Practice Memorandum No. 5

E - Disclosure

Purpose. The effective use of technology in the disclosure process can save time and reduce cost. This practice memorandum and the accompanying checklist and protocol are intended to assist parties in using technology effectively for electronic disclosure ("e-disclosure").

The protocol may be used as a basis for agreement. It too should be taken as a checklist. Each line of the protocol suggests a subject for consideration. Nothing is compulsory.

Present Practice. Rule 16 - Disclosure of Electronic Information seeks to prevent erosion of the ethic of disclosure as a result of the transition from print media to electronic media.

The Rule calls, firstly, for an agreement by the parties. Failing that, it imposes terms that are technically detailed and generic. Thirdly, a judge can impose terms.

Case management judges are often involved when terms for disclosure are being set. Otherwise, judicially imposed terms are rare. It is also rare for a party to insist on all of the default terms imposed by Rule 16.

Almost always, disclosure is by agreement. Sometimes, the agreement is detailed. Often, it is simple. Unfortunately, the terms are not always explicit.

Simple Agreements. Rule 16 does not stand in the way of an explicit, simple agreement. Parties to agreements of this kind should ask themselves a few questions. Am I providing, and receiving, sufficient disclosure? Is disclosure being made too late? Is the risk of losing relevant evidence sufficiently managed?

There are reasons for believing that more detailed arrangements may become more common.

Protocol and Rules. E-disclosure may be more dependent on standardized formats than print disclosure. The protocol can be used for disclosure of electronic information under Rule 16, for e-disclosure of print documents under Rule 15 - Disclosure of Documents, and for the production of a common file under Rule 40.05 or Rule 51.10.

Readily Exchangeable Format. Provision is made in the following Rules for providing the following components of e-disclosure in "readily exchangeable format":

- Schedule A in attached document, Rule 16.09(3)(d)
- Copies of the listed documents, Rule 15.03(1)(3)(e)
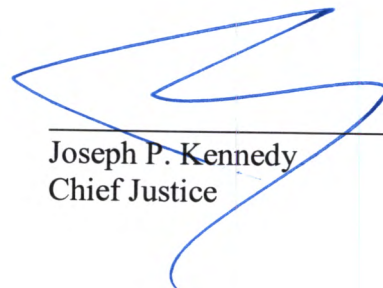- The electronic information, Rule 16.09(3)(e).

"Readily exchangeable format" means that parties agree on common standards and formats for e-disclosure, so that:

- E-disclosure can be made even if parties are using different software
- Disclosure from all parties is consistent and meaningful
- Each party bears the cost of its own disclosure
- Conversion costs may be minimized or avoided altogether
- E-disclosure may be reorganized and re-used if required in future proceedings.

Protocol. Seamless and cost-effective e-disclosure involves attention to technical details. Parties are therefore encouraged to negotiate and agree on the terms of a protocol that sets out in detail the characteristics of all files that will be disclosed electronically. Parties should also bear in mind that e-disclosure should contemplate the potential for the electronic production of documents and electronic information as evidence in trials and hearings.

Costs. The reasonable costs incurred in e-disclosure, including the expenses of retaining or utilizing necessary external or in-house technical consultants, may be claimed as disbursements under Rule 77.10.

Adopted by the court on June 28, 2013.

Joseph P. Kennedy
Chief Justice

Electronic Disclosure Protocol Checklist

- **Determine the software used by each party and how exchange can be compatible**
  - Electronic Information Organizational Software
  - Internet Explorer
  - Adobe Acrobat
  - other

- **Agree on the format to which paper files will be converted for electronic exchange**
  - TIFF
  - PDF
  - Discuss oversized documents

- **Agree on the format for exchanging electronic documents**
  - Native/near-native or converted to TIFF or PDF
  - Searchable or not searchable
  - Metadata included, not included, or partly included

- **Agree on standards to identify how documents were originally grouped together**
  - Use cross-referencing document id to know which document is the "parent" and which are the attachments
  - Identify the folder where the document originated
  - Identify the source/custodian from whom the document was produced
  - Identify the extent to which documents will be produced in bundled or separate form (i.e. family physician or hospital charts containing consulting and lab reports, agreements with appendices, etc.)

- **Agree on how to deal with duplicates**
  - Exchange duplicates and identify them as such
  - Do not exchange duplicates
  - Define "duplicate"
  - De-duplicate within custodians only or across custodians

- **Agree on how to identify redactions**
  - Stamp the reason for the redaction on the document
  - Provide a reason for the redaction in Schedule B

- **Agree on which fields will appear in Schedule A for print documents and the format for the fields. *Note: the first 5 fields are mandatory***
  - *Author [John Smith, Smith, J., etc.]*
  - *Recipient*
  - *Date [2012-01-20, 1/20/2012, etc.]*
  - *Type [memo, deed fax, report]*
  - *Document id number/or identifier*
  - Custodian

- Source
- Copyee
- Parent date
- Parent id
- Attachment
- Folder label

- **Agree on standards for describing document types**
  - Choose a standard common list of identifiers and definitions, for example:
    - "Agreement" – contracts, deeds, etc.
    - "Invoices" – cheques, statements, receipts, etc.
    - "Plan" – engineering drawings, blueprints, etc.

- **Agree on which fields will appear in the Schedule A and the format of the fields for electronic information. (Most litigation software will automatically populate many of these fields when documents are loaded.)**
  - Document id
  - Date created
  - File type (.xlsx, .jpg, etc.)
  - Author
  - Recipient
  - Title
  - Parent date
  - Parent id
  - Attachment id
  - Path
  - Source
  - Cc
  - Bcc
  - Full text
  - Custodian
  - Date last modified
  - Date received
  - Time sent
  - Hash code

- **Agree on how the documents will be exchanged**
  - Format (Internet Briefcase, DLL, IDX, Internet Briefcase)
  - Medium (Secure FTP, CD, DVD, thumb drive, hard drive)

Electronic Disclosure Protocol

Introduction[1]

This protocol is designed to improve the quality and efficiency of disclosure in civil litigation while avoiding unnecessary costs for the parties and the court. Technical terms used are defined in an attached glossary. As explained in the practice memorandum, this protocol is to serve as a model with the accompanying checklist to guide the parties in considering the details and format for electronic exchange of relevant documents and electronic information ("EI").

In general terms, the parties may agree that disclosure will be made electronically in the following way:

1. For paper source documents:
   a. scanned to TIFF image file format
   b. OCR'd for searchable full text
   c. objectively coded in an electronic affidavit.

2. For electronic information:
   a. native format and, for e-mail, near-native format
   b. converted to TIFF image file format where practicable
   c. searchable full text extracted where available
   d. metadata extracted in an electronic affidavit.

---

[1]A protocol ensures that no party is compelled to use any particular software to search, view, or annotate the documents (hard copy), EI, and the electronic Schedule A of other parties. As a model protocol, this document contains sample alternative language in italics and explanatory footnotes.

Part A: Paper Documents

Standards for Separating and Grouping Documents Into Database Records

Physical Groupings

Though each document found in a file folder should be listed individually in the Schedule A, the relationship between the documents will be maintained through cross-referencing of document ids so that the integrity of the original paper collection can be maintained and re-established in evidence. An individual document is usually identified as having an author, title, type or date that differentiates it from others in the same family.

Or

Though each document found in a file folder should be listed individually in the Schedule A, the relationship between the documents will be maintained through the file folder field so that the integrity of the original paper collection can be maintained and re-established in evidence. An individual document is usually identified as having an author, title, type or date that differentiates it from others in the same folder.[2]

A double-sided sheet may be treated as a single document. (See also "Single-sided Images of Double-sided Documents" below.)

Every document referenced as an attachment or enclosure by another document or in the same family should be identified as an attachment, and the first, referencing or attaching document should be identified as a parent.

A letter enclosing a report would be considered a parent. The letter and the report would be listed and described individually in the electronic Schedule A, and would cross-reference each other. On the other hand, a book with multiple chapters would be considered and listed as a single document.

There is no hard and fast rule about whether the contents of binders are considered one document or a family. The correct grouping will depend on the balance between the relevance of each individual document within a binder and the cost of coding. For example, a binder of "conference materials" may contain 15 papers written by different authors on different topics. Theoretically, each document should be coded individually and then cross-referenced to the binder family. However, if the binder is relevant only to show that its custodian attended a conference, it may be coded as one document.

---

[2] In this case the following fields will not be disclosed: parent id and attachment id.

Appendices, attachments, and schedules, which form part of an agreement, report, or legal document, will generally not be coded as separate documents, but will be considered part of the agreement, report, or legal document unless the context requires otherwise.

Copies of Documents

Images
Documents up to 11" x 14" will be imaged as single-page 8-1/2" x 11" black and white TIFF (CCITT Group 4) files, with a resolution of 300 dpi. Blank pages will not be imaged, unless the blank page is relevant.

Or

Documents up to 11" x 14" will be imaged as multi-page 8-1/2" x 11" black and white searchable PDF files, with a resolution of 300 dpi. Blank pages will not be imaged, unless the blank page is relevant.

Oversize Documents
Where numerous oversize documents are relevant, for example blueprints, whiteprints, maps, or posters, and where the cost of scanning is disproportionate to the importance of the evidence, the oversize documents must be listed in the Schedule A but only the title block if any must be imaged.

Colour Documents
Where the colour of a document materially affects its meaning or legibility, it will be imaged to colour JPEG format in 300 dpi.

Single-sided Images of Double-sided Documents
If a hard copy contains meaningful writing on only one side, the other side need not be imaged. For example, where a party is imaging 500 standard form invoices, and one side contains pre-printed boilerplate which is not relevant to the litigation, that side will not be imaged more than once.

Or

If a hard copy contains any writing on a side, that side will be imaged, whether the writing is boilerplate text or otherwise.

Branding (Endorsement) of Each Image
The parties agree to permanently burn in or brand each image with its corresponding document id.

Or

The parties agree to not burn in or brand each image with its corresponding document id.

Adhesive Notes
Where an adhesive note is relevant and is contained within the pages of a document, the adhesive note should be imaged as a separate page within the document. Where the note is privileged, it can be redacted

or eliminated from the produced image; however, the corresponding record for the privileged adhesive note should indicate that it has been redacted. Irrelevant adhesive notes need not be imaged.

Redaction of Images

Privileged or irrelevant portions of documents may be redacted on the TIFF image using black overlay or masking so that the redaction is noticeable. The original non-redacted TIFF image must be readily accessible if required by the court. Where images have been redacted, the fact and reason for the redaction should be indicated on the redaction itself or the corresponding record.

Inadvertent Disclosure of Redacted Information

If any text is inadvertently disclosed or any rendering of a document displays information that was redacted on the TIFF image, recourse may be had to Rule 14.06.

Searchable Full Text

To facilitate full text searching, the parties agree to exchange searchable full text renderings of each document containing OCRable text generated by optical character recognition. The text will be generated using industry-standard OCR software and provided "as-is": it will not be manually edited or enhanced in any way. Parties will not transcribe handwriting.

Or

To facilitate full text searching, the parties agree to exchange documents in searchable PDF format. The text will be provided "as-is": it will not be manually edited or enhanced in any way. Full text will not be provided for redacted documents. Parties will not transcribe handwriting.

Or

The parties agree that searchable full text will only be exchanged for documents that have not been redacted.

Or

The parties agree that searchable full text for documents will not be exchanged.

Standards for Listing and Describing Documents in the Electronic Schedule A

The parties agree that documents will be listed and described in the electronic Schedule A by way of objective coding. (The suggested fields are for consideration by the parties and not all fields will be necessary in all cases.)

Field name	Rule for content and format
document id (single entry, unique) [ProdNo] (single entry, unique)[3]	AAA00000000[4] For rolling productions, documents will be numbered continuously.
date (single entry)	MM-DD-YYYY Undated documents will be coded as 00-00-0000. Documents with partial dates will be coded with partial dates. For example, the month and year (e.g. August 1997) will be coded 08-00-1997.[5] Or Undated documents will be coded as 01-01-1900. Documents with partial dates will be coded with "01" as the day or month and 1900 as the year. For example, a document marked "June" will be coded 06-00-1900.

---

[3] ProdNo is a unique serial identifier for each record irrespective of the number of pages.

[4] This field contains a unique reference to each record. This is a two part alphanumeric entry in the form "AAA00000000" where three upper-case letters represent the party code (see below) and a zero-filled 8-digit serial number referring to each sequential record in the database on a record level or where applicable (for example, where documents have been imaged to multi-page format), or on a page level, in which case the document id is the unique reference for the first page of the document.

For example, a three page document from party "GHG" would be numbered as follows:

        Page 1: GHG00000001
        Page 2: GHG00000002
        Page 3: GHG00000003

The next produced document would be numbered as follows:

        Page 1: GHG00000004
        Page 2: GHG00000005
        Etc.

Attachments are listed and numbered separately from their parent records, not in groups or bundles. (See "Standards for Separating and Grouping Document Families into Database Records" above.)

[5] Note that not all software allows for partial dates. Therefore, usage should be discussed prior to disclosure to ensure that the software being used will accept such dates.

Field name	Rule for content and format
	The parties agree to use the following priority for coding dates: 1. latest revised/updated date – the document must indicate that it has been revised or updated 2. latest date of creation (top, bottom of page, or end of document) 3. latest approved date 4. latest published date 5. latest copyright date 6. latest date from title – If the latest date is a future date then code the latest non-future date 7. latest stamp date 8. latest print date. When coding agreements or contracts the parties agree to use the following priority: 1. execution date 2. effective date 3. signature date. When coding court documents and legal documents the parties agree to use the following priority: 1. effective date 2. contract date 3. execution date 4. filed date 5. stamped date.
type (single entry)	This field can be completed using common document types, for example memo, deed, letter, fax, annual report, etc. See the sample list of document types in the section "Standard for Listing and Describing Documents".
author (multiple entry)	The person(s) or organization(s) who authored the document. To be completed using information indicated on the face of the document. Examples author (person) "MacKay, Andrew" author (organization): "Hilltop Strawberries Inc" Or Person or organization who authored the document together with the person's organizational affiliation. To be completed using information on the face of the document. Last name, first name [organization].

Field name	Rule for content and format
	<u>Examples</u> Author is identified as having an organizational affiliation: "MacKay, Andrew [Hilltop Strawberries Inc]". Author is not identified with an organization: "MacKay, Andrew". Author is an organization: "Hilltop Strawberries Inc". The parties agree that names will be coded in the agreed format using only the information that appears on the face of the document. For example, "Andy MacKay", "A.L. MacKay", and "Andrew M." will be coded as "MacKay, Andy", "MacKay, AL", and "M, Andrew" respectively. Or The parties agree that names will be normalized for consistency. For example, "Andy MacKay", "A.L. MacKay", and "Andrew M." will all be coded as "MacKay, Andrew" if it is ascertained that these variations all refer to the same person. Similarly, "Hilltop", "Hilltop Inc.", and "HSI" will be normalized to "Hilltop Inc".
recipient (multi-entry)	The person(s) or organization(s) who received the document. See rules for author field.
title (single entry)	Exact title of a document such as "Report on Technology" or the "re" line in correspondence. If the title is not obvious from the face of the document this field should be left blank. Or If the title is not obvious from the face of the document enter "untitled".
copyee or cc (multi-entry)	The parties agree to not disclose copyee or cc. Or The parties agree to disclose the person(s) or organization(s) who received the document as copyee or cc. See rules for author field.
parent date/lead date	The date of the parent/lead of the document (used for sorting documents chronologically with their families). Where a document has no parent, the parent date will contain the date of the document.

Field name	Rule for content and format
parent id (single entry)	The document id of a lead or parent document (used for attachments only).
attachment id(s) (multi-entry)	The document id of an attached document (used for parent documents only).
folder label (single entry)	The contents of the folder or binder label in which the document was found.
custodian (single entry)	This field identifies the person or organization from whom the document was obtained. See rules for author field.
redacted (single entry)	If the produced copy has been redacted indicate "yes" or else leave blank.
redacted reason (single-entry) (This field is not required if the reasons for redaction appear on the redaction itself.)	If a document has been marked "yes" in the redacted field above, then indicate the reason. For example, "litigation privilege", "cabinet secret".

Party Codes

Party	Party Code
plaintiff	ABC
defendant 01	DEF
defendant 02	GHI
third party	IJK

Standards for Coding Document Types

The parties agree on a common list of document types set out below. Note that the suggested fields below are for consideration by the parties and not all fields will be necessary in all cases.

Document Type	Description/Example
agreement	Includes contracts, deeds, etc.
agenda	Outline of meeting, business, seminar, or conference events scheduled to take place.
appendix	Includes appendices, schedules, annexes that were originally part of a larger document, usually a report or contract, but have become separated from the body of the larger document.
budget	Material giving financial details or breakdowns of projects, staffing, statement of resources, allocation of resources, etc. Usually called a budget. See also financial document.
certificate	Use for actual certificates such as birth or marriage certificates. Do not use for share certificates – they are corporate documents.
chart/table	Any document in chart or table form separated from a larger report.

Document Type	Description/Example
corporate documents	Share certificates, shareholder agreements, articles of incorporation, other corporate records.
court documents	Includes pleadings, affidavits, motion records
financial document	Use for individual reports containing financial information – the information must be financial and not simply a list of numbers. Examples: balance sheets, operating costs, A/P, A/R, reconciliation records, income statements, all banking documents, exchange rates, consolidated statements.
invoices, cheques, statements, receipts	Enter the invoice, cheque, statement or receipt number in the title field.
letter	Must have an addressee and a signature line, and usually has an address block.
manual	Includes procedural manuals, service manual, maintenance manual, user guide, operating instructions, guidelines, and specifications.
marketing	Includes advertisements, brochures, flyers, and the like.
memorandum	Usually formatted To: From: Re: Date:
note	Brief, informal comments or notations – can be typed or handwritten.
presentation	Materials used for presentations, such as Power Point deck, overheads, and the like.
plan	Engineering, architectural, or building drawings, plans, blueprints.
report	Usually has a formal title and indicates who prepared it (the author) and when. Note that financial reports of any length are coded to financial document. Do not use for corporate documents such as annual reports.

Part B: Electronic Information (EI)

Software Used by the Parties

The parties agree that the following electronic file formats may be read by the corresponding parties and are readily exchangeable formats for the purposes of Part V of the *Civil Procedure Rules*:

Plaintiff

1. *[e.g., Microsoft Excel2010 worksheet]*;
2. *[e.g., Corel WordPerfect 7.0 document]*;
3. ... .

Defendant

1. *[e.g., Quickbooks 5.1 data file];*
2. *[e.g., Microsoft Word 97 document];*
3. ... .

Standards for Separating and Grouping EI into Database Records

Each file in a container file, the container file itself, and all attachments to e-mail messages ("families"), will be listed individually in the electronic Schedule A. The relationship between the files will be maintained through the source/custodian field. (An individual file is readily identified as having a unique path and filename that differentiates it from others in the same family.)[6] The relationship between loose electronic files whose content may refer to a parent or attachment file is not required.

Or

Each file in a container file, and all attachments to email messages ("families"), will be listed individually in the electronic Schedule A. The relationship between the files will be maintained through cross-referencing of document ids. (An individual file is readily identified as having a unique path and filename that differentiates it from others in the same family.)

EI files whose content refers to another file or document are not considered parents. For EI the family relationships are based solely on e-mail/attachment relationships as established in the native e-mail application, and families of files grouped together in a container file.

Every attachment will be identified as an attachment, and the attaching e-mail or container file will be identified as a parent. Where there are multiple container files within container files, the process of identifying families may become less useful and less practicable and is therefore optional.

Copies of Electronic Information

The parties will exchange EI in native format if readily exchangeable, in near-native format where the native format is not readily exchangeable, and where the EI contains privileged information, as redacted TIFF images converted directly from the EI (where practicable), and as searchable text (where available), without affecting the disclosing party's obligation to preserve the mirror image of the non-redacted EI.

Or

The parties will exchange EI as TIFF images (where practicable), in native format where not practicable, and as searchable text (where available).

---

[6] In this case the following fields will not be disclosed: parent id and attachment id.

Format

Where EI is rendered in static format, such as TIFF or PDF, parties will use the same imaging guidelines as for documents. E-mail messages will be extracted from e-mail stores such as Microsoft Exchange, Lotus Notes and PST archives and rendered as individual TXT or HTML, as long as any attachments are kept together in the family relationship.[7]

Colour EI

Where the colour of a file materially affects its meaning or legibility, it will be imaged to colour JPEG format in 300 dpi, or provided in native or near-native format that renders the colour accurately.

Branding (Endorsement) of Each Image

For EI that has been converted to TIFF, the parties agree to permanently burn in or brand each image with its corresponding document id.

Or

The parties agree to not burn in or brand any image with its corresponding document id.

Redaction of Images

Privileged or irrelevant portions of EI may be redacted on the TIFF image using black masking so that the redaction is noticeable. The original non-redacted TIFF image must be readily accessible if required by the court. Where images have been redacted, the fact and reason for the redaction should be indicated on the redaction itself or in the corresponding record.

Inadvertent Disclosure of Redacted Information

Where any TIFF image has been redacted, the parties will ensure that the redaction is burned into the TIFF image and that the extracted text (see below) for the EI will be generated from a redacted version of the image. If any privileged text is inadvertently disclosed in a TIFF or extracted text displays information that was redacted on the TIFF image, recourse may be had to Rule 14.06.

Searchable Full Text

To facilitate full text searching the parties agree to exchange extracted full text from all EI, extracted to ASCII or Unicode format. The extracted text will not be manually edited or enhanced in any way. Full text will not be provided for redacted EI.

Where EI contains an image of text, for example a non-searchable PDF file, the parties agree to not OCR that text.

Or

---

[7] Some files, including most e-mail, cannot be reviewed for production and/or produced without some form of conversion. Most e-mail files must be extracted and converted into individual files for document review and production. As a result, the original format is altered and they are no longer in native format. There is no standard format for near-native file productions. Files are typically converted to a structured text format such as HTML or XML. These formats do not require special software for viewing.

Where EI contains an image of text, for example a non-searchable PDF file, the parties agree to OCR that text.

To facilitate full text searching the parties agree to exchange EI in searchable PDF format. The text will be provided "as-is": it will not be manually edited or enhanced in any way. Full text will not be provided for redacted EI.

Or

The parties agree that searchable full text will only be exchanged for documents that have not been redacted.

Or

The parties agree that searchable full text for documents will not be exchanged.

Standards for Listing and Describing Electronic Information in the Electronic Schedule A

The parties agree that all metadata will be preserved and produced upon request, if relevant, but the parties will not exchange metadata in the preliminary disclosure.

The parties agree that EI will be listed and described through the use of available metadata and supplemented with objective coding only when necessary for meaningful disclosure. For example, the filename of a Word document is often not descriptive. Although metadata for EI is not always accurate, the parties agree to ensure that the metadata to be extracted and disclosed will not be edited by the disclosing party. If objective coding is needed for meaningful disclosure, that coding will be entered into the fields designated for the listing of documents as set out above. For that reason some of the field names for extracted metadata are different from the field names used for objective coding.

Field name	E-file	E-mail
document id	Where copies of EI have been rendered in single-page TIFF format the document id is the same format as for scanned documents. Where copies of EI have been rendered in native or near-native format and are not paginated, the document id shall be a unique sequential number representing each record, i.e. AAA00000001	Where copies of EI have been rendered in single-page TIFF format the document id is the same format as for documents. Where copies of EI have been rendered in native or near-native format and are not paginated, the document id shall be a unique sequential number representing each record, in the same format as for documents, i.e. AAA00000001
date created	Date created metadata field	Date sent metadata field
file type	File extension metadata, e.g. xlsx, jpg	File extension metadata, e.g. msg, eml, txt
author	Contents of author metadata field (not normalized)	Contents of from metadata field (not normalized)

Field name	E-file	E-mail
recipient	[None]	Contents of to metadata field (not normalized)
title	Complete filename and extension	Contents of subject metadata field
parent date/lead date	The date of the parent of a document (used for sorting documents chronologically with their families). Where a document has no parent, the lead date field will contain the date of the document.	The date of the parent of a document (used for sorting documents chronologically with their families). Where a document has no parent, the lead date field will contain the date of the document.
parent id	The document id of a parent document (used for attachments only.) For example a ZIP or other container file, or the document id of an attaching e-mail message.	The document id of a parent document (used for attachments only).
attachment id	The document id of an attachment document (used for parent documents only.)	The document ID of an attachment document (used for parent documents only).
path	Complete file path extracted from system	E-mail account folder label from which e-mail was extracted
source	Filename and path for the container file containing the record (e.g. \personal\xmascards2008.zip)	Container file, path and folder for e-mail records (e.g. \backups\2007\archive2.pst…\In Box)
cc	[None]	Contents of cc metadata field (not normalized)
bcc	[None]	Contents of bcc metadata field (not normalized)
full text	The full text of the file in plain text searchable format (excluding redacted text)	The full text of the e-mail in plain text searchable format (excluding redacted text)
custodian	Name of custodian (not normalized)	Name of custodian (not normalized)
date last modified	Extracted from metadata	[None]
date received	[None]	Extracted from metadata – time zone should be noted
time sent	[None]	Extracted from metadata – time zone should be noted.
time received	[None]	Extracted from metadata – time zone should be noted
hash code (This will be included only if this field is automatically populated by the software being used).	Generated by processing software	Generated by processing software
redacted	Indication of whether the EI static image has been redacted or explanation of why the EI static image was redacted	Indication of whether the EI static image has been redacted or explanation of why the EI static image was redacted

Field name	E-file	E-mail
dupeflag (This will be included only if the parties chose to exchange duplicates).	Indication of whether the EI is a duplicate of another disclosed EI record, based on the hash code or other agreed method	Indication of whether the e-mail message is a duplicate of another disclosed message, based on the hash code or other agreed method

Standards for Preparing Files that will be Exchanged

File Format for Schedule A (for Documents and EI)

The format for the ASCII or Unicode text file containing the electronic Schedule A records is Comma Separated Values ("CSV"). Entries for each field will be delimited by a comma, and entries for records are separated by a new line (carriage return <CR>[8]). Entry contents will be enclosed in quotation marks where necessary to ensure that commas embedded in the content of a field are not interpreted as a delimiter. Multiple entries in a field are delimited by a semi-colon.[9]

Or

The format for the ASCII or Unicode text file containing the electronic Schedule A records is tab-delimited. Entries for each field will be delimited by a tab, and entries for records are separated by a new line (carriage return <CR>). Multiple entries in a field are delimited by a semi-colon.

Header Record

The first row (record) in the Schedule A file will contain the name of each field, to facilitate the loading process.

Load Files – Format

The load file contains a list of document ids for each record in the Schedule A with an entry linking to the images, extracted text, OCR, native files and near-native files, as the case may be, for that record. [10]

---

[8] "<CR>" stands for "carriage return" which derives from the era of typewriters.

[9] Where other delimiters are used to ensure proper data migration, the load files will contain a "read me" file clearly identifying the delimiters used. For example, (ASCII 124 |for field separator and ASCII 094 ^ for "quote character", in Summation.

[10] If a document in Schedule A contains more than one page, its record will contain a reference to more than one image. On the other hand, each record will refer to only one OCR text file, and only one native or near-native format file.

Sample Load File – large volume of files saved in subfolders

DOC ID	Image filename and path	Full text filename and path	Native filename and path
AAA00000034	\images\AAA\01\00000034.tif \images\AAA\01\00000035.tif \images\AAA\01\00000036.tif	\TXT\AAA\01\00000034.txt	\EI\AAA\01\00000034.doc
AAA00000037	\images\AAA\01\00000037.tif	\TXT\AAA\01\00000037.txt	\EI\AAA\01\00000037.xlsx

Sample Load File – files saved in folders

DOC ID	Image filename and path	Full text filename and path	Native filename [11] and path
AAA00000034	\images\AAA00000034.tif \images\AAA00000035.tif \images\AAA00000036.tif	\TXT\AAA00000034.txt	\EI\AAA00000034.doc
AAA00000037	\images\AAA00000037.tif	\TXT\AAA00000037.txt	\EI\AAA00000037.xlsx

General

De-duplication

Where commercially feasible the parties will use technology to de-duplicate EI and duplicate records will not be listed in Schedule A.

The parties agree to de-duplicate imaged document records where practicable.

The parties will de-duplicate both within and across custodians.[12]

Malicious Code

All parties should always perform a check for malware before sending or loading e-disclosure files. However, it is the responsibility of the disclosing party to provide a "clean" disclosure.

Medium

Optical Media

Where volume permits, files will be conveyed on permanent write-once read-many optical media such as CD-R or DVD-R.

---

[11] Note that even if the EI copy disclosed here has a filename that corresponds to the document id, the original filename will always be found in Schedule A.

[12] For example, if two or more custodians have an identical e-mail (with or without attachments), only one will be produced provided the document identifies all recipients.

Magnetic Media

Where volume does not permit the efficient transfer by way of optical disc, files may be transferred on rewritable magnetic or solid state media (external thumb or hard drive), or by way of secure FTP. Files should not be transferred in "read-only" or "viewer" or "run-time" format.

Labels

All transfer media should be clearly labelled to identify the source, matter name and date of disclosure, and allow for safe return if lost. They should also be labelled as confidential.

Quality Control

Parties agree to exercise care when preparing or supervising the preparation of their e-disclosure, to ensure the accuracy and integrity of the evidence. Inattention to small details such as punctuation, or the formatting of an entry, can have a significant impact on the utility of a database.[13]

---

[13] For example, quotation marks should never be used within database field entries because quotation marks are a delimiter in the CSV file.

Glossary

ASCII. Acronym for "American standard code for information interchange". A standard used for encoding alphanumeric characters. Limited to 128 characters. Superseded by Unicode.

Branding. The imprinting of information such as document id onto the static image of a document or EI.

Container File. A file which contains one or more other files, such as a ZIP file.

CSV (comma separated values). The acronym for comma separated values. A structured text file where entries for each field are separated or "delimited"' by a comma, and entries for records are separated by a new line (carriage return <CR>). Entry contents may be enclosed in quotation marks to ensure that commas embedded in the content of a field are not interpreted as a delimiter. For example:

"Smith, Allan", Halifax<CR>

Brown, Truro<CR> [quotation marks optional because no comma in the field]

"Jones, Roger"<CR>

When imported into a table or spreadsheet, the punctuation is removed and the data is placed in separate fields (ex. name and city):

name   city

Smith, Allan	Halifax
Brown	Truro
Jones, Roger	

Database. EI structured into fields and tables to facilitate sorting, searching and reporting.

De-duplicate. The process of flagging or removing duplicate records.

Default Standard. A standard established by this Practice Memorandum to ensure parties prepare and exchange an electronic Schedule "A" using a common set of descriptive elements in readily exchangeable format.

Delimiter. A character chosen to indicate the beginning and end of one data element within a table. Can be a comma, a tab, a pipe (|) or other standard character.

Document. Document has the meaning set out in Rule 14.02(1) of the *Civil Procedure Rules*

Electronic Information ("EI"). Electronic Information ("EI") has the meaning set out in Rule 14.02(1) of the *Civil Procedure Rules*

Endorsement (See Branding)

Field. A field represents a column of data within a database or a spreadsheet and contains an entry representing some element of a document or EI, for example author or date.

Hash Code. A unique file signature generated by certain applications. Used in the de-duplication process.

Header record. The first record in a table, which names the fields for each following record.

Image. An EI file format that portrays a static picture of an electronic file (as it would look if printed to hard copy), or an electronic image of a document that has been scanned. Common image formats for litigation are TIFF and PDF.

Image Resolution. The level of detail captured when a document is imaged. Measured in dots per inch (dpi). Industry standard is 300 dpi for black and white imaging.

JPEG. An image file format.

Load file. A table linking records in the electronic Schedule A with its associated full text, native files, and image files.

Loose files. EI in the form of files that are not attachments to e-mail messages.

Malicious Code. Malicious code is the term used to describe any code in software intended to cause undesired effects, security breaches or damage to a system. Malicious code describes a broad category of system security terms that includes attack scripts, viruses, worms, Trojan horses, backdoors, and malicious active content.

Native (Native Files or Native Format). A reference to EI in its original digital format, prior to processing, or conversion to image form.

Near-Native Format. A digital rendering of a Native File into readily exchangeable format such as HTML, or TXT. Used where Native Files are not readily exchangeable and conversion to TIFF is not practicable.

Non-Printable EI. EI that is normally not capable of being accurately printed either to hard copy or to an image format – for example, multimedia files, databases, spreadsheets, and websites.

Normalization. The process by which various references to a name throughout a collection of documents are standardized to facilitate searching and sorting.

Objective coding. Indexing. The process of entering descriptive information about a document into a Schedule A.

Optical character recognition OCR. A common method of digitizing printed text so it can be searched electronically.

Parent/Attachment Relationships. The relationship between documents or EI that have been grouped together (usually but not necessarily by the author). For example: a report with enclosures; an e-mail with attachments; a binder containing tabbed sections, a ZIP (compressed) file containing other electronic files.

Path. The full location (folders and subfolders) of a file – e.g. f:\data\2009\projects\

PDF. The abbreviation for "portable document format", which is a popular image file format.

Printable EI. EI that renders accurately when printed to paper or static image format such as TIFF, JPG or PDF. Typically includes files created by e-mail systems, word processing applications, presentations, photos and many others.

Protocol. A document outlining the formats and standards to be used in the preparation and exchange of electronic Schedule "A". A protocol may be agreed to by the parties or imposed by order of the court.

Record. A row in a table of structured information. An entry in the electronic Schedule A. Each row or record corresponds to a single document or individual electronic file or e-mail message.

Redaction. The process by which privileged or irrelevant information in a copy of an otherwise producible document or EI is electronically masked.

Tab Delimited. A structured text file where entries for each field are separated or "delimited" by a tab character and entries for records are separated by a new line (carriage return <CR>).

Table. Information presented in a structured format so that each row represents a record and each column represents a field (data element).

TIFF. The abbreviation for Tagged Image File Format which is a common format for Image files.

Unicode. A standard for encoding characters, more robust than ASCII. Can represent more than 107,000 characters, including foreign languages and mathematical or scientific symbols. Popularly implemented in UTF-8 encoding, which is backward compatible with ASCII.

Unitization. The process of separating and grouping individual documents or EI.