GUIDELINES
TO THE TeCSA / SCL / TECBAR
eDISCLOSURE PROTOCOL

Version 0.2
9 January 2015

- ***This document forms part of a series of documents which are intended to be read together. The full set of documents consists of:***

  - ***Introduction to eDisclosure Protocol***
  - ***eDisclosure Protocol***
  - ***Guidelines to eDisclosure Protocol***
  - ***Appendices 1-6: Examples of wording which could be agreed in the Protocol***
  - ***Annex A: Legal/EDRM Timeline***
  - ***Annex B: Suggested pathway to the first CMC***
  - ***Annex C: eDisclosure Checklist***
  - ***Guide to eDisclosure***
  - ***Glossary of technical terms***

  The full set of documents is available at is available at: http://www.tecsa.org.uk/e-disclosure

- ***These Guidelines should be used to assist the parties in agreeing the various elements of the eDisclosure Protocol.***

- ***In preparing the Protocol and referring to these Guidelines, the parties should also refer to the suggested timeline (attached as Annex A). In particular, it is recommended that the parties and their legal representatives start preparing for disclosure as soon as possible, and preferably during the pre-action stage.***

- ***As the draft eDisclosure Protocol may be extensively amended from case to case, it is referred to in these Guidelines as "the template version of the Protocol".***

- ***Appendices 1 to 6 set out detailed examples of different approaches which might be adopted. They are provided only as examples of how the Protocol can be used and they will in many cases need to be amended to meet the particular circumstances of each case. (The examples are shown in italic text.)***

- ***The Parties may wish to work on a joint shared copy of the protocol although there may be circumstances (e.g. where there are numerous parties) when each party may prefer to record their approach to disclosure and areas of agreement/disagreement in separate copies which are exchanged with the other party/parties.***

- ***These Guidelines take account of substantial feedback from litigation support managers, law firms and eDisclosure suppliers. Their contributions have been gratefully received.***

GENERAL

(1) CPR 31.5(4) requires that not less than seven days before the first Case Management Conference ("CMC"), and on any other occasion as the Court may direct, the parties must, at a meeting or by telephone, discuss and seek to agree a proposal in relation to disclosure that meets the overriding objective.

(2) The Disclosure Report which parties must prepare and file not less than fourteen days prior to the first CMC, in order to comply with CPR 31.5(3)(d), must include an estimate of the broad range of costs that could be involved in giving disclosure in the case, including the costs of searching for and disclosing any electronically stored documents. If the parties have produced an Electronic Document Questionnaire ("EDQ"), the EDQ should accompany the Disclosure Report. It is recommended that the parties use and exchange the EDQ as an opportunity to understand the nature and extent of the other party's electronic documentation. This can "kick-start" the dialogue process.

(3) Each party must also submit a costs budget for the claim which includes a section for disclosure. The budget must be served and filed no later than seven days prior to the first CMC.

(4) The purpose of these Guidelines is to assist the parties in reaching agreement in relation to carrying out eDisclosure, with a view to minimising cost, minimising delay and meeting the overriding objective in CPR Part 1.1, and in preparing a budget for eDisclosure**.** The template version can be used as the agenda for a dialogue between the parties in respect of disclosure. As each item is agreed, it can be recorded in the template version of the Protocol.

(5) It is envisaged that it may not be possible to agree all matters set out in the protocol prior to the first CMC and that, in any event, the Court may make an order which varies any of the agreements reached by the parties. Moreover, as the disclosure process unfolds, one or both parties may need to re-visit some areas of agreement as recorded in the protocol. Consequently, the Protocol as agreed between parties should be considered as an organic document that may develop or change over time.

(6) The template version of the Protocol states that the matters set out in the Protocol do not represent a contractually binding and enforceable agreement unless they are expressly stated to amount to a contractually binding and enforceable agreement. This wording is intended to reflect the fact that parties may legitimately be reluctant to commit to a legally binding agreement which leaves no room for flexibility in the event that circumstances change. This is particularly likely to be the position early on in the proceedings, before the first Case Management Conference ("CMC"). However, later variations of the matters set out in the Protocol may have consequences in relation to the parties' liability for costs.

(7) Parties may wish to embody the agreements reached in this Protocol in a direction of the Court in the following terms:

"*Disclosure shall take place in accordance with the agreed eDisclosure Protocol dated [ ], with permission to apply for further directions varying the matters so agreed*."

(8) Any suggested amendment should be agreed between the parties. In some instances, where the amendment carries budgetary consequences or where the parties cannot agree on the proposed amendment or where the agreements reached are embodied in an Order of the Court, the parties may need to refer the issue to the Court.

(9) Attached (as Annex B) is a flowchart showing the suggested pathway for parties to follow in respect of disclosure prior to the first CMC.

(10) Attached (as Annex C) is a suggested checklist of points which you might wish to consider at each stage of the proceedings.

(11) A Glossary of technical terms is available on the TeCSA website, at *http://www.tecsa.org.uk/e-disclosure*, within the Guide to eDisclosure.

1. IDENTIFICATION OF SOURCES OF DOCUMENTATION

1.1 Timing is important – the process of identifying sources of documents needs to begin as soon as possible after litigation is in contemplation if the parties are to achieve agreement on disclosure prior to the first CMC and if each party is going to be able to provide a realistic cost budget for disclosure at the time of the first CMC.

1.2 Before dialogue can commence with the other party or parties, each party and its legal representatives should have undertaken the following steps:

(a) considered and reviewed to the extent practicable all likely sources of possible disclosable documents and all likely custodians and locations;

(b) ascertained what document management policy (if any) is in place within the party's organisation and ensured that all possible disclosable documents are preserved (by, for example, ensuring that any standard or routine document destruction policy is suspended for the duration of the dispute; and that any devices that contain documents in any format including back-up tapes are not destroyed or over-written);

(c) considered whether a third party service provider is required to assist in the identification and collection of documents, and whether an electronic database is required in which to store, process, filter and review all documents collected.

1.3 As a result of taking the steps above, each party should be able to list in Appendix 1 to the eDisclosure Protocol all the information it has ascertained in respect of its documents in order to assist all parties in agreeing a protocol for disclosure. This information will include:

(a) The various locations of documents and who the key custodians of documents are (for example, are documents located in shared network drives within the organisation and/or stored on the hard drives of desktops/personal computers/portable devices and/or stored remotely?).

(b) Identification of any documents that may be stored outside the jurisdiction of England and Wales. (Do you have the right to access the documents? Are there any particular data protection issues?)

(c) Identification of any documents which are not reasonably accessible or which did exist but may no longer exist (for example, what has happened to devices used by relevant employees who may have left the client organisation).

1.4 It is important to consider all possible sources of documentation or data, including "non-traditional" sources such as social media (Twitter, Facebook, LinkedIn etc.), instant messaging, audio data and photographs including, where relevant, associated metadata. A decision can then be made about whether each source is likely to yield disclosable documents and, if so, whether on balance it is proportionate in terms of time and costs to collect such documents. Decisions made on each source should be documented in case a decision needs to be revisited and/or justified at a later date. It should be noted that as new forms of communication emerge, the types of "non-traditional" sources of documents will change and expand.

2. PRESERVATION OF DOCUMENTATION

2.1 See paragraph 1.2(b) above.

3. COLLECTION OF DOCUMENTS

3.1 Given that the Protocol should be finalised no later than 7 days prior to the first CMC, the process of collection of potentially disclosable documents may need to commence before the first CMC.

3.2 For guidance on different IT service providers, please see the "Guide to eDisclosure".

Document Formats

3.3 There is often confusion over the different types of document and how they are processed/produced to the other side. This section provides guidance on this point.

3.4 In general terms, documents fall into the following categories: :

(a) Electronic documents or "soft copy" documents known collectively as "Electronically Stored Information" or ESI (e.g. emails, Microsoft Word documents, Microsoft Excel spread sheets, JPEG photographs etc.).

(b) Paper or "hard copy" documents which can be sub-divided into:

(i) Paper based records that can be changed into electronic form (i.e. into PDF or TIFF format) by scanning. Documents containing print can be made searchable by a process known as Optical Character Recognition (OCR). Documents containing tables, graphics or manuscript notes cannot go through the OCR process which means that it will not be possible to undertake electronic word searches on parts of such documents. Once paper documents become electronic after scanning, they are considered to be *electronic* documents to be managed in the same way as all of the other electronic documents in the disclosure project (see *Re Atrium Training Services Ltd, Smailes v McNally* [2013] EWHC 2882 (Ch) at [56-58]).

(ii) Paper documents that cannot be transformed into electronic form, and so remain in paper format throughout the entire process. .

Native Documents

3.5 The optimum way to collect electronic documents is in its "native format" (i.e. a copy of the original document is made in the format created by the authoring application such as Microsoft Word, Microsoft Excel etc.).

3.6 In extracting documents from their particular sources, care should be taken to ensure that the metadata associated with the documents is not altered. For example, an MS Word document will contain metadata which indicates the date on which the document was created. This date may be important, but the process of extraction could, if not carefully carried out, change that date to the date of extraction, thus destroying potential evidence). The process of extracting documents from their sources will require the assistance of a person who has appropriate I.T. forensic expertise, such as a third party service provider. Some in-house I.T. personnel may have the necessary level of expertise, but many will not – they should not be entrusted with the task of extracting documents without it being checked that they possess the appropriate level of expertise. If in any doubt, obtain advice from an I.T. specialist.

3.7 For native documents, the normal approach is not to input coding information into the document review database manually, but to populate the database with the data which resides within the document's metadata fields. For example emails automatically have a date/time stamp, a subject

line and a list of addressees. Office documents such as Word, Excel and PowerPoint have some information within the metadata, with the key information being the date field.

3.8 In some circumstances parties might agree to provide additional coding for Word documents, such as the actual title of the document (as opposed to what the electronic file is called), the author of the document (as opposed to the owner shown in the metadata field) and the actual date shown on the document (as opposed to the information held in the "Date last Modified" metadata field). This can be an expensive undertaking, with little practical benefit as the search and other tools within eDisclosure software may render this kind of data superfluous.

Non-Native Documents

3.9 In some instances, it may not be possible to collect documents in their native format because, for example, they only exist in hard copy or in scanned PDF format. Generally speaking it is preferable to scan hard copy documents into electronic form, as this means that all the documents in the case can be stored electronically in a single system. In other instances, the documents will have been created using unusual, specialist or bespoke software which may not be readily accessible – in this case, it may be possible (though not always) to convert the document to PDF format. PDF documents should be made searchable by applying the OCR process (as mentioned above at paragraph 3.4(b)). . Disclosing non-searchable PDFs should be the exception, not the rule.

3.10 If the parties decide to convert hard copy documents into PDF format, they need to consider whether the documents should be scanned in colour. Normally, colour versions will only be created if it will be of evidential value to see the colour. The parties will need to determine this in advance of sending documents to a third party to be scanned.

3.11 If hard copy documents are to be scanned and uploaded to a document review database, they will need to be "coded" with associated information to identify each document (bearing in mind Practice Direction 31B paragraph 31(1)). This will include for example the information that a native document would normally carry with it in its metadata, i.e. the date of the document, the author, the document type and file-name or email subject line.

3.12 As a minimum the following coded fields will be required for <u>all</u> non-native documents:

(1) Date of Document (The date should be coded and exchanged in a numeric format, that is DD/MM/YYYY, e.g. 19/02/1957. Once stored in this form it can be displayed in both the litigation support system and any exchange lists according to requirement, e.g. 19 Feb 57, or 19 February 1957)

(2) A field which states whether or not the date has been estimated. If dates are estimated, the parties should explain the convention they will use to show missing day / month / year.

(3) Author of Document

(4) Addressee of Document (if any), and

It may also be considered helpful to include:

(5) Document Title

(6) File Type, and

(7) Names of persons to whom copies were sent.

The required coded information needs to be considered carefully in respect of each document type (e.g. drawings, letters etc.).

3.13 Care should be taken when PDF-ing hard copy documents to retain any host-attachment relationship if possible, so that attachments do not become lost or unidentifiable.

3.14 Occasionally, some parties use TIFF as a format in which to disclose documents. Unless there are good reasons for doing so (for example, because a party's review database can only operate in TIFF), then it is recommended that this format is avoided because TIFFs are not readily searchable.

3.15 Consideration should be given to data stored in electronic or other formats that cannot be readily collected or separated from other extraneous material. This may include material such as electronic data stored in large and/or complex databases and "Cloud" based storage. Expert advice may be necessary in order to determine whether this material can be extracted and disclosed in a usable format, or whether it can otherwise be secured against deletion or modification pending an agreement on how to make this material available to all parties if required.

Choice of Disclosure Approach (paragraph 3.4 of the eDisclosure Protocol)

3.16 CPR rule 31.5(7) refers to various options that the parties can choose in respect of disclosure. The choice depends on many factors, including the value of the overall claim, the likely number of disclosable documents involved, the ease of retrieval, the nature and location of the documents (are there likely to be many privileged documents dotted around the sources?), the likely cost of disclosure etc.

3.17 The choices the parties can make are anything from dispensing with disclosure, to arbitration-style disclosure, to the "keys to the warehouse" approach (i.e. allowing the other party to inspect the whole pool of relevant and irrelevant documents). One option is that of "standard disclosure" which has been the approach applied to disclosure since the introduction of the Civil Procedure Rules in 1999.

4. PROCESSING AND REDUCING THE POOL OF DOCUMENTS

4.1 Generally, the more data that is processed, the higher the cost. It is recommended that consideration is given to whether processing of certain categories or sub-sets of documents can be deferred pending further investigation into the facts of the case.

4.2 In some instances, some of the filtering process can be undertaken before processing, thereby reducing the cost, such as filtering by date ranges and removal of particular file types. Any such filtering should be agreed with the opposing party at the earliest possible stage to avoid the risk of having to repeat the exercise later.

Date Ranges

4.3 The parties should set out in Appendix 1 to the eDisclosure Protocol the date range(s) to be applied to the party's disclosable documents. The date ranges may differ depending on the type of document or the custodian of that document (for example, a particular custodian may not have joined a project until a date after commencement on site and therefore his or her potentially disclosable documents will start at a later date than other custodians who started on an earlier date).

4.4 In some instances, disclosure may need to continue up to the present date. In this case, consideration should be given to how the parties will need to "refresh" the documents they have extracted at a certain date with subsequent documents brought into existence after the date of extraction.

Document/File Type

4.5 It may be possible to remove certain document or file types from disclosure at the outset because it is immediately evident that they will not reveal any disclosable information. For example, this may be the removal of "system files".

Key Word Filters

4.6 Once documents have been extracted and date ranges applied, it is common to produce lists of words which can be used to search the pool of potentially disclosable documents to (i) exclude irrelevant documents and/or (ii) identify disclosable documents. Keywords could also be used to locate and remove privileged material (particularly documents subject to legal advice privilege).

4.7 Filtering by "keywords" should be regarded as an iterative process, because search results may indicate that particular keywords result in too many "false-positives" or in disclosable documents being excluded, or may suggest further words that could be usefully added to the list. Therefore, it is expected that any keyword lists will go through a process of refinement and change until they can be finalised.

4.8 Outlined below are a number of basic points to consider when applying keyword filtering to a pool of documents:

- Personal names are often misspelt. If possible, obtain a list of the permutations of personal names and consider searching for part of a name such as the beginning, and then widen or narrow the search.
- In cross-border cases US spellings should be considered.
- False-positives (documents that meet the search criteria but are of no interest) needlessly increase the number of documents that need to be reviewed. Therefore, consider the use of the "NOT" operator to *exclude* common documents in the pool.
- Consider obtaining a list of the number of times a word is mentioned in the database - this is commonly called a word frequency analysis. This can help frame queries more efficiently.
- Be aware that there are characters that one may not be able to search for such as hyphens, underscores and part of email addresses such as "." and "@". In addition, individual numbers frequently return large numbers of false positives. There are often ways to get around these issues so talk to the third party service provider.
- Using numbers as key words can cause a disproportionate amount of false positives, particularly within a body of other numerical information such as spreadsheets. These should be carefully vetted using word frequency analysis or similar techniques, and perhaps combined with other keyword operators or "proximity" functions, for example "123 within 3 words of 'Project'".

4.9 The aim of key word filtering should be to reduce the pool of documents without eliminating disclosable material. It is essential to avoid the possibility of any misunderstanding in relation to

the use to be made of keywords or other filtering processes. This point is covered in section 5 below. It is usually agreed that after filtering by keywords, further review and analysis of the documents will be carried out.

Duplicates

4.10 De-duplication is the process whereby emails and other electronic files are removed from a population of documents if they are deemed to be a duplicate of another document within the same population.

4.11 Duplicates are not always easy to deal with. Most document review databases can undertake "de-duplication" processes to remove exact copies. For emails, a database will consider the "Hash" value, which is calculated on the following fields: "to", "from", "CC", "BCC", "Subject", body of email and any attachments. Other software may consider the "SHA1" value. If all fields are identical, then the database will remove any duplicates, leaving only one copy. Individual electronic files that are not email have the Hash applied to the binary stream of the file and are removed from the population in such a way as to leave only files with a unique MD5 Hash present in the population. When de-duplicating populations of documents the processing systems should record which other custodians held copies of material that has been de-duplicated out of the disclosure population, in order to provide an accurate representation of which custodians held which documents.

4.12 A difficulty lies with documents which are considered to be "near duplicates", such as a document which exists in its native MS Word format as well as in a scanned PDF or where the same email has been sent to several recipients, all of which have been captured in the extraction process because each recipient has been identified as a key custodian of data (such de-duplication could be done on the basis of comparing the date and time of the email as sent). There are ways to deal with "near-duplicates" which parties should discuss with each other and with the third party service provider, if they have one.

4.13 The following wording might be considered appropriate in relation to de-duplication, though more extensive levels of de-duplication are possible:

"*Duplication will be considered at a family group level – i.e. all the documents within a family group (that is, the host or parent document together with the attachments) will be treated as duplicates if the entire family group is duplicated elsewhere within the collection. An attachment will not be treated as a duplicate if it is merely duplicated elsewhere as an individual, stand-alone document*."

5. REVIEW AND ANALYSIS

5.1 In many cases keyword filtering (if used carefully) is a practical way of reducing the pool of disclosable documents. However, depending on the disclosure option agreed by the parties or ordered by the Court, it will usually be an unreliable way of determining which documents fall within the scope of disclosure and which do not. Keyword searches are rarely sufficient, for example, to ensure that all significant documents have been located and that all irrelevant or privileged documents have been removed.

5.2 As stated above, it is essential to avoid the possibility of any misunderstanding in relation to the use to be made of keywords or other filtering processes.

(a) Is it intended that all documents which contain particular keywords should be disclosed without further review?

(b) Or is it intended (for example) that there should be a further review carried out in order to remove all the documents which do not fall within (for example) standard disclosure?

(c) If the former (i.e. (a) above), is it intended that a party may (if it so wishes) remove documents which contain an agreed keyword but which are nonetheless clearly irrelevant?

5.3 If the parties wish simply to agree that all documents which respond to keywords or which remain after keyword filtering will be disclosed without review, both parties should be clear about the inherent risks of this approach – it may mean a higher volume of disclosed documents which contain irrelevant material not sifted out by keyword filtering, some of which irrelevant material may be commercially sensitive or confidential (such as documents containing personal data); some privileged documents may be missed by keyword filtering; and in other instances, the keywords may fail to capture all disclosable documents. In those circumstances, the parties may not be able to give standard disclosure, or such other level of disclosure as they have agreed in the eDisclosure Protocol (paragraph 3.4 in the template version of the Protocol).

5.4 If the parties do not wish to agree that all documents which respond to keywords or which remain after keyword filtering will be disclosed without review, then each party needs to consider what further work needs to be done on the documents to comply with the chosen disclosure option and avoid the risk of disclosing too much non-disclosable material, and thereby shifting the burden in terms of time and cost onto the receiving party to sift through lots of irrelevant material or the risk of disclosing privileged material. This may take the form of lawyer linear review of documents or categories of documents and/or the use of computer-assisted review such as predictive coding. For example, which (if any) custodians are to be reviewed in full?

6. EXCHANGE AND INSPECTION OF DOCUMENTS

6.1 The parties should agree whether disclosure should be given in a single batch, or whether it will be necessary or desirable to divide disclosure into several batches or stages. There are different ways in which this might be done. For example, drawings may be stored in a separate database used during the project by all parties – it may therefore be possible to agree that all such drawings do not need to be formally disclosed because all parties will already have access to them, or it may be possible to provide access to the other party to the database or to download a copy of all drawings and disclose those particular documents quickly and in a straightforward manner. Where there are a very large number of documents to be disclosed, it may be decided to supply the documents in several stages divided up by date ranges.

6.2 Another approach would be to agree that the documents held by the most significant custodians should be disclosed first, and that the decision whether the documents of other custodians should be disclosed (and, if so, to what extent) should be deferred until after the first tranche of documents has been reviewed and considered. This approach can work well in project-type cases where there are numerous custodians with potentially relevant material but where it is likely that a large proportion of the important documents will be captured by review of a smaller subset.

6.3 If disclosure is to take place in stages, the parties need to identify what each stage of disclosure will comprise (in terms of document type or category, or particular custodians or origin). Further, if disclosure requires an update, the parties should agree when the update or updates should take

place and what updates are required (i.e. will only certain custodians or document types suffice or will each refresh have to be as wide as the original extraction?).

6.4 Unless there is good reason to do otherwise, consistent methodology should be used across each stage, such as sorting, filtering and de-duplication methods. Attention should be drawn to any inconsistencies.

6.5 The parties should agree the date or dates for disclosure and record this in the eDisclosure Protocol. These dates will normally also appear in the directions given at the CMC.

6.6 The same consideration needs to be given to the date(s) for inspection. In this respect it should be noted that where a large volume of documents has been disclosed, it may take some time (several weeks) to produce the documents for inspection. This should be taken into account when agreeing a date for giving inspection of documents.

6.7 Alternatively, the parties could consider providing access to their own disclosure document review platform for use by the other parties to review the disclosed documents. For example, Party A could agree to give Party B access to Party's A database to view Party A's disclosed documents – these disclosed documents would have to be held in a separate part of the database away from Party A's wider pool of documents. In effect, Party A's platform becomes a "shared platform" for these purposes with each party only having access rights to particular areas of the platform. This is being increasingly used in very large scale disclosures where the time taken simply to produce a load file is considerable or it could work where parties have agreed on a "keys to the warehouse" approach to disclosure in accordance with the "menu options" set out in CPR 31.5.

6.8 In advance of the agreed date(s), the parties should consider the logistics of disclosure and inspection in respect of at least the following points:

(a) In addition to the use of Court Form N265, do the parties intend to list each document to be disclosed individually? Does this include all documents over which privilege is claimed (or over documents over which litigation privilege only is claimed)? If such a list is to be produced, it should follow the format set out in Appendix 6 of the eDisclosure Protocol and be compliant with the requirements of the CPR.

(b) Where documents have been collected in their native format, what metadata will be provided with them? Are there any documents which were created using unusual software which is not available to the receiving party, so that the receiving party will be unable to access them?

(c) Will any of the documents be redacted? How will redactions be identified? Can each redaction be labelled so that it is obvious what the grounds of redaction are in each instance?

(d) Are there any reasons why documents will not be listed and produced in date order? Or is there any reason why an attached document cannot also have an identifier to indicate its host document?

(e) Will copies of the documents be provided by means of portable storage or will they be exchanged by way of a network transfer or uploaded to a web-based file sharing facility? What security measures will be applied? (It is good practice to encrypt the portable media used to store documents.)

6.9 It is always important to ensure that documents are supplied in a manner which preserves the relationship between parent and child documents, for example the relationship between an email and its attachments.

Listing Documents

6.10 As can be seen from above, it is not mandatory to list each and every document that is being disclosed. Parties may wish to provide lists for any documents which remain in hard copy only and/or privileged documents but dispense with listing the usually considerably larger number of electronic documents, copies of which can be and usually are exchanged in their entirety without any need to review the list beforehand (see rule 31.10(8) which provides that, at least where standard disclosure is being used, the parties may agree in writing to disclose documents without making a list). The decision about whether or how to list documents will depend on the approach to disclosure (i.e. what "menu option" has been ordered or agreed).

6.11 In some cases it might be helpful to provide a list in advance of inspection so that the parties can review the nature and extent of the documents disclosed. This can assist in planning inspection, in identifying any anomalies or gaps in chronology for example, and can be used as an index to cross-check against the copies when they are produced for inspection.

6.12 In particular, in respect of privileged documents, it is usual to disclose the existence of such documents by category only in Court Form N265. However, the parties may wish to consider providing a list of each document over which privilege is being claimed (at least, in respect of documents over which litigation privilege is being claimed) so that each party can review and assess the documents and challenge the claim to privilege if appropriate (this is sometimes referred to as a "Privilege Log"). If the parties opt for this approach, careful consideration should be given to how much information can be provided to describe each document over which privilege is claimed to enable the receiving party to make its own assessment of whether it is likely to be privileged without revealing the privileged content. The provision of a List of Documents is not mandatory in every case (CPR 31.5(8)(b) and 31.10(8)(a)). If there is to be no List of Documents, parties will need to consider in what other document or documents their claim for privilege should be made.

The Disclosure Statement

6.13 Rule 31.10 (5) states that a disclosure list must include a disclosure statement and rule 31.10 (6) states that the disclosure statement is to be made by "the party disclosing the documents". Rule 31.10(9) permits the disclosure statement to be signed by a person who is not a party where this is permitted by a relevant practice direction – this applies to an insurer or the Motor Insurers' Bureau acting on behalf of a party where the insurer or the Motor Insurers' Bureau has a financial interest in the result of proceedings brought wholly or partially by or against that part (PD 31A 4.7). Rule 31.10(8) goes on to provide that the parties may agree in writing to dispense with a disclosure statement.

6.14 The Disclosure Statement should be signed by the client or a person within the client organisation who has oversight of the disclosure process. In the vast majority of cases, it should be the client who signs the Disclosure Statement. Where the client is a company, firm, association or other organisation, the Disclosure Statement must identify the person making the statement and explain why he/she is considered an appropriate person to make the statement (see rule 31.10(7) (a) and (b)).

6.15 However, a client's legal representative has a duty to explain the client's disclosure obligations to the client and to ensure that the client understands them (PD31A paragraph 4.4).

6.16 The disclosure statement must:

6.16.1 Set out the extent of the search made to locate documents which the person signing it is required to disclose;

6.16.2 Certify that he/she understands the duty to disclose documents; and

6.16.3 Certify that to the best of his/her knowledge he/she has carried out that duty.

6.17 PD31A paragraph 4.2 adds further details stating that the disclosure statement should:

6.17.1 Expressly state that the disclosing party believes the extent of the search to have been reasonable in all the circumstances; and

6.17.2 In setting out the extent of the search draw attention to any particular limitations on the extent of the search which were adopted for proportionality reasons and give the reasons why the limitations were adopted, e.g. the difficulty or expense that a search not subject to those limitations would have entailed or the marginal relevance of categories of documents omitted from the search.

7. CONFIDENTIALITY AND PRIVILEGE

7.1 The template version of the Protocol provides at paragraph 7.1 for the listing of documents over which privilege is claimed if the parties so wish to agree. This is similar to a "privilege log" of the type commonly used in arbitration, where parties list documents over which litigation privilege is claimed (with parties usually agreeing to dispense with the need to list documents over which legal advice privilege is claimed) (see paragraph 6.12 above). In many cases, the parties may prefer not to agree to list each individual privileged document as long as they are adequately identified in the Disclosure List.

7.2 Paragraph 7.2, of the template version of the Protocol, dealing with inadvertent disclosure of privileged documents, goes further than CPR 31.20, which only states that "where a party inadvertently allows a privileged document to be inspected, the party who has inspected the document may use it or its contents only with the permission of the court". Paragraph 7.2 states that no use may be made of a privileged document which has been inadvertently disclosed and there will be no waiver of privilege – the receiving party cannot seek the Court's permission to use such a document. Where large quantities of electronic documents are involved it may be impossible or impracticable to be completely certain that every privileged document has been withheld, and there is a greater risk of inadvertent disclosure. Parties may therefore wish to include a greater level of protection against inadvertent disclosure than would be provided by CPR 31.20. (This type of agreement is known in the USA as a "clawback" agreement,)

7.3 It should be noted that whatever may be agreed between the immediate parties in the eDisclosure Protocol, or directed by the court, the inadvertent production of a privileged document may still amount to a waiver of privilege vis-a-vis third parties or parties joined in the action after the Protocol has been agreed.

7.4 Parties may wish to agree that the paragraph(s) dealing with no waiver of privilege should form part of an Order by Consent for directions or are the subject of a legally binding agreement.

APPENDICES & ANNEXES

GUIDELINES
TO THE TeCSA / SCL / TECBAR
eDISCLOSURE PROTOCOL

Version 0.2
09 January 2015

APPENDICES 1-6
EXAMPLES OF WORDING WHICH COULD BE AGREED

(Full set of documents available at http://www.tecsa.org.uk/e-disclosure)

A1. APPENDIX 1 – LOCATION AND NATURE OF DOCUMENTS AND KEY CUSTODIANS

Custodian Lists

A1.1 List the custodians involved in the case. If individuals have joined the organisation during the period of litigation, show the dates for which documents will be provided. Also use the list to show custodians who are no longer members of the party's organisation. It may also be helpful to put together an organisational chart to show where custodians sit within an organisation or corporate group, their job title and to show any reporting lines up and down the chain of command.

Party 1

Custodian 1

Custodian 2

Custodian 3 (From DD MMM YYYY)

Custodian 4 (No longer works at [Party 1])

Document Location

A1.2 In order to provide context to the following sections, it might be useful to provide (at the appropriate level of detail) a description of where the documents are located. This would be based on information contained within the data map that lawyers are advised to draw up, but it is not suggested that the map itself is replicated here. One of the objectives of this section is to highlight

any documents that are held in locations/systems outside the general "run of the mill" infrastructures. For example:

A1.2.1 *Party 1's documents exist on an email server and shared network drives on servers all located at Party 1's headquarters in XXX. All this information is mirrored on a continual basis to business continuity servers located in YYY and was collected from those servers. Some information is held within a specialist stock control system called STOCKMAN. Part of the project was outsourced to a German based firm and details on how information has been obtained from this jurisdiction are show below*.

Documents which are located outside the jurisdiction of England and Wales

A1.3 Discuss any documents held outside England and Wales and how these might be treated taking account of issues such as data protection and data security rules in other jurisdictions.

(a) *[Role 1 of Party 1 is located in [Location outside England and Wales]. Accordingly, some potentially relevant documents are located in [Country outside England and Wales].*

Documents which are not reasonably accessible

A1.4 This area is normally focused on documents that might theoretically exist on back-up tapes, but the cost and time to establish their existence is usually considered disproportionate to their potential value to the case. A full explanation of the particulars of the individual circumstances should be given here.

A1.5 Where back-up tapes might well be used to provide information from individuals who have left the organisation, see the following section.

Documents that may no longer exist

A1.6 This area is normally concerned with individuals who have left the Party's organisation and whose electronic information was not preserved. In the absence of their data being available on back-up tapes, it might not be possible to retrieve their information. Examples of some scenarios are shown below as an indication of how you might explain these points.

(a) *Prior to Party 1 moving to Location 1 employees were based in Location 2. Their emails and document systems were migrated over a period of months in or around the latter part of YYYY to the systems at Location 1. Party 1 is neither in control nor possession of the back-ups of emails and documents created at Location 2. Any documents or emails which did not for any reason migrate will not be accessible by Party 1.*

(b) *It is Party 2's practice to retain a copy of email accounts of personnel leaving the employment of Party 2 for 2 months, unless an email account is provided to the ex-employee's line manager on request. In addition, Party 2 has a retention policy of 14 months for all backups. Accordingly, email accounts cannot be obtained for personnel whose line manager did not request a copy of their email account and who left the employment of Party 2 more than 14 months ago. Therefore, there are no email accounts for Custodian 1, Custodian 2.*

(c) *Prior to YYYY, it was Party 3's practice to delete the emails of personnel leaving the employment of Party 3 and no emails were archived. Accordingly, it is not be possible to recover email accounts for Custodian 3, Custodian 4.*

(d) *Party 4 Custodian 5 does not have an exchange account/email account, and any emails or documents were saved locally to his personal PC. Custodian 5's PC was replaced 3/4 years ago*

following the failure of his earlier PC, and no back-ups exist. Any data that may have existed before Custodian 5 current PC will no longer exist.

(e) *There is no category of Party 5's documents that will no longer exist. There may be individual documents that no longer exist but these could only be identified by specific searches.*

Documents in native format which were created using relatively unusual software

A1.7 List any documents created with relatively unusual software. For example:

(a) *Party 1 Construction drawings created and saved in CAD format.*

(b) *Party 2 Documents created in the Borland dBase Database.*

(c) *Party 3 AutoCAD, Microsoft Project and Argus Developer files.*

Documents/locations/custodians which remain and are subject to further investigation

A1.8 Under this section detail any on-going work. For example:

(a) *Party 1 is investigating the availability of email accounts for Custodian 1, Custodian 2 and Custodian 3, who left the employment of Party 1 prior to YYYY.*

Hard Copy Documents

A1.9 This can be an area of confusion. Please refer to paragraph 3.4 above.

A1.10 Hard copy documents which are not capable of being converted into electronic format could be offered for physical inspection, or photocopied and provided for inspection in accordance with CPR rule 31.15 methods of disclosure.

Either:

(a) *Hard copy documents that (1) are not capable of being converted into electronic format or (2) are unusable if converted, will be made ready for physical inspection.*

Or:

(b) *Hard copy documents that (1) are not capable of being converted into electronic format or (2) are unusable if converted, will be photocopied and exchanged in accordance with the procedure in Appendix 6.*

A1.11 Hard copy documents that are capable of being converted into electronic form by means of a process involving scanning, OCR and coding.

(a) *Hard copy documents that are capable of being converted into electronic format will be processed and exchanged in accordance with the procedure in Appendix 6.*

A2. APPENDIX 2 – KEY WORD FILTER SEARCHES AND DATE RANGES

Use of Keywords

A2.1 It is important to understand the different types of searches that each particular electronic documents database can perform and the limitations of each search type. It is also important to test key words and to understand that it may require several key word searches to refine or amend the list (see paragraphs 10.6 to 10.9 below).

A2.2 Keywords can be used in different ways. An approach to using keywords is to agree the keywords that will be used to identify potentially relevant documents. It may also be possible to agree (but not always advisable given that key word searching is an imprecise exercise that should usually only be used as one of several tools to reduce down the pool of documents for review) that any documents that are not "hit" by the keywords can be excluded from consideration.

A2.3 Key words can also be used to exclude irrelevant material. An example of this approach is:

(a) *The following key words have been agreed as a means of removing non-disclosable documents from the wider pool of documents collected.*

(b) *[….] will identify non-disclosable documents by its exclusion from the list identified at paragraph 2 below.*

A2.4 Finally, key words can also be used to identify potentially privileged documents although it would be inadvisable to rely solely on key words to identify privilege – any document "hits" arising from the search would need to be reviewed to decide if they are in fact privileged.

Keyword Lists

A2.5 There are two main approaches to using keywords:

(a) The agreed keywords are listed by the parties with some explanatory text. In this approach the implication is that the words have been agreed before they have been tested against the collected data.

(b) The parties hold back from agreeing their keywords until they have been able to test possible words against the data.

A2.6 Examples of these approaches are:

(a) *The following key words have been agreed as a means of identifying disclosable documents within the wider pool of documents collected: [These key words will be tested against the data once processed and refined as necessary]:*

Keyword 1.

Keyword 2.

Keyword 3 AND Keyword 4.

(b) *[ ] will issue its proposed list of key words for the purpose of identifying potentially disclosable documents in due course following the testing of the list against data collected from the custodians identified at Appendix 1.*

A2.7 In respect of privilege key word searches, a party might wish to record in the protocol the following:

(a) *The following key words have been agreed as a means of identifying potentially privileged documents within the wider pool of documents collected and include but are not limited to documents/correspondence:*

(b) *For [……], sent from/to/CC:*

Name 1, Name 2, Name 3.

(c) *Marked 'privileged' or 'without prejudice'.*

(d) *Containing the words:*

court

proceedings

pleadings

witness AND statement

technology AND construction

"staff appraisal"

"personal review"

*litigat**

"legal advice"

Searching for Key Words - Stemming and "Fuzzy" Search

A2.8 When applying search terms there are some techniques that are helpful to understand. Before adding the following to the protocol it would be advisable to check with your litigation support supplier as to the capabilities in this area of the software you are using.

(a) Stemming is the concept of automatically finding all the words (stems) that use the initial root word. So the root word "fish" would give you stems of "fishing", "fished", and "fisher".

(b) Some search tools allow you to look for a degree of "fuzziness" in the search. This approach can be useful when there are miss-spelling of a word, but too great a degree of fuzzy search will just produce nonsensical results.

A2.9 Examples of these approaches are:

(a) *When Party 1's key words were applied to its wider pool of documents, stemming was enabled (to return grammatical variations of any of the key words) and the fuzzy searching level was set to Level 3 (to ensure that misspelt variations of the key words were included in the pool of documents to be reviewed). However, nonsensical variants of the above keywords and words that are correctly spelt but unrelated words have been removed.*

OR

(b) *The Party 2 keyword list includes a number of wildcards within a set of Boolean searches. The use of stemming was explored, but discounted, as it did not materially improve the accuracy of the search results.*

(c) *In terms of fuzzy searching, the following approach has been adopted.*

(d) *The way in which most of the Party 2 search terms have been run is in long Boolean strings. An example of this is "Keyword 1" AND ("Keyword 2" OR "Keyword 3" OR "Keyword 4" OR "Keyword 5")". A general fuzziness level would apply to all words in the search request, leading to a disproportionate number of irrelevant documents being returned from the search.*

(e) *Party 2 selectively added fuzziness to specific words by placing the % character into the term being varied. The number of % characters inserted determines the number of possible variations and the place within the word the first % is placed determines the point from which variations are assessed.*

(f) *The terms used are set out below,*

Ke%%ord 1

Ke%%ord 2

Ke%%ord 3

Using Date Ranges in Searches

A2.10 Specific date ranges will be derived from the facts of the case. It is important to be aware that electronic documents can have a number of dates within different metadata fields and therefore it is advisable to agree the field which will be used. For example:

(a) *The parties have agreed to apply the following date range(s) to the documents collected:*

DD MMM YYYY to date.

OR

DD MMM YYYY to DD MMM YYYY

OR

The date range below is the range which, in general, applies, however there are likely to be more limited date ranges for a number of the organisations, some of which have only been involved since YYYY and certain categories of documents.

DD MMM YYYY (the date from which the agreed milestone event for the start of the case) to date.

(b) *For electronic documents the metadata date field used to filter documents will be the Date Last Modified field.*

A3. APPENDIX 3 – DE-DUPLICATION / IRRELEVANT DOCUMENTS

Approach to De-duplication

A3.1 The different approaches to de-duplication reflect different views on providing context to a document within a family group of documents (e.g. an email and its attachments). Examples are:

(a) *[….] will identify identical documents through the use of a Hash Algorithm (exact algorithm to be confirmed) and will then remove duplicates from the document collection on a global top level basis, meaning that where a document is a duplicate at attachment level, the duplicate will remain so as to keep the context of the family of documents. Appropriate steps will be taken to ensure that custodians that were in possession of duplicate documents will be subsequently identifiable.*

OR

(b) *[…..] will identify identical documents through the use of a Hash Algorithm (exact algorithm to be confirmed) and will also remove duplicates from the document collection on a global basis. Appropriate steps will be taken to ensure that custodians that were in possession of duplicate documents will be subsequently identifiable*

OR

(c) *Both parties should apply an industry standard de-duplication method to remove as many exact duplicates from their disclosure as reasonably possible. Our ESI provider will carry out de-duplication using the [……] algorithm. Only standalone duplicates or entire duplicate families will be de-duplicated, whilst duplicate documents that are attached to non-duplicate documents will not be de-duplicated.*

Approach to Irrelevant documents

A3.2 Similar criteria apply to how irrelevant documents are handled. Examples are:

(a) *Where an irrelevant document is a parent document or an attachment to a disclosable document, the irrelevant document will be disclosed for the purposes of providing context.*

OR

(b) *Where an irrelevant document is a covering or attachment document to a disclosable document, the irrelevant document will not be disclosed and a tiff placeholder will appear in its place.*

A4. APPENDIX 4 – COMPUTER ASSISTED REVIEW OR AUTOMATED METHODS OF REDUCING THE POOL OF DOCUMENTS

Use of Computer Assisted Review / Other Automated Methods

A4.1 The optimum way to approach this is to let your supplier do the hard work of providing a methodology to be agreed with the other parties. Such documents are specific to the technologies being employed and should be a key part of the service being provided by your supplier. Examples of wording are:

(a) *[….] reserves its position in respect of the use of computer assisted review and analytical tools, pending the results of its initial key word searches. To date, [….] has used the email threading feature which forms part of the Product X tools package.*

OR

(b) *[….] will employ the use of the analytics package of Product Y as detailed in the attached Supplier A protocol entitled "Product Y Assisted Review - Process".*

OR

(c) *[…] will employ the use of the computer assisted review tool available with the Product Z software.*

A5. APPENDIX 5 – DOCUMENTS TO BE FURTHER REVIEWED

Details of categories of documents which are to be reviewed to ensure that they do in fact fall within the agreed scope of disclosure

A5.1 The purpose of this section is to provide clarity on what documents will be reviewed. For example:

(a) *A single copy of all documents which respond to the keyword word search / date filters will be reviewed for relevance and to determine whether or not they fall within the requirements of standard disclosure. Documents which as a result of the review are considered to fall outside standard disclosure will not be disclosed.*

OR

(b) *The parties will review all documents which are responsive to the keywords set out in Appendix 2 to ensure that they do in fact fall within the agreed scope of disclosure. Documents which as a result of the review are considered to fall outside the agreed scope of disclosure will not be disclosed.*

(c) *In addition the parties will review the non-responsive covering or attachment documents of responsive documents where the responsive document is considered to be disclosable.*

OR

(d) *In addition all documents identified as being potentially disclosable as a result of the processes described in Appendix 4 will be reviewed for relevance. Documents which as a result of the review are considered to fall outside the agreed scope of disclosure will not be disclosed.*

Details of categories of documents which need not be reviewed at all (using any methods)

A5.2 Documents that are mentioned under this category fall into two main areas:

(a) Documents that all parties have access to.

(b) Documents deemed non-responsive by the technologies detailed in Appendix 4.

A5.3 So for example you might have:

(a) *Emails sent from/to the legal team (solicitors and counsel) from/to Party 1 or any of Party 1's experts.*

(b) *Inter-solicitor correspondence and correspondence with the Court, as these will be disclosed without being reviewed.*

(c) *All documents identified as being not disclosable as a result of the processes described in Appendix 4. This pool of documents will be sampled by Party 1 to confirm that the documents are not disclosable [any such sampling can be provided to other parties on request].*

A5.4 Where it is agreed that no review is required, the following should nevertheless be considered:

Although review of documents identified by automated methods as being disclosable is not required, some or all of these documents may nevertheless be reviewed for relevance. Documents which as a result of the review are considered to fall outside the agreed scope of disclosure will not be disclosed.

A6. APPENDIX 6 – DISCLOSURE, EXCHANGE AND INSPECTION

Use of Lists

A6.1 See paragraphs 6.8 to 6.10 in the main Guidelines in respect of the disclosure list. Parties may often wish to agree to dispense with a list altogether and simply provide copies of the electronic documents for inspection.

Inspection of Hard Copy documents

A6.2 Although these Guidelines and the accompanying eDisclosure Protocol are concerned with electronic disclosure, it is important not to overlook any hard copy documents in a disclosure exercise. If the parties have agreed on disclosure lists, then hard copy documents would usually be listed in a separate list to that which lists the electronic documents (for practical reasons: the electronic documents list will have been system-generated whereas a hard copy list will have to be typed up by someone).

A6.3 If hard copy documents are being disclosed by way of standard disclosure then the procedure for standard disclosure in rule 31.10 should be followed with inspection in accordance with rule 31.15.

Exchange of Hard Copy documents converted into electronic format

A6.4 The aim of this section of the appendix is to set out details of the methods to be used to identify the documents to be exchanged and the information that will be provided with each document. The actual mechanism of exchanging the documents is covered in the section below on the ESI Exchange Protocol.

A6.5 Normally hard copy documents are converted into PDF files, though occasionally single page tiffs are used. Whatever the format, it is again normal that the documents be searchable and have a certain level of information associated with them. The discussion at paragraph 3.10 in the main Guidelines has a direct bearing upon the information that has been agreed will be coded. It is normal practice to set out the information that will be coded here in Appendix 6. Parties might wish to provide additional coding on their documents to assist each other in identifying items that relate to specific issues, if such an approach is being considered it needs to be agreed before the review/coding process starts.

A6.6 An example of the coding description is shown below:

PDF documents:

PDF documents should have the following information provided with them, unless it is inappropriate to that document type:

1) *Date of Document (in numeric format as DD/MM/YYYY i.e "15/10/2013").*

2) *A field which states whether or not the date has been estimated. If dates are estimated, the parties should explain the convention they will use to show missing day / month / year.*

3) *Author of Document.*

4) *Addressee of Document (if any).*

5) *Document Title (or file name, email subject line or brief description).*

6) *File Type.*

7) *Names of persons to who copies were sent.*

Exchange and Inspection of Electronic Documents

A6.7 The aim of this section of the appendix is to set out details of the methods to be used to identify the documents to be exchanged and the information that will be provided with each document. The actual mechanism of exchanging the documents is covered in the section below on the ESI Exchange Protocol.

A6.8 Because of the volume of electronic material it might be appropriate to dispense with the production of lists as in practice all the ESI material will be exchanged.

Handling Privileged Documents

A6.9 If the parties agree that privileged documents should be listed (or agree that litigation privileged documents should be listed, excluding legal advice privileged documents), then they could do so in the same way that a disclosure list might be generated using the database where the electronic documents are stored. The only additional field (or column) to be added to the list would be to explain the basis for the claim of privilege. Thus, each privileged document listed would have the following information: document identification number, date, document type, document title, parent/attachment, basis for claiming privilege (the latter need only be a few words to explain, for example, that a document was created on or after the date identified for when litigation was a reasonable prospect, it was and remains a confidential document and it was created for the dominant purpose of obtaining evidence for and/or legal advice in respect of litigation).

ESI Exchange Protocol

A6.10 The exchange of ESI data requires a detailed protocol to be agreed so that the information can be loaded into litigation support systems. There are two main approaches that can be adopted, either let your supplier deal with this, or become involved in specific details.

A6.11 For the first approach the TCC protocol should reflect what electronic database parties are using to store, process, filter and review all documents collected, and then pass the technical burden of working out an exchange protocol to suppliers, for example:

Organisation	***Supplier***	***Product***
Party 1	*Supplier 1*	*Product 1*
Party 2	*Supplier 2*	*Product 2*
Party 3	*Supplier 3*	*Product 3*

Each supplier warrants that their product is capable of:

1) *Exporting data in a file format specific to the other products, e.g. a load file for a Product 1 or Product 2 database.*

2) *Exporting data, in a format to be agreed between suppliers, such that information is capable of being loaded into the other products.*

It is the responsibility of each supplier to articulate to the others the format of their load file and how they wish to receive data.

A6.12 The second approach is more complicated and might require assistance from an external organisation, if you have not already formed a partnership with one. Here you detail the manner in which information will be exchanged. An example of such an approach is shown below, it duplicates some of the information shown in other sections of this protocol as it is intended to be a standalone document that can be shared with both the other parties and their litigation support organisations / suppliers.

System

1. Our current intention is to host the disclosure documents on Product 1.

File Format

2. Disclosure documents are to be exchanged in native format where possible. For those documents that are not disclosed in a native format, they should be in single page tiff format. [As to emails, these will be displayed in HTML or MHTML format with attachments disclosed as separate documents.]

3. Any native format documents that are not inherently text searchable (eg, non-searchable PDFs) and all tiff format documents are to be provided with a separate OCR file, so far as it is possible to do so. [For the avoidance of doubt, our Product 1 database uses the native document for the purposes of text searching. As such, we do not create a separate OCR file for native format documents that are already text searchable.]

4. The image files, and OCR or native documents, should be named so as to be identifiable both within the load-file and against each other.

5. Each tiff image should be provided in:

5.1 the same orientation as the document it mirrors;

5.2 black and white unless colour is necessary for the understanding of the document; and

5.3 300 dpi resolution.

6. [Any tiff images will be branded with the Disclosure List Number of the document.]

[Note: Native documents cannot be branded in this way so, where disclosure is a mix of native and tiff documents, the branding will be inconsistently applied. It is also worth bearing in mind whether the presence of the branding on tiff images will become a problem later in the proceedings (eg, preparation of trial bundles). If so, the branding should be placed away from the bottom right corner of the document where a trial bundle page number will usually be placed.]

De-duplication

7. Both parties should apply an industry standard de-duplication method to remove as many exact duplicates from their disclosure as reasonably possible. Our ESI provider will carry out de-duplication using the [ ] algorithm. Only standalone duplicates or entire duplicate families will be de-duplicated, whilst duplicate documents that are attached to non-duplicate documents will not be de-duplicated.

Irrelevant documents

8. *[Where an irrelevant document is a covering or attachment document to a disclosable document, the irrelevant document will be disclosed for the purposes of providing context.]*

   OR

   [Where an irrelevant document is a covering or attachment document to a disclosable document, the irrelevant document will not be disclosed and a tiff placeholder will appear in its place.]

Data exchange

9. *The parties shall provide a list of the disclosure documents in Excel format on [DATE]. The list will contain the information set out in paragraphs 13 to 30 below for each of the documents. [The documents in the list will be sorted into chronological order, save that attachments will follow their source documents in the order in which they originally appeared.]*

10. *To allow this information to be loaded on to a database, it will also be included in an appropriately formatted Excel Comma Separated Values ("CSV") load-file which will be provided on [DATE], at the same time as the native documents/images and OCR (if applicable).*

11. *In addition to a CSV file, an Image Cross Reference File ("Image x-ref") Excel CSV file will be provided on [DATE] which will link the CSV data with the relevant native electronic documents or tiff images. Further details on the content of this file are set out in paragraph 31 below. A similar Excel CSV file will be provided on [DATE] which will link the CSV data with the relevant OCR file ("OCR x-ref"). Further details on the content of this file are set out in paragraph 32 below.*

12. *Wherever possible, each document shall have coded data for the fields below.*

"Disclosure list number"

13. *The parties shall list documents sequentially and assign each electronic document a consecutive six digit disclosure list number ("0000001"), preceded by a way of identifying which party's document it is (so, [GIVE EXAMPLES]) and which delivery ("A" for the first delivery, "B" for the second and so on). So, for example, the first document in our client's first delivery would be "[????]A000001".*

14. *Attachments or enclosures shall be listed and numbered separately from their parent or covering document (see "Attached to document" below). Appendices, annexes, schedules and exhibits shall be treated as forming part of the document to which they relate.*

"Document type"

15. *The document type shall be an unambiguous, readily identifiable and consistent description of the nature of the document (e.g. "letter", "email" or "memorandum").*

[Note: Obviously, this field is normally:

- only available for documents that have been manually coded (so, primarily former hard copy documents) and emails (because these are easily identifiable); and
- not available for attachments to emails or 'loose' electronic documents (because, for example, the metadata for a Word document would not record that it was actually a letter).]

"File extension"

16. The document format shall be the document's original file extension (e.g. .xls, .doc, .pdf, .msg etc.).

[Note: This will, of course, only be applicable to native electronic documents.]

"Document title"

17. Wherever possible, the parties shall provide data in respect of the document title of each document. This field will be completed with information taken:

17.1 from any manual coding (so, usually formerly hard copy documents);

17.2 verbatim from the "subject" line of emails; and

17.3 from the native filename for attachments to emails or loose electronic documents.

"Date"

18. Wherever possible, the parties shall provide the date of documents. For email items the date will be taken from the date the email was sent or, if not available, the date it was received. For attachments to emails or loose electronic documents, the date should be taken from the "Date Last Modified" metadata. For manually coded documents (so, usually, the former hard copy documents), the date should be as taken from the face of the document.

19. Dates shall be expressed using the format DD MMM YYYY (e.g. 01 Jan 2000).

20. For manually coded hard copy documents, if the day (DD) or month (MMM) part of the date is missing, then "01" will be used to record the missing day element and/or "JAN" used to record the missing month element. If the year (YYYY) element of the date is missing, the document should be deemed undated and the date field left blank.

[Note: Be aware of the different dating convention used in the US where the month is usually provided first and then the day, i.e. 3 February 2014 would in the US be written numerically as 02/03/14 whereas in the UK it is 03/02/14.

"Time"

21. Wherever possible, the parties shall provide the Time for documents. For email items the time will be taken from the time the email was sent or, if not available, the time it was received. For attachments to emails or loose electronic documents, the time should be taken from the "Date Last Modified" metadata. No time will be provided where documents have been manually coded (again, usually the former hard copy documents). [The supplied data will be formatted as HH:MM:SS and based on a 24 hour clock].

[Note: Where documents have been collected from time zones other than GMT it may be necessary to specify whether the time information is provided "as is" or has been standardised to GMT.]

"From", "To", "CC" and "BCC" ("parties information")

22. So far as possible, the parties shall describe an individual and his or her organisation in the parties information using one standardised format throughout, namely that of "[Surname], [First Name] of [Organisation]".

23. The chosen format shall use consistently two delimiters, one (a comma) between the surname and first name, and the other (the word "of") between the individual and the organisation.

24. Where available, an individual's organisation shall be provided and organisations shall be consistently and uniquely described. Where no such information exists, or it cannot be ascertained, the entry should be left blank.

25. Where more than one individual is to be entered into a field, the parties shall separate the descriptions of each individual and his or her organisation by the use of a pipe (|) (e.g. "Smith, John of ABC Ltd|Doe, John of XYZ Plc").

"Attached to document"

26. When listing an attachment the Disclosure list number of its parent or covering document should always be included in the "Attached to document" field.

"Redacted"

27. Where a document had been redacted, the parties shall indicate this by entering "Yes" in this field.

28. Where a document has been supplied in a redacted format, the redaction should be made in white with a black surround. [The grounds for redaction should also be provided. If possible, the grounds for redaction should be inserted in the redaction box.]

"CSV delimiters"

29. All multi-value entries will be separated by the use of a pipe (|).

30. In the event that either party is unable to provide an Excel CSV file and is, instead, providing a TXT CSV file, all TEXT data should be encapsulated with the ^ character (e.g. at the beginning and end of the title).

Image x-ref file

31. The Image x-ref file will contain an entry for each page to be uploaded into the database and will include the following information:

31.1 Disclosure List Number - as set out in paragraph 13 above.

31.2 Filename - for native files this would be the Disclosure List Number followed by the file extension. If the disclosed document is in the format of single paged tiff files, each tiff file should be listed with the Disclosure List Number followed by an underscore ("_") and a sequential page number (padded to four digits) starting with "0001" for the first page of each document.

31.3 File path - this will identify the folder within which the page(s) for the document can be found. We will generally use a top level folder named "Images" and then, depending on

volume, the next level will be subdivided by every 1,000 documents with the first Disclosure List Number being the name of the folder. Each document will then reside in a folder named as the Disclosure List Number.

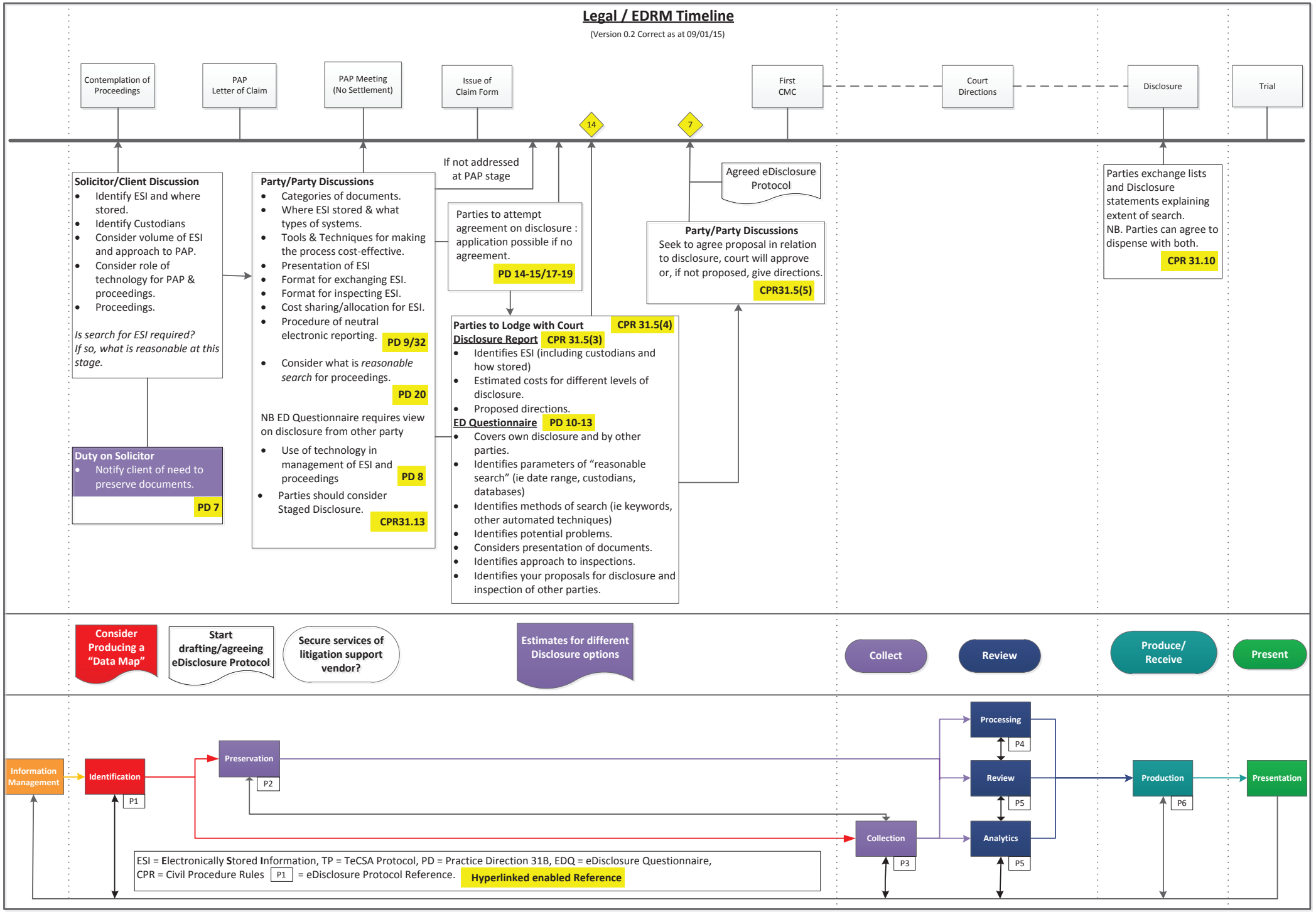
OCR x-ref file

32. The OCR x-ref file will contain an entry for each document to be uploaded into the database and will include the following information:

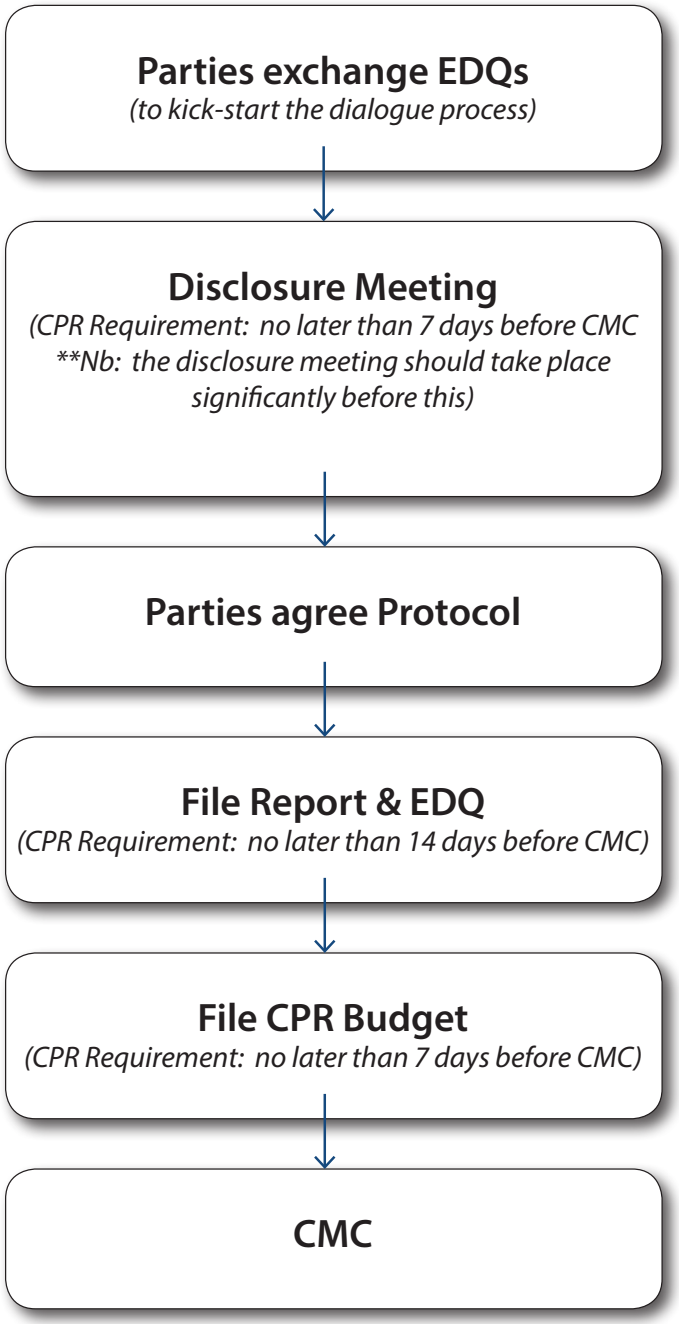
32.1 OCR Filename - The OCR filename should be listed in this field and named as the Disclosure List Number followed by the extension.

32.2 File path - this will identify the folder within which the page(s) for the document can be found. We will generally use a top level folder named "Images" and then, depending on volume, the next level will be subdivided by every 1,000 documents with the first Disclosure List Number being the name of the folder. Each document will then reside in a folder named as the Disclosure List Number.

ANNEX A



ANNEX B
Suggested Pathway to the First CMC



ANNEX C
Project Checklists

(Full set of documents available at http://www.tecsa.org.uk/e-disclosure)

Item	Task	Status	As Of	Notes
A.	**Initial discussions with Client**			
1.	Establish responsibility for maintaining disclosure audit trail			
2.	Determine date range for the matter			
3.	Obtain organization charts for relevant periods			
4.	Identify corporate contacts			
5.	Identify key corporate IT personnel			
6.	Identify key records managers			
7.	Identify key witnesses/custodians			
8.	Identify other witnesses/custodians			
9.	Obtain copy of document retention policy			
10.	Determine if policy has been followed			
11.	Identify and address any imminent document destruction issues			
12.	Gain detailed understanding of e-mail systems			
13.	Determine whether any changes have been made to the method of data storage during the relevant period that might have caused a change to metadata.			
14.	Determine which e-mail systems are involved			
15.	Determine how long e-mail remains on system			
16.	Determine how e-mail is backed up			
17.	Determine if there is a disaster recovery plan			
18.	Determine how other data backed up			
19.	Determine if there are legacy systems involved			
20.	Determine which voice mail system(s) and instant messaging systems involved			
21.	Locate or create data map			

22.	Identify sources of privileged information			
23.	Develop list of all inside and outside counsel			
24.	Determine if client has preferred vendors			
25.	Develop outline of legal and factual issues			
B.	**Litigation Hold/Document Preservation**			
26.	Confirm that destruction policies have been suspended			
27.	Conduct litigation hold strategy meeting			
28.	Determine scope of hold			
29.	Determine recipients of hold			
30.	Coordinate with HR re incoming/departing employees subject to hold			
31.	Determine if third parties have relevant data			
32.	Consider preservation notice to third parties			
33.	Determine where and how to hold data			
34.	Issue litigation hold communication			
35.	Schedule periodic follow-up reminders re litigation hold			
36.	Receive confirmation of hold instructions from recipients			
37.	Inventory data sources			
38.	Determine reasonable/unreasonable accessibility of data sets.			
39.	Determine if back-up tapes are implicated			
40.	Consider setting aside system snapshot			
41.	Determine if data exists in the cloud; develop strategy for preserving same			
42.	Determine if home computers or personal e-mail accounts are implicated			
43.	Determine if Instant Messaging implicated			
44.	Send opposition appropriate preservation demand			
C.	**Preparing for discussions with other parties**			
45.	What are the issues in the case?			
46.	Who are the key players in the case?			

47.	Who are the persons most knowledgeable about ESI systems?			
48.	What events and intervals are relevant?			
49.	When did preservation duties and privileges attach?			
50.	What data are at greatest risk of alteration or destruction?			
51.	Are systems slated for replacement or disposal?			
52.	What steps have been or will be taken to preserve ESI?			
53.	What third parties hold information that must be preserved, and who will notify them?			
54.	What data require forensically sound preservation?			
55.	Are there unique chain-of-custody needs to be met?			
56.	What metadata are relevant, and how will it be preserved, extracted and produced?			
57.	What are the data retention policies and practices?			
58.	What are the backup practices, and what tape archives exist?			
59.	Are there legacy systems to be addressed?			
60.	How will the parties handle voice mail, instant messaging and other challenging ESI?			
61.	Is there a preservation duty going forward, and how will it be met?			
62.	Is a preservation or protective order needed?			
63.	What e-mail applications are used currently and in the relevant past?			
64.	Are personal e-mail accounts and computer systems involved?			
65.	What principal applications are used in the business, now and in the past?			
66.	What electronic formats are common, and in what anticipated volumes?			
67.	Is there a document or messaging archival system?			
68.	What relevant databases exist?			
69.	Will paper documents be scanned, and if so, at what resolution and with what OCR and metadata?			
70.	What search techniques will be used to identify responsive or privileged ESI?			
71.	If keyword searching is contemplated, can the parties agree on keywords?			
72.	Can supplementary keyword searches be pursued?			
73.	How will the contents of databases be disclosed? Queries? Export? Copies? Access?			

74.	How will de-duplication be handled, and will data be re-populated for production?			
75.	What forms of production are offered or sought?			
76.	Will single- or multipage .tiffs, PDFs or other image formats be produced?			
77.	Will load files accompany document images, and how will they be populated?			
78.	How will the parties approach file naming and unique document identification numbering?			
79.	Will there be a need for native file production? Quasi-native production?			
80.	On what media will ESI be delivered? Optical disks? External drives? FTP?			
81.	How will we handle inadvertent production of privileged ESI?			
82.	How will we protect trade secrets and other confidential information in the ESI?			
83.	Do regulatory prohibitions on disclosure, foreign privacy laws or export restrictions apply?			
84.	How do we resolve questions about printouts before their use at the trial?			
85.	Will it be necessary to authenticate documents to be used at the trial? How will this be done?			
86.	What ESI will be claimed as not reasonably accessible, and on what bases?			
87.	Who will serve as liaisons or coordinators for each side on ESI issues?			
88.	Will technical assistants be permitted to communicate directly?			
89.	Would it be helpful to engage a neutral person to supervise eDisclosure and to mediate/arbitrate if differences arise?			
90.	Can any costs be shared or shifted by agreement?			
91.	Can cost savings be realized using shared vendors, repositories or neutral experts?			
92.	How much time is required to identify, collect, process, review, redact and produce ESI?			
93.	How can production be structured to accommodate depositions and deadlines?			
94.	When is the next Case Management Conference (more than one CMC may be needed)?			
D.	**Collection**			
95.	Develop data collection plan			
96.	Identify and retain collection vendor if required			
97.	Identify sources of data			
98.	Determine who will collect the data			
99.	Prepare and interview IT staff re systems, back-ups, etc.			

100.	Estimate amount of data			
101.	Determine data formats			
102.	Identify any problems in relation to databases			
103.	Determine if any data is encrypted			
104.	Determine if there are unique software applications			
105.	Anticipate which data may require native production			
106.	Determine if computer forensics implicated			
107.	Identify and interview key custodians			
108.	Develop plan for hard copy data collection			
109.	Determine OCR strategy for paper			
110.	Determine extent of coding required for paper			
111.	Maintain chain of custody for data gathered			
112.	**Processing**			
113.	Determine culling software and strategies; keyword list; date range limitations, etc.			
114.	Select processing vendor(s)			
115.	Obtain cost estimates for processing			
116.	Obtain time estimates for processing			
117.	Insure chain of custody for data			
118.	Determine which metadata fields should be extracted			
119.	Determine procedures for dealing with exceptions			
120.	Confirm load file formats			
F.	**Review**			
121.	Determine review platform and process			
122.	Formulate and test keyword search terms; test with key custodian(s)			
123.	Determine review team composition			
124.	Determine if second review required/warranted			
125.	Train review team			

126.	Conduct intensive review of key custodian(s) data			
127.	Develop budget estimate for review			
128.	Develop time estimate for review			
129.	Load data for review			
130.	Review data for relevance and privilege			
131.	Develop protocol for redaction			
132.	Create privilege log			
G.	**Production**			
133.	Determine priority of data to be produced; consider rolling productions.			
134.	Negotiate and obtain appropriate protective order re data including clawback agreement			
135.	Determine desired production format(s)			
136.	Negotiate production format(s) with opposition			
137.	Negotiate timetable for production(s)			
138.	Negotiate timetable for receiving production(s)			
	Roles			
	IT = Client IT staff			
	IC = In-house legal tem			
	OC = Outside solicitors			
	LS = Outside solicitors' Litigation Support team			
	V = Outside Vendor			