Federal Court of Australia

Practice Note No 17: Example of an Advanced Document Management Protocol

1.   Purpose of this Document

1.1   This is an example of an Advanced Document Management Protocol prepared in accordance with Practice Note No 17. It is current as at 29 January 2009.

1.2   The Protocol sets out the agreement of the parties in the matter of *(insert name)* in relation to the scope, means and format in which both Paper Documents and Electronic Documents are to be:

- exchanged between the parties during the discovery process; and

- delivered to the Court for inclusion in the Court Book.

2.   Document Descriptions

2.1   All Documents to be exchanged between the parties and delivered to the Court will be described in a List of Documents containing the following information for each Document:

(a)   Document ID *(see Schedule 1 for details)*

(b)   Document Title

(c)   Document Type *(see Schedule 7 for details)*

(d)   Document Date

(e)   Author *(see Schedule 2 for details)*

(f)   Recipient *(see Schedule 2 for details)*

(g)   Host Document ID *(see Schedule 3 for details)*

(h)   Folder and Filename[1] *(Refer Schedule 4 for details)*

2.2   In addition to the mandatory information outlined above, the parties may agree to include further descriptive information in the List of Documents, for example:-

(a)   Redacted *(to indicate whether or not file has been redacted, values may be 'Yes', 'No' or blank)*

(b)   Privileged *(to indicate whether the whole or part of the Document is subject to a claim of privilege, values may be 'Yes', 'No', 'Part' or blank)*

(c)   Privilege Reason *(to indicate basis upon which privilege claim is made)*

(d)   Confidential *(to indicate whether the whole or part of the Document is subject to a claim of confidentiality, values may be 'Yes', 'No', 'Part' or blank)*

(e)   Discovery Category *(where the parties have agreed or the court has ordered discovery by category)*

---

[1] This refers to the file name and folder of the renamed electronic image files (For example, PDF or TIFF or Native Electronic Documents) not the source path and name of the original file. This field is included in order to facilitate automated (and more efficient) compilation of an Electronic Court Book.

(f) Document Description *(for example, the parties may agree that certain fields can be combined to make a description field)*

(g) Estimated Date *('yes' if date is estimated, otherwise 'no')*

(h) Container Type *(see Schedule 9 for details)*

(i) Container Path *(see Schedule 9 for details)*

(j) Document Category *(see Schedule 9 for details)*

(k) File Type *(see Schedule 9 for details)*

3. Document Structure and Format

3.1 The List of Documents to be exchanged between the parties will be in accordance with the format and structure described in Schedule 6.

3.2 The Court may also require Documents to be delivered to the Court in the format and structure described in Schedule 6, for example, to facilitate an electronic trial. However, where no such request is made by the Court, the parties will deliver the List of Documents to the Court in the format specified in the Default Document Management Protocol.

3.3 Parties will avoid converting Native Electronic Documents to paper for exchange purposes and will instead exchange them as Searchable Images or, if agreed by the parties or ordered by the Court, as Native Electronic Documents.[2]

3.4 Where Documents are to be provided or exchanged as Searchable Images, Native Electronic Documents should be rendered *directly* to Portable Document Format (PDF) to create Searchable Images. They should not be printed to paper and scanned or rendered to Tagged Image File Format (TIFF) format and then converted to PDF. Rendering Native Electronic Documents directly to PDF will minimise the costs and avoid inaccuracies associated with the Optical Character Recognition (OCR) process.

3.5 Where Documents are to be provided or exchanged as Searchable Images, the parties may agree or the Court may order that the Documents also be exchanged or delivered as:

(a) Single Page TIFF Images;[3] and/or

(b) Native Electronic Documents.

3.6 Electronic Documents that do not lend themselves to conversion to PDF (for example, complex spreadsheets, databases, etc.) may be exchanged or delivered to the Court as Native Electronic Documents or in another format agreed by the parties and the Court.

3.7 Native Electronic Documents that are imaged files in their native form may be rendered with OCR to improve their searchability where this is technically possible and the parties agree that it is necessary.

3.8 Paper Documents should be exchanged as Searchable Images wherever practicable.[4] However, where the parties agree that OCR is unnecessary they may exchange Paper Documents as Unsearchable Images.

---

[2] Retaining Electronic Documents in an electronic form is always preferable to printing and/or photocopying Electronic Documents. Further, due to inherent support for Searchable Electronic Images and redaction, PDF is a preferable Electronic Image format to TIFF.

[3] See definition in the Glossary.

3.9 Colour versions of Documents will only be created if it will be of evidential significance to see the colour in a Document.

3.10 Blank, irrelevant pages will be removed where practicable and will not be allocated Page Number Labels.

3.11 Subject to this section, all Documents to be included in an Electronic Court Book will be delivered to the Court as Searchable Images.

4. Page Numbers

4.1 Subject to this section, a unique Page Number Label in the format described in Schedule 1 will be placed on each page of every Document as described in Schedule 5.

4.2 The Page Number assigned to the first page of a Document will also be assigned as the Document ID for that Document.

4.3 Native Electronic Documents will be assigned a single Document ID and individual Page Number Labels are not required.

5. Expert Reports & Statements

5.1 A Document that is an Expert Report or a Statement will be described, named, imaged, exchanged and delivered in the same way that Discoverable Documents are described, named, imaged, exchanged and delivered.

5.2 Each Expert Report and Statement should be exchanged in accordance with the timetable agreed by the parties or ordered by the Court:

(a) as Native Electronic Documents (for example, in Microsoft Word format); and

(b) as Searchable Images retaining any court applied markings such as signatures, stamps, and annotations.

5.3 Where a Document is referred to in an Expert Report or Statement, the references will, wherever possible, be to the Document ID for the Document.

5.4 If a Searchable Image of a Document referred to in an Export Report or Statement is available, this will, wherever possible, be hyperlinked to each reference to the Document ID of the Document in the Expert Report or Statement. If no Searchable Image of the Document is available, the hyperlink will, wherever possible, link to an Unsearchable Image of the Document.

5.5 An Expert Report will be described with a Document Type of 'Expert Report'.

5.6 A Statement will be described with a Document Type of 'Statement'.

5.7 Each exhibit to a Statement or Expert Report will be treated as an Attached Document in relation to the Host Document (the Statement or Expert Report).

---

[4] As mentioned in Practice Note No 17, the Court has a strong preference for Searchable Images, or, where this is not feasible, for Unsearchable Images. This is because of the increased functionality available with Electronic Images as compared with Paper Documents and due to the relative high costs associated with photocopying documents multiple times when compared with the cost of converting them to an Electronic Image once.

6. Submissions

6.1 Each submission will be described, named, imaged, exchanged and delivered in the same way that Discoverable Documents are described, named, imaged, exchanged and delivered.

6.2 Each Submission should be exchanged in accordance with the timetable agreed by the parties or ordered by the Court:

(a) as Native Electronic Documents (for example, in Microsoft Word format); and

(b) as Searchable Images retaining any court applied markings such as signatures, stamps, and annotations.

6.3 Where a Document is referred to in a Submission, the references will be to the Document ID for the Document.

6.4 If a Searchable Image of a Document referred to in a Submission is available, the Submission should include a hyperlink to the Document ID for that Document. If no Searchable Image of the Document is available, the hyperlink should, wherever possible, link to an Unsearchable Image of that Document.

7. Minutes of Meetings

7.1 A Document that has not been sent from an Author to one or more Recipients but has been distributed or tabled as a summary of the proceedings or the outcomes of a meeting will have the attendees at that meeting recorded as Recipients for the Document.

7.2 The Author field for such a Document may remain blank.

7.3 Such a Document will generally be classified with a Document Type of 'Minutes of Meeting'.

8. Agreements and Contracts

8.1 A Document that constitutes, or is claimed to constitute, an agreement or contract between two or more parties will have each party to the agreement or contract recorded as a Recipient of the Document.

8.2 The Author field for such a Document may remain blank.

8.3 Such a Document will generally be classified with a Document Type of 'Agreement' or 'Contract'.

9. Electronic Exchange Media

9.1 Unless otherwise agreed or ordered by the Court, the information to be exchanged between the parties and delivered to the Court will be contained on read-only optical media (for example, CD-ROM, DVD-ROM) or portable hard drive.

9.2 Where portable hard drives are used, they will be returned to the supplying party as soon as the data has been copied by the recipient party.

10. Data Security

10.1 A party producing data to another party will take reasonable steps to ensure that the data is useable and is not infected by Malicious Software.

10.2 Notwithstanding paragraph 10.1, the onus is on each party receiving the data to test the contents of any exchange media prior to its use to ensure that the data does not contain Malicious Software.

10.3 If data is found to be corrupted, infected by Malicious Software or is otherwise unusable, the producing party will, within 2 working days of receipt of a written request from a receiving party, provide to the receiving party a copy of the data that is not corrupted, infected by Malicious Software or otherwise unusable (as the case may be).

11. Errors in exchanged documents

11.1 If errors are found in any exchanged Document, the producing party must provide a corrected version of the Document to the receiving party.

11.2 If errors are found in more than 25% of the exchanged Documents, the producing party must, if requested by the receiving party, provide a correct version of all Documents to the receiving party.

11.3 In addition to the requirements of paragraphs 11.1 and 11.2, if errors are found in any exchanged Document a written explanation will also be sent to each receiving party setting out the reasons for the errors in the Documents and describing the data affected.

12. Redaction for Privileged or Confidential Documents

12.1. If the whole or part of a Document is subject to a claim of privilege or confidentiality, the parts of the Document that are subject to the claim should be identified or, if appropriate, Redacted pending determination of the claim. If the whole or part of the Document is Redacted, the party producing the Document must retain an unredacted version of the Document which must be produced to the Court if required to do so.

12.2. If the Court makes an order that the whole or part of a Document is subject to privilege, the copy of the Document to be exchanged between the parties and provided to the Court may be permanently Redacted in accordance with that order.

12.3. If the Court makes an order that the whole or part of a Document is confidential, arrangements will be made to ensure that access to the Document, or to the confidential parts of the Document, is restricted in accordance with that order.

12.4. If the whole or part of a Document is subject to a claim of privilege or confidentiality it will be:

   (a)   allocated a Document ID;

   (b)   given a Document Description that does not disclose the information that is the subject of the claim of privilege or confidentiality; and

   (c)   if the claim of privilege or confidentiality relates to the whole Document – represented by a single Placeholder Page with the words 'Document subject to claim of privilege/confidentiality' inserted under the Document ID.

12.5. If the whole or part of a Host Document is subject to a claim of privilege or confidentiality it will be:

    (a)   identified as a Host Document;

    (b)   allocated a Document ID;

    (c)   given a Document Description that does not disclose the information that is the subject of the claim of privilege or confidentiality; and

    (d)   if the claim of privilege or confidentiality relates to the whole Document – represented in the Document Group to which it belongs by a single Placeholder Page with the words 'Document subject to claim of privilege/confidentiality' inserted under the Document ID. .

12.6. If the whole or part of an Attached Document is subject to a claim of privilege or confidentiality it will be:

    (a)   identified as an Attached Document;

    (b)   allocated a Document ID;

    (c)   given a Document Description that does not disclose the information that is the subject of the claim of privilege or confidentiality; and

    (d)   if the claim of privilege or confidentiality relates to the whole Document – represented in the Document Group to which it belongs by a single Placeholder Page with the words 'Document subject to claim of privilege/confidentiality' inserted under the Document ID..

13.   De-Duplication of Documents

13.1   Where appropriate, each party will take reasonable steps to ensure that duplicated Documents are removed from the exchanged material ('De-Duplication').

13.2   However, the Court acknowledges that there may be circumstances where Duplicates need to be identified and retained for evidential purposes.[5]

13.3   Duplication will be considered at a Document Group level. That is, all the Documents within a Document Group (that is, a Host Document and Attached Documents) will be treated as Duplicates if the *entire* Document Group is duplicated elsewhere within the collection. An Attached Document in a Document Group will not be treated as a duplicate if it is merely duplicated elsewhere as an individual, stand alone Document that is not associated with another Document Group.

13.4   The method of de-duplication is described in Schedule 8.

---

[5] For example, it may be relevant to retain multiple copies of an email in sender and recipient email boxes due to the fact that it will be of evidential relevance to know who actually received the email after it was sent.

Schedule 1 – Document IDs and Page Numbers

1.1. Document IDs and Page Numbers will be *unique* because it is the sole means by which Documents will be referenced

1.2. Document IDs and Page Numbers will be in the following format

SSS.*BBB***.FFFF.NNNN_***XX***   (italics represent optional elements)*

1.3. This format is described in the table below.

Level	Description
SSS	The Party Code (also, often referred to as 'Source') identifies a party to the proceedings. It should comprise three alpha characters. The determination of the Party Codes to be used for a particular case will take place prior to the commencement of discovery in order to ensure that all Document IDs will be unique (i.e. to ensure that no two documents have the same Document ID so that each Document can be uniquely referenced). Refer to Schedule 1.4 for the list of available Party Codes.
BBB	The Box Number identifies a specific physical archive box, email mailbox or any other Container or physical or virtual classification that is appropriate for the party to use. Use of the Box Number is optional. The box number should comprise 3 digits
FFF	The Folder Number identifies a unique folder number allocated by each party in their own Document collection.[6] The Folder Number should be padded with zeros to consistently result in a 3 digit structure. The Folder Number may, where appropriate, correspond to the Box Number of any Container in which the Document is contained.
NNNN	This refers to each individual page within each Folder for Paper Documents, Unsearchable Images and Searchable Images. For Native Electronic Documents, this number applies to the whole Document irrespective of the number of pages within it. In such cases, it therefore operates as a Document Number rather than a Page Number because individual pages are not numbered. This number is padded with zeros to consistently result in a 4 digit structure.
_XX	This number is optional and is only required where additional pages need to be inserted into a Document. A suffix will be used, preceded by an underscore, padded with zeros to consistently result in a 2 digit structure.

---

[6] A Party may allocate loose or unsorted Documents, either hard-copy or electronic, to 1 or more folders. This is acceptable providing that the originals of such Documents are able to be promptly sourced for inspection if required. It may also identify an electronic folder (as part of a directory structure) or a folder within an e-mail mailbox.

Depending on the volume, format and structure of the material that is to be discovered, the parties may agree to a different Document ID format. For example, for small Document collections parties may wish to omit the box level while for larger collections the Folder Number may be increased in length to 4 digits (FFFF). For large electronic collections, the Page Number level may be increased to 5 (NNNNN) or 6 (NNNNNN) digits. Parties may also agree to use different Party Codes for specific material that has been prepared for the purposes of the litigation (such as Statements, Expert Reports and Submissions).

1.4. Party Codes for the Document ID

For the purposes of the Document ID, the following Party Codes are available.

(Insert table here tailored to suit the case at hand, for example...)

Party Code	Party
ASC	Australian Securities and Investments Commission (Applicant)
ABC	AB Corporation Pty Ltd (First Respondent)
XYH	XY Holdings Pty Ltd (Second Respondent)

Schedule 2 – Describing People

2.1. People names may be referenced using:

    (a)   email addresses (for example, jcitizen@abc.com.au); or

    (b)   Surname [space] Initial (for example, Citizen J) where email addresses are not available; or

    (c)   by reference to a position (for example, Marketing Manager) where email addresses and Surname, Initial is not available; or

    (d)   by reference to an organisation associated with the person where email address, Surname, Initial and Position are not available.

2.2. Multiple Recipients will be entered as separate rows in the Parties Table.

Schedule 3 – Document Hosts and Attachments[7]

3.1　Every Document that is attached to or embedded within another Document will be called an Attached Document.

3.2　A Container is not a Host Document for the purposes of this Protocol.[8]

3.3　Attached Documents will have the Document ID of their Host Document in the descriptive field called 'Host Document ID'.

3.4　Host Documents and Attached Documents are jointly referred to as a 'Document Group'.

3.5　Subject to paragraphs 3.6 and 3.7 below, in a Document Group the Host Document will be immediately followed by each Attached Document in the order in which the Attached Documents are numbered in their Document ID. If a Document Group includes Documents that are subject to a claim of privilege or confidentiality, the Documents should be treated in accordance with Section 12 of this Protocol.

3.6　If a Document is contained within a Container (for example, a single ZIP file) that is attached to an email then the email should be treated as the Host Document and the Document in the Container should be treated as an Attached Document to that Host Document (that is, the Host Document will be the email and not the Container within which the Document is contained).

3.7　If the Document Group consists of a number of Paper Documents fastened together, the first Document will be treated as the Host Document and the remaining Documents will be treated as the Attached Documents within the Document Group unless those Documents are not related, in which case each Document will be treated as a separate Document without a Host Document.

3.8　Annexures, Attachments and Schedules that are attached to an Agreement, Report, Legal Document or Minutes of a Meeting may be described as separate Attached Documents associated with the relevant Host Document.

---

[7] May be referred to as Document Delimiting.

[8] See the Glossary to Practice Note No 17 and Related Materials for further information on Host Documents and Containers.

Schedule 4 – Electronic Folders and Filenames

4.1   This schedule specifies how Electronic Images are to be located and named for the purposes of Document exchange. It does not relate to the capture and exchange of the original source location of an Electronic Document.

4.2   The Folder containing all Documents will be named either '\Documents\' or '\Images\'[9]

4.3   Documents produced as Searchable Images will be named 'DocumentID.pdf'

4.4   Documents produced as multiple Single Page TIFF Images will have each TIFF image file named 'PageID.TIFF'.

4.5   Documents produced as Native Electronic Documents will be named 'DocumentID.xxx(x)' where 'xxx(x)' is the original default file extension typically assigned to source Native Electronic Files of that type.[10]

4.6   The Documents folder will be structured in accordance with the Document ID hierarchy, for example:

The Document produced as a Searchable Image called 'ABC.001.0004.00392.pdf' would be located in the folder called '*Documents\ABC\001\0004*'. So, it will appear in the directory listing as '*Documents\ABC\OO1\0004\ABC.001.0004.00392.pdf*'.

Where this same Document has also been produced as many Single Page TIFF Images, the second page will be called 'ABC.OO1.0004.00393.TIFF' and will be located in the folder called '*Documents\ABC\001\0004*'. So, it will appear in the directory listing as '*Documents\ABC\001\0004\ABC.OO1.0004.00393.TIFF*'

Where this same Document has been produced as a Native Electronic Document, and, assuming it is a Microsoft Excel spreadsheet file, for example, it would be called 'ABC.OO1.0004.00392.xls' and will be located in the folder called '*Documents\ABC\001\0004*'. So it will appear in the directory listing as '*Documents\ABC\001\0004\ABC.001.0004.00392.xls*'

---

[9] The term 'Images' is becoming increasingly obsolete in light of the increasing trend for Documents to be exchanged as Native Electronic Documents rather than as Images. This directory name also reflects the requirements of a proprietary commercial application so it may be desirable to replace it with the more neutral and contemporary term '\Documents\'.

[10] For example, Microsoft Word documents will have a '.doc' extension, Microsoft Excel spreadsheets will have a '.xls' extension, so Native Electronic Documents will be named along the following lines *ABC.001.003.0456.xls (Excel Spreadsheet), XYZ.099.456.0093.doc (Word Document)* A four character extension may be required for particular file types.

Schedule 5 – Page Number Labels

5.1 Wherever possible, Page Number Labels will be placed on the top right corner[11] at least 3 millimetres from both edges of the page

5.2 If there is insufficient space for a Page Number Label on a Searchable Image or an Unsearchable Image, the electronic image of the page will, if possible, be reduced in size to make room for the Page Number Label.

5.3 Page Number Labels may also include machine readable barcodes.

5.4 Where feasible, landscape pages of Searchable Images, Unsearchable Images and Paper Documents should be positioned so that the title is on the left side of the page[12] and the Label is oriented to the text, preferably at the bottom right corner of the original page so it appears down the top right side edge of the rotated page. .

5.5 The parties may apply Page Number Labels to the following Paper Documents where they contain relevant content:

   (a) folder covers, spines, separator sheets and dividers

   (b) hanging file labels

   (c) the reverse pages of any Document

5.6 Adhesive notes should not normally be labelled but should be scanned in place on the page to which they were attached. If this cannot be done without obscuring text, the adhesive note should be numbered as the page after the page to which it was attached and the page should be scanned twice – first with and then without the adhesive note.

---

[11] This ensures that upon electronic retrieval, images will not need to be scrolled down manually on the screen in order to view the Page Number Label.

[12] This generally involves a 90 degree anti-clockwise rotation.

Schedule 6 – Document Descriptions

Document Descriptions are to be structured in the following tables in Microsoft Access Database format.

Table Name	Table Description
Export	Main Document information
Parties	People and organisation information for each Document
Pages	Listing of electronic image filenames for each Document
Export_Extras	Additional data fields for each Document

Export Table

Field	Data Type	Explanation – Document Types and Coding Method and possible values	
Document_ID	Text, 255	Document ID in accordance with Schedule 1.	
Document_Type	Text, 255		
		Paper Documents	Refer Document Types in Schedule 7.
		Electronic Documents (including email, email attachments, loose files etc)	Either Native File Type in Schedule 9 or Document Type in Schedule 7 as determined on the basis of the face of the Document (parties to agree on the approach to be used)
Document_Date	Date, 11	DD-MMM-YYYY	
		Paper Documents	Determined on the basis of the Date appearing on the face of the Document
		Undated Documents	Leave field blank
		Incomplete Date (Year Only)	For example, 01-JAN-1900
		Incomplete Date (Month and Year Only or Day and Month only	For example, 01-MMM-YYYY, DD-MMM-1900

Field	Data Type	Explanation – Document Types and Coding Method and possible values	
		emails	Electronic Metadata – Sent Date [13]
		Unsent emails	Last Saved Date
		Other Electronic Documents	Electronic Metadata – Last Saved Date; or Date appearing on the face of the Document (parties to agree on the approach to be used)
Estimated	Text, 3	Yes OR No OR Blank	
		Default	No or Blank
		Undated Documents	No or Blank
		Incomplete Date	Yes
Host_Reference	Text, 255	If the Document is an Attachment, this field contains the Document ID of its Host Document. Please refer to Schedule 3.	
Title	Text, 255	Paper Documents	Determined on the basis of the title appearing on the face of the Document
		Email	Subject Field
		Other Electronic Documents	Electronic Metadata – File Name or determined on the basis of the Title appearing on the face of the Document
Level_1		The Party level of the Document ID (see Schedule 1)	
Level_2		The Box level of the Document ID (see Schedule 1)	
Level_3		The Folder level of the Document ID (see Schedule 1) under which the Searchable Images or Native Electronic Documents are stored.	

Parties Table

This table holds the names of people associated with a particular Document and their relationship to the Document.   It may also hold organisation information for these people. There is a one-to-many relationship between the Export table containing the primary Document information and the Parties table because multiple people could be associated with a single Document.

---

[13] The concept of time zones can be difficult to manage where emails are sent from one location and time zone and received in many different locations and time zones. The emerging convention seems to be to record the time zone of the server that sent the e-mail in the primary date field for an email. The received date associated with the local email server for the recipient of a 'Duplicated' e-mail may also be captured in other metadata date fields (that is, other than the primary Date field). New conventions are likely to emerge in this area over time.

Field	Data Type	Explanation	
Document_ID	Text, 255	Document ID in accordance with Schedule 1.	
Correspondence_Type	Text, 100	Correspondence Type (Sent or Received)	
		Paper Documents	AUTHOR, RECIPIENT BETWEEN, ATTENDEES, CC To be determined on the basis of the face of the Document.
		emails	FROM, TO, CC, BCC
		Other Electronic Documents	AUTHOR, RECIPIENT, BETWEEN, ATTENDEES, CC To be determined on the basis of the face of the Document.
Organisations	Text, 255		
		Paper Documents	Name of organisation the produced the Document as determined on the basis of the face of the Document.
		emails	Parties may: agree to keep this field blank; or use automated manipulation of data after automated Metadata Extraction although this is not always reliable (email domain is not always an indicator); or determine the field on the basis of the face of the Document.
		Other Electronic Documents	To be determined on the basis of the face of the Document.
Persons	Text, 255	Please refer to [Schedule 2 – Describing People].	
		Paper Documents	To be determined on the basis of the face of the Document.
		emails	Electronic Metadata – email addresses or email alias names.
		Other Electronic Documents	Electronic Metadata – Author value or may be determined on the basis of

Field	Data Type	Explanation
		the face of the Document (parties to agree approach)

Pages Table

There will be an entry in the Pages table for every TIFF page or PDF document that relates to a single Document in the Export table. i.e. There is a one to many relationship between the Export table and the Parties table. Where Native Electronic Documents only are exchanged (no TIFF, PDF files or placeholder pages), there will only be one entry in the pages table corresponding to each Native Electronic Document.

Field	Data Type	Explanation
Document_ID	Text, 255	Document ID
File_Name	Text, 128	Filename, including extension of each indexed Document
Page_Label	Text, 32	Page number plus file extension (for example, 'ABC.001.023.pdf'). Where single page TIFF files are exchanged no file extension is necessary.
Page_Num	Number, Double	An integer indicating the order in which the files related to the Document ID should be sequenced when viewing the full Document. For example, if a TIFF page is available for each page, the record in this table associated with the first TIFF page should have this value equal to '1'. The record in this table associated with the second page will have this value set to '2' and so on. A multi-page PDF file will have this value set to the value of the last TIFF page plus 1 (where TIFF pages have been produced) or set to '1' if no TIFF pages were produced. A Native File will have a value equivalent to the PDF file value plus one unless there is no PDF file in which case it will receive a value of '1'.

Export_Extras Table[14]

This table holds any additional metadata the parties wish to exchange that is not held in the other three tables mentioned above.

Field	Data Type	Explanation
Document_ID	Text, 255	Unique Document Identifier (Document ID)
theCategory	Text, 50	Text OR Date OR Numb OR Bool OR Pick
theLabel	Text, 255	Custom Field Name
theValue	Text, 255	Custom Field Contents

---

[14] Where the parties agree, an 'Electronic Document Source' field will be included where possible to specify the original source directory and filename of the original electronic Document.

Field	Data Type	Explanation
9999Value	Number, Double	If Numb, Custom Field Data
textValue	Text, 255	If Text OR Pick, Custom Field Data
boolValue	Number, Double	If Bool, Custom Field Data

By way of example only, the parties may agree to include custom fields in the Export Extras table such as those below:

- Privilege Status

- Basis for Privilege

- Confidentiality Status

- Redaction Status

- Reason for Redaction (Privileged, Irrelevant, Confidential)

- Group (Host, Attachment, Loose)

- Path and Name of Original File *(i.e. Original location prior to collection and processing)*

- Document Objective Title (i.e. The File Name is not an indication of the content in most electronic files so the parties may agree that this should be objectively coded and exchanged as a separate field)

- Electronic Document Source

- MD5 Hash Values

- Document Container Path (if the Document was 'contained' within an electronic folder or directory or a compressed electronic file such as a ZIP file – see Schedule 9)

- Document Container Category (if the Document is a 'container' – see Schedule 9)

- Document Category (see Schedule 9)

- Native File Type (see Schedule 9)

Schedule 7 – Document Type List

The following table should be completed in accordance with the particular needs of the case. As an alternative to the Document Type values set out in the table, parties may agree to use File Type values such as those mentioned in Schedule 9.

Document Type	Description
Agreement	
Affidavit	
Annual Report	
Article	
Authority	
Board Papers	
Brochure	
Cheque Remittance	
Company Search	
Contract	
Court Document	
CV	
Diagram – Plan	
Diary Entry	
Drawing	
Expert Report	
Fax Transmission	
File Note	
Form	
Invoice – Statement	
Letter	
List	
Manual	
Map	
Meeting Agenda	
Memorandum	
Minutes of Meeting	
Notice	
Photograph	

Document Type	Description
Receipt	
Report	
RFI – RFO	
Search	
Specification	
Spreadsheet	
Statement	
Submissions	
Timesheet	

Schedule 8 – De-duplication Methodology

8.1   The parties will use MD5 hash values to identify and, where appropriate, remove Duplicates from their exchanged Document collections based on the approach agreed during the Pre-Discovery Conference.

8.2   The Metadata fields to be used to generate the MD5 hash value for emails are 'Sender', 'To' and 'Date Sent', 'Body' and 'Number of Attachments' (or MD5 hash values of Attachments).[15]

8.3   MD5 hash values will be stored in the export extras table.

---

[15] There is a general trend to simply use the fields 'Sender', 'To' and 'Date Sent' for de-duplication however, the additional field 'Number of Attachments' is recommended to address the potential problem associated with 'Sent' times being rounded to minutes rather than seconds by some e-mail servers.   On such servers it would be possible for the same author to send two entirely different emails to the same recipients at what appears to be the same time.

Schedule 9 – Document Containers, Categories and File Types

Each of the following tables should be completed in accordance with the needs of the case.

Container Type	Container Description
Directory	An electronic folder or directory on a computer file system that contains electronic files. This includes folders or directories inside an email store.
Compressed File	An Electronic Document that contains one or more compressed Electronic Documents that are considered 'Documents' in their own right and may be extracted to their original size.
Email Store	A single file (or 'email box') containing multiple emails, email attachments and other items such as diary appointments and tasks. The most common email store file types are PST and NSF files.
Diary	A paper based personal calendar containing multiple appointments over a period of time.
Folder	A Physical, Hard Copy Binder containing multiple Paper Documents.
Database	An electronic file that contains data that may be considered to be multiple Documents.
Etc.	

Document Category	Description
Email	An email – usually contained within an email store (e.g. an email box) but may be extracted to reside within a directory or folder on a file system.
Email Attachment	An Electronic Document attached to an email.
Loose File	An Electronic File that is not attached to an email but rather resided in its original state in a directory on a file system.
Paper	A Document that is in paper format in its original state (where the electronic version of the Document is not available).
Diary	A single entry in a Diary or Calendar. For example, Appointment, Meeting etc.
Task	A single task on an electronic To-Do list – usually contained in an email box.

File Type	Description
Wordprocessing	For example, a Microsoft Word Document with a '.doc' or a '.docx' extension or a Wordperfect Document with a '.wpd' extension etc.
Spreadsheet	For example, A Microsoft Excel Document with an xls extension
Presentation	For example, a Microsoft Powerpoint file with a .ppt extension
Database	A database file. For example, a Microsoft Access file with an .mdb extension
Picture	A file containing a graphic or image such as a file with a '.jpeg' or '.bmp' or a '.tiff' extension
PDF	A file with a .pdf extension
Email	An email that has been extracted from its email store. For example a file with a .msg or .eml extension etc..
Music	
Text	For example, a file with a .txt extension
Web	For example, a file with a .html or .htm extension
Wordprocessing Template	For example, a file with a .dot extension
Compressed	For example, a file with a .zip extension