- ***These Guidelines should be used to assist the parties in agreeing the various elements of the eDisclosure Exchange Protocol.***
- ***It is recommended that the parties start the process of agreeing the eDisclosure Exchange Protocol in sufficient time before the finalisation date to enable the production of material to the required specifications. If the matter falls within the remit of the Disclosure Pilot for the Business and Property Courts in England and Wales, that came into force on 1st January 2019, it is recommended that this Protocol should, so far as possible, be finalised with the completed Disclosure Review Document; that is to say 5 days before the first CMC. If this is not possible, it should be finalised no less than 7 days prior to the date of exchange.***
- ***As the draft eDisclosure Exchange Protocol may be extensively amended from case to case, it is referred to in these Guidelines as the "template version of the Protocol".***
- ***Some of the elements of the eDisclosure Exchange Protocol will be determined by the functionality of your litigation support system. If this is provided by an external supplier, make sure you involve them in the drafting process.***
- ***The level one heading numbering in this Guidelines Document, mirror those used in the template version of the Protocol.***
- ***Please check the ILTA site to ensure you have the latest version of this document. Please email any feedback to iltaprotocol@iltanet.org.***

1 MATTER DETAILS

1.1 Use the matter specific table to complete the required information. As well as the core information, a row has been provided to record any additional significant dates or milestones within the project. This is optional.

1.2 It is envisaged that the template version of the protocol might be used in two main ways; first, as an individual document by one party to articulate how its disclosure will be carried out, or second, as an agreed joint protocol that will be used by all parties in a matter.

1.3 The specifics of disclosure will be determined by the functionality of the litigation support packages used by the individual parties. Where all parties use the same product, agreeing common standards should be easier, though there might still be differences in the available metadata, depending upon decisions taken by the legal teams at the start of the project. Where parties use different software products, there might be areas that cannot be agreed because the underpinning software does not support the specific requirement. In such cases, each party might draft its own variant of an agreed common document, with differences as needed, or agree on a single document, with specific elements that reflect the capabilities of each party's software.

2 PROTOCOL COMPONENTS

2.1 The template version of the protocol is split into three discrete areas. These contain the details on the specifics of the document and are:

- The Disclosure List; what is going to be exchanged,
- The Disclosure Data; the format of the exchanged data

- The Disclosure Load File; the technical requirements to load the data into a litigation support system.

2.2 It is envisaged that the legal team will be involved in agreeing the components of sections 1 and 2, with the more specific technical information being covered in the final area. For parties operating without a litigation support system, the final section is superfluous.

3 DISCLOSURE LIST

3.1 The template version of the protocol contains a paragraph describing the approach to producing the list. A choice here is what format the list itself should be provided it. The options shown in the template version of the protocol are either an Excel spreadsheet a Word table or a CSV file. The preferred default, from a technical viewpoint, is an Excel spreadsheet, as this is the easiest way to manipulate the data exported from a litigation support system. Some legal teams have a preference for a Word table, hence its inclusion as an option, but normally this involves additional effort to produce and is not as user friendly as a spreadsheet. Electronic formats, such as a PDF rendition of a spreadsheet should be rejected as these do not allow for any data manipulation on receipt of the document. A CSV file is the least useful format, but is better than a PDF.

3.2 There is a “Terms” section which provides descriptions for the key definitions used within the rest of the template version of the protocol. These are for; email, electronically stored information (ESI), native, metadata, image and text. Effectively anything that is electronic data, and not email, is referred to as ESI. Some examples of specific ESI formats are provided, as the range of data collection expands, you might want to add some matter specific formats in here.

3.3 There can be debate on the composition of a native email, which is why a definition of this has been included. If your litigation support system/supplier/in-house team have a different approach, then it should be reflected here. Specifically, if your supplier can provide a “native” version of a parent email which **does not** contain the attachments to the email, then this should be reflected in this paragraph by deleting the word [Attachments]. This is a highly significant piece of functionality that is discussed in much greater detail in the section below on email families (4.12).

3.4 There is then a table to show the data that will be supplied within the disclosure list. The table consists of two columns; the field name and then a description of that field. The use of field names is discussed in the following paragraph. The data that is available to be shared in this list will vary from product to product. The aim of the template version of the protocol is to give a default set of fields, and then leave the rest of the table to individual users, the capabilities of their litigation support software and the demands of the specific matter.

3.5 **Use of field names.** The field names used in the template version of the protocol will depend upon the name for that field used by your litigation support software, and will need to be edited to reflect what they are. For example, the first row in the table refers to the unique number applied to the first page of a document when it is produced. In different litigation support systems this field is called different things, for example:

- ProductionBegDoc

- Beg Bates
- Bates Start

3.6 **Unique Document Number | Proposed Format.** The final sentence of the descriptive text in the first box in the table is in bold and should be used to give an example of the proposed format. The text prefix (shown as XXX) can be whatever your standard is. Some organisations use the initials of the name of the law firm supplying the disclosure; e.g. in a matter between Squire Patton Boggs (UK) LLP and RPC (Reynolds Porter Chamberlain), the disclosure from Squires might be labelled SPB000001 onwards, and from RPC RPC000001 onwards. If your firm only uses two letters instead of three, that is not an issue. Other organisations prefer to use the name/initials of the end client. Others will include a geographic reference such as UK, FRA, US. See paragraphs 4.9 - 4.10 below for a discussion on page numbering within individual items.

3.7 Edit the list so that it reflects the names used by your system when it produces an export file. You and the other parties will map the opposing sides disclosure onto your own fields, so it doesn't matter, per se, what they are called, so long as they accurately reflect the text labels used in your system.

3.8 **Dates | Family Date.** It is a requirement of the template version of the protocol that the data in the disclosure list should be sorted in chronological order. This is achieved by having a date field for all of the documents in an email/family group in addition to their own date information, so that all of the items in a group can be kept together in a chronological order.

3.9 Different systems/suppliers call this field different names. Normally, it is automatically generated when an email is imported into a litigation support system so that all attachments to an email, have the same date as the parent email, irrespective of when the attachments were created / last modified.

3.10 The Disclosure list needs to reflect this date in whatever format and name your system provides it in.

3.11 **Dates | Other Dates.** Different organisations have different approaches to providing / receiving other date information that might be contained within the metadata of native files, such as Date Created, Sent, Last Modified, and Last Printed. If an item is not part of an email family, and thus does not have a Family Date value, the field "Date Last Modified" is normally used for sorting purposes. Therefore it is not unusual for this information to also be requested. This is why there are a second pair of date field rows within the Disclosure list table.

3.12 **Date Format.** Some systems provide dates in a combination of Date and Time, others separate this data into two fields; one for the Date, the second for Time. Dates should always be in a day / month sequence. The reason for being specific about this is that the majority of litigation support software is developed in the US, were their default date sequence is month followed by day and then year. This shouldn't be an issue for either suppliers or your own systems, but it is worth defining it just to guard against technical slips in production.

3.13 When there is no date information, the normal default is to leave the field blank, however there might be systems that do not accept blank date fields, or occasions where you want to ensure the undated documents appear at either the start or end of the list. In these circumstances the documents are given a value that will ensure they are at the top or bottom of the list, for example 01/01/1900 or 31/12/2050 You don't have to use these precise dates, just some combination of Day, Month, Year that places the documents where you want them on the disclosure list.

3.14 There is a protocol for assigning objective information to hard copy material that might also be encountered in this topic.

3.15 When objective data is coded for hard copy material, the normal protocol for partially or undated information is to 1) show that the date is an estimated one via a tag, and 2) follow an agreed protocol for missing dates of:

- No day = 01/MM/YYYY
- No month = 01/01/YYYY
- No year = 01/01/1900 or 01/01/9999 or something similar

3.16 This approach is not needed for emails / ESI, as litigation support systems can accept a blank field for a date and automatically put any blanks at the top or bottom of any date list depending on its sort order (ascending = blank dates first, descending = blank dates last).

3.17 If you are asked to code email / ESI blank dates as 01/01/1900 or something similar, you would normally reject that request, citing the previous paragraphs as your reason, unless you wish to adopt the approach cited in 3.13 above.

3.18 **Core Data | Subject / File Name.** The remaining fields are those which you would normally expect to exchange. There are two fields to reflect the "subject" of an item. For emails this is the data held in the Subject field, for all other ESI, this is the file name. Some suppliers like to combine this information into a single field, on the basis that the file name for an email is often superfluous and incomprehensible, and that other ESI will not have a value for the Subject. Either approach is valid, though the evolving best practice seems to be supplying a single field. You should discuss your approach with your litigation support supplier.

3.19 **Core Data | From / Author.** As with the previous field, this field contains data from one of two sources. For emails is it the From information, for other ESI, it is the Author data.

3.20 **Optional Fields.** These fields will depend upon what your system can provide, and any agreements with the other parties.

3.21 **Optional Fields | Issues.** Bear in mind that the Disclosure Pilot for the Business & Property Courts in England and Wales, that came into force on 1st January 2019, is very focused on the issues of a matter. It is to be expected that most legal teams will code the relevant items in an electronic support system with the issues that they refer

to. Exchanging this information would seem to offer a mutual benefit to all parties in a matter. If this is the case, then the issue codes would form one of the optional fields being listed. If this is an option that is being considered, then you need to confirm with the legal team whether the text used to describe an issue within the eDisclosure system needs to be a set of mutually agreed words between all involved parties.

3.22 **Optional Fields | Relevance.** See below for the discussion on production of family email groups and how this field might be used.

3.23 **Optional Fields | Redacted.** A Yes / No / Blank field that shows if an item is redacted. There might be an additional request from the opposing party to show **why** an item has been redacted; Privilege, Commercially Sensitive, Personal Information. This level of detail can be a contentious issue and should be discussed in advance with the legal team.

3.24 **Optional Fields | Placeholder.** A Yes / No / Blank field that shows if an item is a placeholder image. See the discussion below on email families to understand when this field might be employed.

3.25 **Optional Fields | Others.** There are a number of other fields that have been suggested by suppliers. Rather than provide a list of all possible options, the template version of the protocol and these Guidelines acknowledge that they exist and can be added as agreed between the parties.

3.26 **Redaction of Metadata :** There are circumstance where the metadata associated with items might be redacted. These are explored below. You might show this by adding a column to the metadata table to indicate with a Y/N that some or all of the information might be redacted. Alternatively you might have a discrete paragraph at the end of table to explain the circumstances under which redactions might be applied to the metadata.

If an item is wholly privileged (legal advice OR Litigation) and a placeholder has been provided. (See 4.1)	Remove metadata e.g. titles / date / MD5 hash / Author / Email from / Email to / CC / BCC / Redacted (information that will be viewable in the review platform); the title can be altered to Withheld for Privilege.
If an item is part privileged and therefore, redaction has been carried out	Removal of metadata should be confirmed with the legal team – do not redact automatically.
If an item has commercially sensitive data within the body of the document	Removal of metadata should be confirmed with the legal team – these items may need to be referenced in legal correspondence.
If an item has a confidential custodian or a court agreed individual / entity	Removal of metadata should be as per the instruction from the legal team and the relevant court documentation.

4 ESI FORMATS

4.1 **Placeholders | Circumstances for use.** The options shown are mainly derived from the discussion below on family email groups. See 4.12 - 4.27 below for more information. There is also an option for when you have a file which you do not wish to supply in its native format, but it is impossible or impracticable to render it into an image. (For example media files with metadata you do not wish to exchange.)

4.2 There might be times when you use a placeholder to identify a privileged document on the basis that the privilege call might be over turned and you need to then supply the document in the correct location in the disclosure list. Explore if this scenario might apply with your legal team at the start of the matter.

4.3 If you can remove superfluous logo's / text from email address blocks / footers then there is no need to replace them with a placeholder.

4.4 **Placeholders | Format.** As ever this depends upon what your supplier/eDisclosure system can export / ingest and the same for the opposing party. Talk to your supplier/in-house support team. The format should not be an issue, it's the manner in which they are used which is the key question, with the actual text on the page a secondary matter.

4.5 **ESI Formats.** As earlier mentioned, the preferred disclosure format is native. However, there will be cases where a document may need to be produced as an image, normally because of the need to apply redactions, or where a document is a relevant parent that has irrelevant children. An image is the equivalent to taking a photo of the original document.

4.6 Disclosing a document in image format allows for the use of document stamping. It is sometimes seen that the unique document identification number and/or a message such as "Privilege and Confidential" are stamped on a document. This allows for easy reference to cover letters and indexes.

4.7 Not all document types are able to be rendered into images. An example of this is with media files. Where this is the case, a placeholder will need to be provided for the document if any information needs to be withheld.

4.8 File types such as Excel can pose formatting issues when being disclosed as an image due to the use of multiple sheets, comments and hidden columns to name a few. This can potentially be addressed by ensuring the correct settings are used for imaging Excels. It may also be a preference to have the Excel sheet resized to fit one page during imaging, though the resulting text might be rendered into a very small lfont an be difficult to read Some systems offer the ability to redact the native version of an Excel spreadsheet, discuss options with your supplier before you start turning spreadsheets into images.

4.9 There are two main formats in which images are disclosed, these are TIFF and PDF. If the image needs to be supplied in colour, then the JPEG format might be used, but this requirement is unusual. The format used is normally determined by the supplier's preference and the system that is being used. Both TIFF and PDF can be provided in

either single or multi-page form. Single page gives an electronic file per page of a document, multi-page a single electronic file per document. Be aware that different eDisclosure systems treat the allocation of unique document identification numbers to image files in different ways. This is particularly significant when handling single page tiffs.

4.10 For example a 10 page word document when produced as a native file would have a single unique document identification number, so LLL001.docx. If produced as a multi-page TIFF or PDF, it would again have a single number so LLL001.TIF or LLL001.PDF. If produced as a set of single page images, it can be numbered as LLL001 to LLL010 PDF or TIFF. This change in numbering approach can cause confusion when looking at the disclosure list, and might need to be explained to the legal team.

4.11 When it comes to exchanging image files, it is quite normal for differing sides to have different formats depending upon the requirements of their eDisclosure software. There is no need for each side to have the same format for image exports. Some software, ingest images as single page TIFFS, other systems prefer multi-page PDF. Packages that take imports as single page TIFFS, are capable of exporting ESI as multi-page PDF's, so this would be reflected in the appropriate section of each party's protocol.

4.12 **Relevant parent / irrelevant children | Format for Parent.** The issue here is as follows. When an email family is brought into an eDisclosure system, the attachments within it are copied out into ESI instances in their own right. So an email with two attachments; one a Word document, the second an Excel spreadsheet, would result in 3 items in the eDisclosure system, the parent and the two children. BUT the email itself still contains the two attachments. Normally producing an email parent in native mode, also means producing the attachments within that email family. When the children are irrelevant or privileged (either part or full), that means they will be disclosed to the opposing party irrespective of the decisions made on the renditions of the children within your eDisclosure system. The remainder of this section explores the implications of this. First, there are technology differences between eDisclosure systems, which obviate this point, so it is extremely important you discuss this specific capability with your supplier at the start of any matter.

4.13 Most eDisclosure systems store native emails with their attachments still embedded within them. Some systems are starting to appear on the marketplace which can store an email in a native format which does not embed attachments. If your system is one of these then most of the following debate does not apply.

4.14 Almost all eDisclosure review exercises involve making decisions on how to deal with email families containing a relevant parent and irrelevant attachment(s) or vice versa.

4.15 There is no single approach which will be appropriate for all cases. Ideally, before commencing a review you should:

- Decide how to you wish to approach the production of such items and try to agree this with your opponent; and

- Produce clear guidelines for reviewers as to how they should approach the task of reviewing and coding such items, so that it is consistent with how you intend to produce them

4.16 However, in practice the question is not properly addressed at the outset of a matter and/or views on what is appropriate for any given case evolve as a review progresses, meaning it has to be dealt with as a QC and production issue. Therefore it may be more practical and realistic to aim to inform your opponent at the time of, or shortly before, giving production of the approach which you have taken.

4.17 Set out below are the scenarios which arise in relation to mixed relevant/irrelevant family groups (which for the sake of completeness include the scenario where the entire family is relevant). The question of how privileged documents should be approached in this context is also touched upon below but a full consideration of the interplay of privilege is beyond the scope of these guidance notes. There is a summary table shown at 4.27 which you might want to print out now and refer to whilst reading the following paragraphs.

4.18 **Entire family is relevant.** If an entire family is relevant then all family members should be produced as natives, subject to privilege/confidentiality:

- If a parent is privileged it should be either placeholdered or not produced at all, depending upon your approach to the scenario of an irrelevant parent with relevant children.

- If one or more children is privileged, then the privileged child(ren) should either be placeholdered or not produced, and the parent should be produced as an image (unless the parent is also privileged in which case it should be dealt with as above.)

- If one or more of the parent or child(ren) is part privileged then the part privileged item(s) should be produced as a redacted image, and the parent should always be produced as an image regardless of redactions (unless the parent is privileged in which case it should be dealt with as above).

- If an entire family is privileged, it would not normally be disclosed (i.e. listed) at all and clearly would not be produced.

4.19 **Parent relevant, one or more children not relevant.** There seems to be no real consensus on what the most appropriate approach for dealing with a relevant parent but one or more irrelevant children. Should the entire family be produced (subject to privilege), or should a family member be produced only if it is itself relevant, and any non-relevant items be withheld?

4.20 On the one hand, if family members are withheld from production on the basis of non-relevance, there may be concerns that an opponent will use this as the basis for challenging the scope of disclosure and to allege that documents which they consider to be likely relevant have improperly been withheld or have been mistakenly omitted from the production. A producing party may conclude that it is therefore 'safer' just to hand over irrelevant family members in order to avoid this type of challenge and they have nothing to lose by doing so.

4.21 On the other hand, the view may reasonably be taken that an opponent should be given nothing to which they are not strictly entitled (and where the litigation is between, e.g., competitor companies, there may be a risk in handing over seemingly anodyne commercial information which actually turns out to have some sort of value to an opponent. Document reviewers will often not be well placed to spot this).

4.22 In addition, the forthcoming Disclosure Pilot further underlines that the parties must avoid 'dumping' voluminous quantities of irrelevant documents on their opponent. This again suggests that producing entire families containing a mixture of relevant and irrelevant items may not be appropriate.

4.23 In this scenario the parties may propose and attempt to agree a number of options, including:

a) Produce only the relevant parent (image) and any relevant children (native). Non-relevant children, plus any privileged documents (parent or child) are then placeholdered. Placeholdering may be easier than what is suggested at option b) below and also demonstrates to an opponent that documents have not been accidentally omitted from a production. However, it does involve 'dumping' a sometimes large number of placeholder documents on an opponent and may lead to an opponent querying whether items have been properly withheld on the grounds of irrelevance.

OR

b) Produce the entire family group with image parent comprising both relevant and irrelevant items, subject to any privileged items. If a child is privileged it should be placeholdered and the parent should be produced as an image. If the parent is privileged it should be placeholdered, and only relevant children should be produced (in native form).

If option b) is selected, then the parties should consider including a load file field of REL/NOT REL so as to indicate to their opponent whether a document is being produced because it is considered to be within the scope of disclosure (REL), or whether it is being produced simply for the sake of providing complete family groupings (NOT REL). It is suggested that this approach is more compliant with the Disclosure Pilot than simply 'dumping' complete families on an opponent.

If option b) is selected, it is also worth noting that **document reviewers must be instructed to review non-relevant documents for privilege.** Otherwise there is a clear risk of irrelevant but privileged family members being swept into a production.

In this regard document reviewers could be asked to review <u>all</u> documents for privilege, regardless of whether or not they consider them to be relevant. This is a simpler instruction to follow (because the reviewer only needs to consider each document on its face) but obviously will involve incurring costs in relation to the privilege review of documents which will never be in the prospective disclosure pool (i.e. where an entire family is not relevant). Alternatively, reviewers could be instructed to code for privilege only where one or more family members have been coded relevant. This involves the reviewer performing greater 'mental gymnastics' in terms of keeping a clear focus on family groupings (as soon as a relevant item is identified a different approach must be

taken to other family members) and therefore there is greater scope for error. However, it may be a more appropriate approach to take with a data set that contains a relatively high proportion of irrelevant documents (and therefore it would seem disproportionate to ask reviewers to spend time privilege reviewing large numbers of non-relevant documents).

OR

c) Produce only the relevant family members, without any placeholdering for non-relevant members (albeit again subject to privilege). For this option, parents would be produced as images, and relevant children would be produced as natives. Relevant privileged documents are not disclosed unless a parent email is relevant and privileged; if so, the parent would be placeholdered. Non relevant children would neither be produced nor placeholdered.

This option is again arguably more consistent with the Disclosure Pilot and the obligation not to 'dump' irrelevant documents (including placeholders, presumably) on an opponent.

4.24 **Parent not relevant, one or more children relevant.** Again, there seems to be no real consensus as to how the above should be dealt with. On one view, if a child is relevant then its parent should always be disclosed and produced (subject to privilege) on the basis that the parent gives context to the relevant child even if the parent is not itself relevant. (The same is arguably not true in relation to relevant parents and irrelevant children, meaning you could quite legitimately adopt a different approach to the two scenarios).

4.25 However, the same counter-arguments may be run as at Scenario 2 – an opponent should not be given any document to which they are not, on the face of that document, entitled. Further, to produce irrelevant parents again arguably runs counter to the parties' obligations not to 'dump' irrelevant material on the other.

4.26 In summary, the options may be as follows:

a) Produce the relevant child (native) and the parent for context (image), and placeholder any irrelevant siblings. If there are no irrelevant siblings, then (subject to privilege) the entire family can be produced as natives.

b) Produce the entire family group including irrelevant siblings (native, subject to any privileged or part privileged items). Again, if selecting this option consider including a load file field of REL/NOT REL.

c) Produce only relevant/non privileged and relevant/part privileged (with redactions) items. This is often referred to as disclosure on a broken family basis. Standalone documents and children of a non-disclosed parent would (subject to redactions) be produced as natives. If providing production on this basis then no family structure would be disclosed (and in particular there would be no 'beg attach' number in the list). Child attachments may also end up being listed in the 'wrong' place chronologically, if the date field and date ordering had been done by reference to a (non-disclosed) parent.

4.27 Email Families : A summary table.

Option	Relevant parent / relevant and non-relevant children	Parent not relevant / relevant and non-relevant children	Comment
1 Disclose complete family, either producing or placeholdering irrelevant items	Produce relevant parent (image) and any relevant children (native) and either: a) Placeholder non-relevant family members; OR b) Produce non-relevant family members (native). Consider including a load file field to indicate whether individual family members are relevant/non-relevant.	c) If all children relevant, produce entire family in native. OR If children are a mix of relevant and non-relevant, produce relevant children (native) and parent email (image); OR d) Produce only the relevant children (native) and placeholder the parent	In c), parent treated as relevant on the basis that it gives context to relevant children even though parent is not itself actually relevant
2 – Disclose complete family but produce only relevant family members (no placeholdering other than for privilege)	Produce relevant parent (image) and any relevant children (native). Non-relevant children neither produced nor placeholdered	If all children relevant, produce entire family in native (subject to priv) If children are a mix of relevant and non-relevant, disclose relevant children (native) and parent email (image).	
3 – Broken family basis, disclose and produce only relevant documents.	Disclose and produce parent(image) Relevant children disclosed and produced (native) Non-relevant children not disclosed or produced or placeholdered	No disclosure of not relevant parent. Relevant children produced as native. Standalone docs produced as native	Under this option there would be no begin attachment number, i.e. no family structure. Any orphan children could end up sitting anywhere in the production, i.e. not in date order, so query the date field to be produced

4.28 **OCR documents.** When a document is produced in image format, it is often not text searchable. In most cases text will be provided as part of the disclosure. However, if it is not, you are able to make the image text searchable by carrying out Optical Character Recognition (OCR). If you are providing redacted images as part of your disclosure you should OCR the document after the redactions have been made and provide the OCR'd text in your disclosure. This will ensure redacted text is not provided to the other side. If your matter falls under the remit of the Disclosure Pilot for the Business and Property Courts in England and Wales, that came into force on 1st January 2019, then use the second reference in this paragraph, if outside of this then the first.

4.29 **Paper Documents.** If hard copy documents form part of the disclosure, they should have been scanned and turned into text searchable images, either TIFF or PDF. In the case of TIFF's this means providing the OCR text as a separate file. If the document needs to be captured in colour, you might use the JPEG image format instead of TIFF.

4.30 A decision should be made in conjunction with the legal team on the provision or otherwise of coded data for scanned images. The normal data provided with a scanned image is known as objective coding and consists of:

- Document title / Subject
- Document type
- From / Author
- To / Distribution
- CC
- BCC
- Date (including an indication if it is an estimated date)

4.31 **Translated Documents.** A number of suppliers have raised the topic of translated documents. At present this does not seem a sufficiently mainstream requirement to warrant a paragraph in the template protocol document, should one be required, it could be added as Paragraph 4.9 in that document. The following points are suggested as a start point for this area, as ever, discuss this with your supplier.

4.32 The translation of documents may have been created manually (human translation) or automatically (machine translation); there may be an obligation to provide these translated documents to accompany the original version, especially if this is to be used later in trial. To simplify the locating of documents within a disclosure there are some recommendations for all parties:

- Pre-fix the title of the translated document with TRANSLATION AND the Disclosure Begin Bates;

- Use the same Disclosure Begin Bates with the addition of _T (underscore T) to identify this is the translation of the original item;
- The MD5 does not require populating;
- The Author should be removed;
- The date fields can be included however, ensure these match the original documents.

5 LOAD FILE FORMAT

5.1 The text used here is chosen so that it works irrespective of the decision made in the previous section on using either PDF's , TIFF or colour JPEG for the image format. These are industry standard formats, so there should not be any issue with your supplier/software complying with this format. The use of UTF-8 for text files is only specifically required when there is a need to display characters from certain languages such as Chinese or Arabic. However there is no harm in having this as a default for the text file, speak to your supplier for more information on this.

6 LOAD FILE CONTENTS

6.1 Again, this is an industry standard approach but check with your supplier/software first. In particular, find out if they will need/use the image path information, or will this data be subsumed within the .OPT load file.

6.2 Note that the provision of the MD5 Hash values will enable suppliers to make a "best efforts" attempt to carry out de-duplication. This is not an exact process, and you might find using the near-duplication functionality in some software packages a better solution.

7 EXCHANGING DATA

7.1 If you have a secure FTP site, then this is the quickest and most secure way to transfer data. Normally you can offer this facility to opposing parties as well. All external suppliers have this capability, and you should avail yourself of it.

7.2 If the data volumes are such that an FTP exchange is not viable then you will need to provide an encrypted medium, such as an external hard drive. There are a number of encryption packages available, your in-house IT department and/or external supplier will be able to help you with this.

7.3 Encryption means that the actual physical device has been formatted so that you cannot access it without the correct password. Only once the medium is encrypted, do you copy data to it. If you have compressed your delivery into a Zip file, you might also password protect that as well. Putting a password protected zip file on unencrypted media is NOT as secure and should normally not be used as an approach.

7.4 Some suppliers have reported clients requesting that Zip files be password protected, and then transferred on unencrypted devices. For example, the package 7zip encrypts the data with an AES-256 encryption which for some organisations is an acceptable approach. However, for the purposes of this protocol, this is not

recommended as a best practice, and should be discouraged if encountered, though ultimately the responsibility for data security will rest with the organisation dictating the security policy.

7.5 Some suppliers use encryption products (an example of which is software called VeraCrypt) to transfer their data. This is fine and should be dealt with by your in-house litigation support staff or external supplier.

7.6 As a matter of course, the password for the encrypted medium should always be sent in a separate email to the delivery. Watch out for users who want to send a covering letter with the medium, that contains the password. You should never send passwords and the hard drive together.

7.7 A common practice is to exchange the data in advance of the due date, and then exchange passwords when all sides have the disclosed information on site.