on a monthly basis over a one year period and deducted from the 0-12 month monthly ARSPR. For example, a $300 gift card would reduce 0-12 ARSPR by $25 per month.

9. The term "Bundled Services" means, and information shall be provided separately for, (a) MVPD service, Internet Access Service and Telephone Services; (b) MVPD Service and Internet Access Service; (c) MVPD Service and Telephone Services; and (d) Internet Access Service and Telephone Services, each sold as a package of services for a single price.

10. The term "Cable System" means a facility consisting of a set of closed transmission paths and associated signal generation, reception, and control equipment that is designed to provide cable service which includes Video Programming and which is provided to multiple subscribers within a community, as defined in Section 76.5 of the Commission's rules.

11. The term "Capacity" means capacity for efficient production of service, stated in the measure customarily used in the particular industry involved. When stating capacity, state each assumption used to calculate capacity, including the customary period over which capacity is measured, and the customary unit of time in which the service is provided. If you know of no customary measure for a service, state the maximum amount of the service that could be provided efficiently over a one month period given existing facilities, equipment and personnel.

12. The term "Cluster" means commonly owned or managed Cable Systems in close proximity that are operated on an integrated basis through the use of common personnel, marketing, or shared use of technical facilities.

13. The term "Content Delivery Network" or "CDN" means a distributed system of servers that cache content closer to end users.

14. The term "DBS" means direct broadcast satellite.

15. The term "Designated Market Area" or "DMA" means unique, county-based geographic areas designated by The Nielsen Company.

16. The term "discussing" when used to refer to documents means analyzing, constituting, summarizing, reporting on, considering, recommending, setting forth, or describing a subject. Documents that contain reports, studies, forecasts, analyses, plans, proposals, evaluations, recommendations, directives, procedures, policies, or guidelines regarding a subject should be treated as documents that discuss the subject. However, documents that merely mention or refer to a subject without further elaboration should not be treated as documents that discuss that subject.

17. The term "documents" means computer files and written, recorded, and graphic materials of every kind in the possession, custody, or control of the Company. The term "documents" includes without limitation, drafts of documents, copies of documents that are not identical duplicates of the originals, and copies of documents the originals of which are not in the possession, custody, or control of the Company. In addition, the term "documents" includes without limitation any amendments, side letters, appendices, or attachments. The term "computer files" includes without

limitation information stored in, or accessible through, computer or other information retrieval systems. Thus, the Company should produce documents that exist in machine-readable form, including documents stored in personal computers, portable computers, workstations, minicomputers, mainframes, servers, backup disks and tapes and archive disks and tapes, and other forms of offline storage, whether on or off the Company's premises. Electronic mail messages should also be provided, even if only available on backup or archive tapes or disks. Computer files shall be printed and produced in hard copy or produced in machine-readable form (provided that Commission staff determine prior to submission that it would be in a format that allows the Commission to use the computer files), together with instruction and all other materials necessary to use or interpret the data. Unless otherwise specified, the term "documents" excludes bills of lading, invoices, purchase orders, customs declarations, and other similar documents of a purely transactional nature and also excludes architectural plans and engineering blueprints. Where more than one identical copy of a requested document exists, the Company shall only produce one representative copy.

18. The term "Edge Provider" means entities that provide content, applications, or services over the Internet, and include but are not limited to OVDs, Voice over Internet Protocol ("VoIP") providers, and gaming companies.

19. The term "End User" means a residential or business customer (a private or public entity or an institution) that purchases a Relevant Service. With respect to Internet Access Service, the term "End User" means a Person who originates, receives, or otherwise interacts with information, data and content on the Internet.

20. The term "hurdle rate" means the required rate of return in a discounted cash flow analysis, above which investment makes sense and below which it does not. Often, this is based on a firm's cost of capital or weighted average cost of capital, plus or minus a risk premium to reflect the project's specific risk characteristics.

21. The term "identify," when used with reference to a Person or Persons, means to state his/her full legal name, current or last known business address, current or last known telephone number, current or last known organization, and position therewith. "Identify," when used with reference to a document, means to state the date, author, addressee, type of document (e.g., the types of document, as described above), a brief description of the subject matter, its present or last known location, and its custodian, who must also be identified. "Identify," when used with reference to an entity other than a Person, means to state its name, current or last known business address, and current or last known business telephone number. The term "identify" when used with reference to a census block means to geocode the census block using the 2010 Census Block Identifier.

22. The term "Interconnection" means the linking together of interoperable systems, and the linkage used to join two or more communications units, such as systems, networks, links, nodes, equipment, circuits and devices.

23. The term "Interconnection Agreement" means an agreement for the purchase of Transit Service, the sale of Transit Service, Settlement-Free Peering, Paid Peering, or equivalent agreements.

24. The term "Internet Access Service" means the provision to end users of connectivity to the Internet by any means, including, for instance, hybrid-coaxial, optical fiber or coaxial cable,

xDSL, satellite systems, fixed or mobile wireless services, ultra-high frequency microwave (sometimes referred to as "LMDS"), or multichannel multipoint distribution services ("MMDS").

25. The term "Internet Backbone Services" means services that route traffic between Internet Access Service , other Internet Backbone Services, Edge Providers, and CDNs, and, when exchanging traffic with third-party services or networks, the exchange of traffic by means of Settlement-Free Peering, Paid Peering or Transit Service.

26. The term "Internet Traffic Exchange" means interconnection between Internet Access Service, Internet Backbone Services, Edge Providers, CDNs, and other Internet networks and services, where traffic is exchanged pursuant to an interconnection agreement or arrangement.

27. The term "IP Point of Presence" means any point where an Internet Access Service or an Internet Backbone Services interconnects, offers to interconnect, or intends to offer interconnection with other Internet networks, services or Edge Providers.

28. The term "Multichannel Video Programming Distributor" or "MVPD" means an entity, including but not limited to, a cable operator, which is engaged in the business of making available for purchase, by subscribers or customers, multiple channels of Video Programming.

29. The term "MVPD Service" means the distribution of Video Programming or Online Video Programming by an MVPD to consumers, including but not limited to, distribution as part of a programming package or tier or on an individual basis, and distribution as part of TV Everywhere or Specialized Services, regardless of whether such programming is distributed inside or outside of the MVPD's footprint.

30. The term "Navigation Devices" means converter boxes, interactive communications equipment, and other equipment used by consumers to access multichannel Video Programming and other services offered over multichannel video programming systems via the use of authentication or another form of conditional access.

31. The term "Non-Broadcast Programming Network" means without limitation network Video Programming delivered directly to MVPDs, whether delivered in standard or high definition.

32. The term "Online Video Programming" means Video Programming that is intended for distribution by means of the Internet or other IP-based transmission path.

33. The term "Online Video Distributor" or "OVD" means any entity that provides Video Programming by means of the Internet or other IP-based transmission path provided by a person other than the OVD. Unless otherwise stated, an OVD does not include an MVPD inside its MVPD footprint or an MVPD to the extent it is offering Online Video Programming as a component of an MVPD subscription to customers whose homes are inside its MVPD footprint.

34. The term "Paid Peering" is peering in which one peer pays another for the exchange of traffic. Paid peering is sometimes referred to as "Non Transit" interconnection.

35. The term "Pay Per View" or "PPV" means a service that allows MVPD subscribers, for an additional fee, to order individual programs, generally live event programming.

36. The term "Peer" means a Person who is a party to a peering arrangement.

37. The term “Peering” means an interconnection arrangement between Persons, pursuant to which Internet traffic is exchanged between Persons and their customers, however, there is no Transit Service through networks to other Peers or Transit Service providers.

38. The term “Person” includes the Company, and means any individual, partnership, corporation (including a business trust), joint stock company, trust, unincorporated association, joint venture, limited liability company or other entity, or a government or any political subdivision or agency thereof.

39. The term “plans” means tentative and preliminary proposals, recommendations, or considerations, whether or not finalized or authorized, as well as those that have been adopted.

40. The term “relating to” means in the whole or in part constituting, containing, concerning, discussing, describing, analyzing, identifying, or stating.

41. The term “relevant area” means, and information shall be provided separately for,

    a. with respect to Internet Access Service, the Cable System in which the Company provides such services;

    c. with respect to Internet Backbone Services, the United States;

    d. with respect to Internet Traffic Exchange, the United States; and

    e. with respect to an OVD, the United States. and

    h. with respect to Video Programming Distribution, the United States.

42. The term “relevant service” as used herein means, and information shall be provided separately for, each service in the following categories:

    a. Content Delivery Network;

    b. Internet Access Service, and information shall be provided separately for:

        i. Subscribers; and

        ii. Edge Providers; and

    c. Internet Traffic Exchange, and information shall be provided separately for interconnection between the Company and each provider of:

        i. Internet Access Service;

        ii. Internet Backbone Services;

        iii. Edge Providers;

        iv. CDNs; and

v. other Internet networks and services;

vi. OVDs.

43. The term "Service Plan" shall mean a package of specific tiers of one or more Cable Services, sold as a package of services for a single price.

44. The term "Settlement-Free Peering" means peering where there are no payments exchanged between network providers for the exchange of traffic.

45. The term "Specialized Services" means services that share capacity with Internet Access Service over providers' last-mile facilities.

46. The term "Standalone Services" means, and information shall be provided separately for: (a) MVPD Service; (b) Internet Access Service; and (c) Telephone Services, each sold as a separate service for a single price.

47. The term "Subsidiary" as to any Person means any corporation, partnership, joint venture, limited liability company, or other entity of which shares of stock or other ownership interests having ordinary voting power to elect a majority of the board of directors or other managers of such corporation, partnership, joint venture, limited liability company or other entity are at the time owned, or the management of which is otherwise controlled, directly or indirectly, through one or more intermediaries, or both, by such Person.

48. The term "System Technology" means techniques including but not limited to encoding, frequency banding, modulation, physical connections, protocols, and security (including authentication and conditional access).

49. The term "Time-Shifted and Place-Shifted Video Programming" means Video Programming that may be viewed by consumers at the time and place of their choosing, usually via an Internet-connected device or on a VOD basis.

50. The term "Transit Service" means a service arrangement where customer pays Internet Backbone Services provider to send and receive traffic to and from destinations that can be either on or off the provider's network.

51. "United States" or "U.S." means the United States, its possessions, territories, and outlying areas.

52. The term "Video-on-Demand" or "VOD" means a service which allows subscribers to MVPD Service to view individual programs at the time of their choosing including, but not limited to, motion pictures, professional video, Broadcast Programming Networks, or Non-Broadcast Programming Networks. VOD includes without limitation both free programs and programs for which there is a charge.

53. The term "Video Programming" means programming provided by, or generally considered comparable to programming provided by, a television broadcast station or cable network, regardless

of the medium or method used for distribution, and includes but is not limited to: programming prescheduled by the programming provider (also known as scheduled programming or a linear feed); programming offered to viewers on an on-demand, point-to-point basis (also known as VOD or PPV); short programming segments (also known as clips); programming that includes multiple video sources (also known as feeds, including camera angles); programming that includes video in different qualities or formats (including high-definition, 3D and 4K); and films.

54. The term "Video Programming Distribution" means the provision of Video Programming to consumers.

Instructions

1. Submit responsive documents (including materials containing Highly Confidential or Confidential Information) in electronic form only, unless otherwise specified, as set forth in the Instructions for Submission of Electronic Documents. The Commission does not require the submission of paper copies of these documents at this time, but reserves the right to require their submission at a later time.

2. Each responsive document shall be submitted in its entirety, even if only a portion of that document is responsive to a request made herein. This means that the document shall not be edited, cut, or expunged, and shall include all appendices, tables, or other attachments, and all other documents referred to in the document or attachments. All written materials necessary to understand any document responsive to these Requests shall also be submitted.

3. Data provided in response to this Request should include a list of all parameters/assumptions on which the data are based.

Instruction for Electronic Production of Documents and Electronically Stored Information

Introduction

This document contains the standard specifications and procedures for submitting to the Federal Communications Commission electronic versions of documents in response to an Information Request.

- In many cases, it is useful for the party producing the documents, its vendor and the FCC staff to review the technical details of the production prior to submitting the response.
- Take care to ensure that all responsive data and metadata are preserved in the collection process.
- Submit two copies of the response: file one copy in the docket with the Secretary's Office and provide one copy to FCC staff or the FCC's contractor, as directed, for loading and review.

1. Categories of Documents

There are seven major categories of relevant documents: (1) email and other electronic messages (*e.g.*, instant messaging, text messages), (2) other electronic documents, (3) hard copy documents, (4) shared resources, (5) databases, (6) audio and video data, and (7) foreign-language materials. Typically, responsive documents will come from the first four categories, although databases will be used to respond to Data Requests. The general requirements for each category of document are outlined below. The attached Metadata Table of Requested Fields provides information regarding document-specific metadata and bibliographic information (identifying information).

A. <u>Email, Attachments, and Other Electronic Messages</u>

Email and other electronic messages (*e.g.*, instant messages (IMs), text messages) should be produced as image files with related searchable text, metadata and bibliographic information. Depending on how the Company's systems represent names in email messages, IMs or text messages, we may require a table of names or contact lists from custodians.

Each IM or text conversation should be produced as one document.

Email repositories, also known as email databases (*e.g.*, Outlook .PST, Lotus .NSF), can contain a variety of items, including messages, calendars, contacts, tasks, etc. For purposes of production, responsive items should include the "Email", "Other ESI", or "Calendar Items" metadata/database fields outlined in the Metadata Table, including but not limited to all parent items (mail, calendar, contacts, tasks, notes, etc.) and child files (attachments of files to email or other items), with the parent/child relationship preserved. Similar items found and collected outside an email repository (*e.g.*, .MSG, .EML, .HTM, .MHT) should be produced in the same manner.

Graphic objects embedded in emails such as logos, letterheads, or backgrounds should remain as part of the email and not be extracted as separate attached documents.

Pay special attention to the PARENTID, ATTACHMENTIDS, and FAMILYRANGE fields, which are used to track email families. While the example below relates to email families, all attachment relationships for all responsive documents are to be produced in this format.

EXAMPLE: Consider ABC-JD-00000001 a 10-page parent email, with records ABC-JD-00000011 to ABC- JD-00000015, ABC-JD-00000016 to ABC-JD-00000020, and ABC-JD-00000021 to ABC-JD-00000025 as its attachments. Fields should be populated exactly as follows using the semicolon as the multi-entry delimiter for ATTACHMENTIDS:

BEGDOC#	**ENDDOC#**	**PARENTID**	**ATTACHMENTIDS**	**FAMILYRANGE**
ABC-JD-00000001	ABC-JD-00000010		ABC-JD-00000011; ABC-JD-00000016; ABC-JD-00000021	ABC-JD-00000001 – ABC-JD-00000025
ABC-JD-00000011	ABC-JD-00000015	ABC-JD-00000001		ABC-JD-00000001 – ABC-JD-00000025
ABC-JD-00000016	ABC-JD-00000020	ABC-JD-00000001		ABC-JD-00000001 – ABC-JD-00000025
ABC-JD-00000021	ABC-JD-00000025	ABC-JD-00000001		ABC-JD-00000001 – ABC-JD-00000025

B. Electronic Documents

Electronic documents include word-processing documents, spreadsheets, presentations, and all other electronic documents not specifically discussed elsewhere. Production of these items should include image files with related searchable text, metadata, and bibliographic information. All passwords and encryption must be removed from electronic documents prior to production. Note that the following apply to both attachments of files to email or other items and loose native files:

1. *Spreadsheets*: Spreadsheets should be produced in native format (*e.g.*, as .XLSX files), with searchable text for the entire document, metadata, and bibliographic information. Provide only a single image of the first page of the spreadsheet or provide a single placeholder image. The placeholder image must contain at a minimum the BEGDOC#, FILENAME, and FILEPATH. The Identification range for a spreadsheet should be a single number (*e.g.*, ABC-JD-00000001 – ABC- JD-

00000001). The linked native file name should match the BEGDOC#/DOCID with the appropriate file extension.

2. *Presentations*: Presentations should be produced in full slide image format along with speaker notes (which should follow the full images of the slides) with related searchable text, metadata, bibliographic information and linked native file. Presentations should also be produced in native format (*e.g.*, as .PPT files). The linked native file name should match the BEGDOC#/DOCID with the appropriate file extension.

3. *Hidden Text*. All hidden text (*e.g.*, track changes, hidden columns, hidden slides, mark-ups, notes) shall be expanded and rendered in the extracted text file. For files that cannot be expanded linked native files shall be produced with the image files.

4. *Embedded Files.* All embedded objects (*e.g.*, graphical files, Word documents, Excel spreadsheets, .wav files) that are found within a file shall be produced so as to maintain the integrity of the source document as a single document. For purposes of production the embedded files shall remain embedded as part of the original source document. Hyperlinked files must be produced as separate, attached documents upon request. Any objects that cannot be rendered to images and extracted text (*e.g.*, .wav, .avi files) must be produced as separate extracted files with linked native files and placeholder images and be treated as attachments to the original file.

5. *Image-Only Files.* All image-only files (non-searchable .PDFs, multi-page TIFFs, Snipping Tool screenshots, etc., as well as all other images that contain text) shall be produced with associated OCR text, metadata, and bibliographic information.

6. *Proprietary File Types and Non-PC or Non-Windows Based Systems*. Proprietary file types, such as those generated by financial or graphic design software, should be discussed with Commission staff in advance of production to determine the optimal format of production. Also, file types from non-PC or non-Windows based systems (*e.g.,* Apple, UNIX, LINUX systems), should be discussed with Commission staff in advance of production to determine the optimal format of production.

7. *Archive File Types.* Archive file types (*e.g.,* .zip, .rar) must be uncompressed for processing. Each file contained within an archive file should be produced as a child to the parent archive file. If the archive file is itself an attachment, that parent/child relationship must also be preserved.

8. *Processing Errors.* The text of the placeholder image should also be contained in the searchable text of the document in the event of uncorrectable processing errors.

C. <u>Hard-Copy (or Paper) Documents</u>

Hard-copy documents are to be produced as black-and-white image files, except where noted below, with related searchable OCR text and bibliographic information. Special attention should be paid to ensure that hard-copy documents are produced as they are

kept, reflecting attachment relationships between documents and information about the file folders within which each document is found. In addition, multi-page documents must be produced as single documents (*i.e.*, properly unitized) and not as several single-page documents. Where color is required to interpret the document, such as hard copy photos, and certain charts, that image must be produced in color. These color images are to be produced as .jpg format. Hard-copy photographs should be produced as color .jpg format files, if originally in color, or black-and-white .tif files if originally in black-and-white.

D. Shared Resources

Shared Resources should be produced as separate custodians if responsive custodians have access to them or if they contain responsive documents. The name of the group having access would be used as the custodian name, *i.e.* Marketing Execs or Accounting Dept. The Company will separately provide a brief description of each shared resource that includes a list of the custodians who have access to that shared resource.

E. Database Productions

Production of enterprise databases are not addressed in these specifications and must be discussed with the appropriate government legal and technical staff to determine the optimal production format; these will usually fall outside the scope of an image-based production. Care must be taken to ensure that all responsive databases and their metadata are preserved.

F. Audio/Video Data

These specifications do not address the production of audio/video data. Care must be taken to ensure that all responsive audio/video data and their metadata are preserved. These data types may be stored in audio or video recordings, voicemail text messaging, and related/similar technologies. However, such data, logs, metadata, or other files related thereto, as well as other less common but similar data types, should only be produced after consultation with and consent of Commission staff as to the format for the production of such data.

G. Foreign-Language Materials

Foreign language materials should be produced after consultation with Commission staff.

2. De-duplication

Before doing any de-duplication, provide Commission staff with a written description of the method used to de-duplicate (including which elements are compared and what hash codes are used), and what is considered a duplicate. Then confirm that your approach is acceptable to the Commission. The Commission does not allow de-duplication of hard-copy documents, or that of "loose" electronic documents (*e.g.,* presentation slides located on the custodian's C: drive) against email attachment versions of those same documents. The integrity of any produced email and any related "document family" must be maintained except as limited by any claim of privilege. Email attachments may not be deduplicated against attachments in other email

families. De-duplication should occur both vertically within each custodian and horizontally across custodians. Vertical de-duplication is crucial when a production includes electronic documents from back-up tapes. Horizontal de-duplication must be done in a way that preserves (and produces) information on blind copy (Bcc) recipients of emails and other custodians whose files contain the duplicates that will be eliminated from the production as well as original filepath and mailbox folder information.

Custodian Append File. A Custodian Append file is to be produced when de-duplicating ACROSS custodians (*i.e.,* horizontal de-duplication) and data is produced on a rolling basis. The file must be provided on an incremental basis starting with the second submission; as more custodians are discovered for previously produced documents, this file is updated with only the new custodian information. The Custodian Append File is a four-field delimited file consisting of the DOCIDs of the previously delivered document, the new custodian names, the Filepath(s) prepended with Custodian Name (multi-entry), and the FolderLabel(s) prepended with Custodian name (multi- entry) for the duplicates of those records that would otherwise have been produced in the subsequent (new) submissions.

These specifications do not allow for near de-duplication or email threading. These forms of de-duplication must be discussed separately with Commission staff and consent obtained prior to the use of such techniques for production.

3. Document Numbering

Documents must be uniquely and sequentially numbered across the entire production, with an endorsement burned into each image. Each number shall be of a consistent length, include leading zeros in the number, and unique for each produced page. Numbers should contain no more than three segments connected by a hyphen. The leading segment must be the Company identifier, a middle segment identifying the custodian, and a sequential page counter with connecting hyphens (*e.g.*, ABCCO-CEO-00000001). The number of digits in the numeric portion of the format should not change in subsequent productions, nor should spaces, hyphens, or other separators be added or deleted. Under no circumstances should Identification numbers contain embedded spaces, slashes (/), backslashes (\), carats (^), underscores (_), ampersands (&), hash marks (#), plus signs (+), percent signs (%), dollar signs ($), exclamation marks (!), pipes (|), any character used as a delimiter in the metadata load files, or any character not allowed in Windows file-naming convention (,\ / : * ? “ < > | ~ @ ^).

4. Privilege Designations

Documents redacted pursuant to any claim of privilege will be designated “Redacted” in the EPROPERTIES field as described in the Metadata Table. Appropriately redacted searchable text (OCR of the redacted images is acceptable), metadata, and bibliographic information must also be provided.

All documents that are part of a document family that includes a document withheld pursuant to any claim of privilege will be designated “Family Member of Privileged Doc” in the EPROPERTIES field as described in the Metadata Fields table for all other documents in its family. Placeholder images with BEGDOC#, FILENAME, FILEPATH and reason withheld

(*e.g.,* "Privileged") should be provided in place of the document images of the privileged document.

5. Sample

Before beginning production, a sample production covering files of all types, including emails with attachments, loose files including spreadsheets and presentations, redacted documents, etc., should be provided, as directed by Commission staff. The sample size should be between 500 to 1000 records to be large enough to be representative and small enough to review quickly. The Commission will take a few business days to evaluate the sample and provide feedback. If there are any problems, corrected samples will need to be resubmitted until the Commission can confirm the problems are resolved.

6. Load File Set/Volume Configuration

Each production must have a unique MEDIAID name associated with it. This MEDIAID name must also appear on the physical label. The MEDIAID naming scheme should start with a 2 or 3 letter prefix identifying the Company followed by a 3-digit counter (*e.g.,* ABC001). Each separate volume delivered on that media must also have a separate VOLUMENAME associated with it. On the root of the media, the top level folder(s) must be named for the volume(s). VOLUMENAME(s) should also be indicated on the physical label of the media. The volume naming scheme should be based on the MEDIAID name followed by a hyphen, followed by a 3-digit counter (*e.g.*, ABC001-001). Load file volumes should be as large as practical but not contain more than 100,000 records each. The VOLUMENAME should increase sequentially across all productions on the same MEDIAID.

Under the VOLUMENAME folder, the production should be organized in 4 subfolders:

1. DOCLINK (contains linked native files , may contain subfolders, with no more than 5,000 files per folder)

2. IMAGES (may contain subfolders, with no more than 5,000 image files per folder)

3. FULLTEXT (may contain subfolders, with no more than 5,000 document-level text files per folder)

4. LOADFILES (should contain the metadata, DII, OPT, LST, and custodian append files)

7. Deliverables

A submission index spreadsheet must be delivered with each submission and should provide statistical information about the volume(s) and media produced. Provide this in hard copy format and electronically on the deliverable media. A sample is included in this PDF.

It is expected that all productions will conform to the structure of the final, approved sample production. Any changes to production procedures that alter output format will require prior submission of another sample production as outlined in section 5. Sample.

The Commission accepts electronic productions loaded onto hard drives, CD-ROMs, or DVD-ROMs; however, production on hard drives minimizes costs and delay and is preferable. Where the size of the production exceeds the capacity of a single DVD-ROM, hard drives should be used as the delivery medium. For each piece of media a unique identifier (MEDIAID) must be provided and must be physically visible *on the exterior* of the physical item.

If the media is encrypted, supply the tool for decryption on or with the same media, as well as instructions for decryption. Provide the password separately.

All documents produced in electronic format shall be scanned for, and free of, viruses. The Commission will return any infected media for replacement, which may affect the timing of the Company's compliance with this Information Request.

The Commission does not accept load file productions via email or those that are posted on download sites (*e.g.,* FTP, secure server).

The Commission recognizes that occasionally unforeseen issues will arise that require replacing documents or data from a previously delivered production with new documents or corrected data. Substantive corrections may require the reproduction of the entire production volume. The production format for all corrections must be agreed upon prior to any submission of corrections. Any productions that have been created but not delivered when the need for corrections is detected must be corrected prior to delivery. Each replacement or corrected production must be named based on the production volume that is being replaced or corrected. For example, if a corrected metadata file replaces data in the previously delivered volume ABC-001-001 then the file name ABC-001-001-fix.txt should be used. Replacement document image file names, likewise, must be labeled with the same Identification number as the image being replaced. A separate correcting file is required for each volume to be corrected. These replacement or corrected productions must be delivered on separate media from any new productions. In the event that corrections alter the statistical information previously reported in the cover letter spreadsheet, an updated submission index spreadsheet with the corrected final statistical information must accompany each replacement or corrected submission. <u>Provide this in hard copy format and electronically on the deliverable media.</u>

METADATA TABLE OF REQUESTED FIELDS

Review carefully as fields have been added or modified from the Commission's last set of specifications.

A "**X**" indicates that the field should be populated in the load file produced. "Other ESI" includes non-email files, such as, but not limited to MS Office files, WordPerfect files, etc.

Field Name	Field Description	Field Type	Hard-Copy	Email	Other ESI	Calendar Items
COMPANIES	Company submitting data	Multi-Entry		**X**	**X**	**X**
MEDIAID	The unique identifier on the physical piece of media (*e.g.*, ABC001)	Note Text		**X**	**X**	**X**
VOLUMENAME	Production volume number (*e.g.*, ABC001-001)	Note Text		**X**	**X**	**X**
CUSTODIAN	Custodian(s) / source(s) -- format: *Last, First* or *ABC Dept.* Be consistent across all ESI sources/productions.	Multi-Entry		**X**	**X**	**X**
TIMEZONE	The TimeZone in which the custodian is located.	Note Text		**X**	**X**	**X**
SPEC#	Subpoena/request paragraph number to which the document is responsive	Multi-Entry		**X**	**X**	**X**
HASHMD5	Document MD5 hash value (used for deduplication or other processing)	Note Text		**X**	**X**	**X**
HASHSHA	Document SHA1 hash value (used for deduplication or other processing)	Note Text		**X**	**X**	**X**
SEARCHVALUES	List of search terms used to identify record as responsive (if used)	Multi-Entry		**X**	**X**	**X**
BEGDOC#	Start Identification number (including prefix) -- No spaces or special characters	Note Text		**X**	**X**	**X**
ENDDOC#	End Identification number (including prefix) -- No spaces or special characters	Note Text		**X**	**X**	**X**
DOCID	Must equal the value appearing in the BEGDOC# field and be UNIQUE	Note Text		**X**	**X**	**X**
NUMPAGES	Page count	Integer		**X**	**X**	**X**
PARENTID	Parent record's BEGDOC#, including prefix (populated ONLY in child records)	Note Text		**X**	**X**	**X**
ATTACHMENTIDS	Child document list: BEGDOC# of each child (populated ONLY in parent records)	Multi-Entry		**X**	X	**X**
FAMILYRANGE	Range of the BEGDOC# value of the parent record to the ENDDOC# value (including prefix) of the last child record (for example, ABC-JD-00001201 - ABC-JD-00001220); populated for all documents in the group. Empty if the record is NOT in family grouping	Note Text		**X**	**X**	**X**

Field Name	Field Description	Field Type	Hard-Copy	Email	Other ESI	Calendar Items
EPROPERTIES	Indicate all that apply : Record Type: E-Doc, E-Doc Attachment, Email, Email Attachment, Hard Copy, Calendar Appt Other Notations: Translation of [DOCID of original], Translated as [DOCID of Translation] Privilege Notations: Redacted, Privileged, Family Member of Priv Doc	Multi-Entry		**X**	**X**	**X**
FOLDERLABEL	Email folder path (sample: Smith,James-Inbox\Active); or Hard Copy folder/binder title/label prepended with Custodian Name.	Multi-Entry		**X**		X
FROM	Author of the Email or Calendar item (as formatted on the original)	Note Text		**X**		**X**
TO	Recipients of the Email or Calendar Item (as formatted on the original)	Multi-Entry		**X**		**X**
CC	Names of the individuals who were copied on the Email or Calendar Item (as formatted on the original)	Multi-Entry		**X**		**X**
BCC	Names of the individuals who were blind-copied on the Email or Calendar Item (as formatted on the original)	Multi-Entry		**X**		**X**
SUBJECT	Email or calendar subject	Note Text		**X**		**X**
DATE_HC	Date of hard copy documents, if coded. Format: YYYYMMDD.	Date				
DOCDATE	This is a multipurpose date field. Populate with: DATESAVED for E-Docs; DATESENT for Emails; DATEAPPTSTART for calendar appointments; DATE_HC for hard copy documents, if available. Format: YYYYMMDD.	Date		**X**	**X**	**X**
DATECREATED	Date electronic file was created. Format: YYYYMMDD.	Date			**X**	
DATESENT	Date the Email or Calendar Item was sent. Format: YYYYMMDD.	Date		**X**		**X**
TIMESENT	Time Email or Calendar Item was sent -- Format: HH:MM:SS (use 24 hour times, *e.g.*, 13:32 for 1:32 pm; timezone indicators cannot be included)	Time		**X**		**X**
DATERECEIVED	Date Email or Calendar Item was received. Format: YYYYMMDD.	Date		**X**		**X**
TIMERECEIVED	Time Email or Calendar Item was received. Format: HH:MM:SS (use 24 hour times, *e.g.*, 13:32 for 1:32 pm; timezone indicators cannot be included)	Time		**X**		**X**
HEADER	The internet header information for Email sent through the internet;	Note Text		**X**		
INTERNETMSGID	Globally unique identifier for a message which typically includes messageid and a domain name. Example: <0E6648D558F338179524D555@m1p.innovy.net	Note Text		**X**		**X**

Field Name	Field Description	Field Type	Hard-Copy	Email	Other ESI	Calendar Items
MESSAGEID	Unique system identification number for the e-mail message assigned by the proprietary email database/mailstore/post office file associated with centrally managed enterprise email servers. EntryID for Microsoft Outlook, the UniqueID (UNID)	Note Text				
INREPLYTOID	Internet message ID of the Email replied to	Note Text		**X**		
CONVERSATIONINDEX	Email Thread Identification	Note Text		**X**		**X**
IMPORTANCE	Email flag indicating priority level set for message	Note Text		**X**		**X**
DELIVRECEIPT	Delivery receipt request notification for Email messages	Note Text		**X**		**X**
READRECEIPT	Read Receipt request notification for Email messages	Note Text		**X**		**X**
SENSITIVITY	Sensitivity field from Email messages	Note Text		**X**		**X**
REVISION	Revision number extracted from metadata of native file	Note Text			**X**	
DATESAVED	Date native file was last modified. Format: YYYYMMDD.	Date			**X**	
DATEPRINTED	Date native file was printed (metadata derived from Word documents, etc.)	Date			**X**	
EORGANIZATION	Company field extracted from the metadata of a native file	Note Text			**X**	
EAUTHOR	Author field value extracted from the metadata of a native file	Note Text			**X**	
LAST_AUTHOR	Last Saved By field value extracted from metadata of a native file	Note Text			**X**	
ESUBJECT	Subject field value extracted from metadata of a native file	Note Text			**X**	
FILESIZE	File size in Bytes (integer value only - do not include unit of measure, thousands character, or decimal places - *e.g.*, 1008 not 1,008)	Integer		**X**	**X**	**X**
FILENAME	File name of native file (E-Docs or attachments to Email)	Note Text		**X**	X	**X**
APPLICATION	Application used to create native file (*e.g.*, Excel, Outlook, Word)	Note Text		**X**	**X**	**X**
FILEEXTENSION	File extension of native file	Fixed Length 5 chars		**X**	**X**	**X**
FILEPATH	File path to native file as it existed in original environment, prepended with Custodian Name.	Multi-Entry		**X**	**X**	**X**
DOCLINK	File path location to the current native file location on the delivery medium	Note Text			**X**	

Field Name	Field Description	Field Type	Hard-Copy	Email	Other ESI	Calendar Items
DATEAPPTSTART	Start date of calendar appointment. Format: YYYYMMDD.	Date				**X**
TIMEAPPTSTART	Start time of calendar appointment. Format: HH:MM:SS (use 24 hour times, *e.g.*, 13:32 for 1:32 pm; timezone indicators cannot be included)	Time				**X**
DATEAPPTEND	End date of calendar appointment. Format: YYYYMMDD.	Date				**X**
TIMEAPPTEND	End time of calendar appointment. Format: HH:MM:SS (use 24 hour times, *e.g.*, 13:32 for 1:32 pm; timezone indicators cannot be included)	Time				**X**

Submission Index Spreadsheet for ABC Company Production

Matter: ABC's Acquisition of XYZ

Custodian	Physical Media Name	Volume Name	Begin I.D. Number	End I.D. Number	Intentionally Left Blank	Number of Records	Number of Images	Number of Native Files	Number of Extracted/OCRed Text Files	Volume Size in GB	Date Produced
Doe, John K.	ABC001	ABC001-001	ABC-JKD-00000001	ABC-JKD-00005825		258	5825	13	258		
Doe, Jane B.	ABC001	ABC001-001	ABC-JBD-00000001	ABC-JBD-00003711		365	3,711	52	362		
					TOTALS FOR THIS VOLUME	**623**	**9,536**	**65**	**621**	**.50**	mm/dd/yyyy
Daniels, James	ABC001	ABC001-002	ABC-JAD-00000001	ABC-JAD-00005727		1,150	5,727	156	1,149		
Wise, Barry	ABC001	ABC001-002	ABC-BLW-00000001	ABC-BLW-00003010		600	3,010	20	598		
					TOTALS FOR THIS VOLUME	**1,750**	**8,737**	**176**	**1,750**	**1.39**	mm/dd/yyyy
Brown, Charlene J.	ABC002	ABC002-001	ABC-CJB-00000001	ABC-00006952		1,315	6,952	68	1,310	1.02	mm/dd/yyyy
Mahan, Freda	ABC002	ABC002-001	ABC-FRM-00000001 ABC-FRM-00000698	ABC-FRM-00010400 ABC-00000982		2,023	10,115	85	2,023	2.00	Mm/dd/yyyy
					TOTALS FOR THIS VOLUME	**3,338**	**17,067**	**153**	**3,333**	**1.39**	
					GRAND TOTALS	**5,711**	**35,340**	**394**	**4,121**	**3.28**	

*** This spreadsheet can be provided on a current production-specific or a cumulative production basis, but the totals per *volume* must be provided**

**** This spreadsheet is a sample of format and does not reflect the names/Identification numbers of the sample files provided**

***** Volume size is actual space used not total available space on the delivery media**

THIS PAGE

INTENTIONALLY

LEFT BLANK