GUIDELINES
TO THE TeCSA / SCL / TECBAR
eDISCLOSURE PROTOCOL

Version 0.1
1 November 2013

- ***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 Appendix 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".***

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 Appendix B) is a flowchart showing the suggested pathway for parties to follow in respect of disclosure prior to the first CMC.

1. IDENTIFICATION OF SOURCES OF DOCUMENTATION

1.1 Timing is important – the process of identifying sources of documentation 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 documentation 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 documentation is preserved (by, for example, ensuring that any standard or routine documentation destruction policy is suspended for the duration of the dispute; and that any devices that contain documentation in any format are not destroyed);

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

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".

Native Documents

3.3 The optimum way to collect electronic documentation 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.).

3.4 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.

Non-Native Documents

3.5 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. Non-

searchable PDF documents should be made searchable by applying an Optical Character Recognition ("OCR") process. Disclosing non-searchable PDFs should be the exception, not the rule.

3.6 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.

3.7 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.

3.8 As a minimum the following coded fields will be required for <u>all</u> documents:

(1) Date of Document (in alphanumeric format as "01 Jan 2013")

(2) Author of Document

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

(4) a field which states whether or not the date has been estimated.

It may also be considered helpful to include:

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

(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).

3.9 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.

3.10 Occasionally, some parties use TIFF as a format in which to disclose documentation. 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.

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

3.11 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.

3.12 The choices the parties can make are anything from dispensing with disclosure, to arbitration-style disclosure, to the "keys to the warehouse" approach (ie 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.

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 documentation 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 the date range(s) to be applied to the party's disclosable documentation. 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 documentation. 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.

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.

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 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 documentation which contains 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 documentation may be missed by keyword filtering; and in other instances, the keywords may fail to capture all disclosable documentation. 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 documentation 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 documentation 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 documentation.

6.7 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 3 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 documents stored on portable media.)

6.7.2 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

6.8 As can be seen from above, it is not mandatory to list each and every document that is being disclosed. It may assist, however, to provide a list in advance of inspection so that the parties can review the nature and extent of the documentation 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.

6.9 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 documentation over which litigation privilege is being claimed) so that each party can review and assess the documentations 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 privileged 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.

6.10 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 100% sure 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,)

6.11 It should be noted that whatever may be agreed between the immediate parties in the eDisclosure Protocol, 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. Even if the agreement is repeated in a direction of the Court, in the absence of legislation this cannot be assumed to be effective vis-a-vis third parties or additional parties.

6.12 Parties may wish to agree that the paragraph(s) dealing with no waiver of privilege are repeated in a direction of the Court and/or are the subject of a legally binding agreement.

APPENDIX A

Legal / EDRM Timeline

(Version 1.0 Correct as at 01/11/13)

Contemplation of Proceedings

PAP Letter of Claim

PAP Meeting (No Settlement)

Issue of Claim Form

14

7

First CMC

Court Directions

Disclosure

Trial

Solicitor/Client Discussion

- Identify ESI and where stored.
- Identify Custodians
- Consider volume of ESI and approach to PAP.
- Consider role of technology for PAP & proceedings.
- Proceedings.

Is search for ESI required?
If so, what is reasonable at this stage.

Duty on Solicitor

- Notify client of need to preserve documents.

PD 7

Party/Party Discussions

- Categories of documents.
- Where ESI stored & what types of systems.
- Tools & Techniques for making the process cost-effective.
- Presentation of ESI
- Format for exchanging ESI.
- Format for inspecting ESI.
- Cost sharing/allocation for ESI.
- Procedure of neutral electronic reporting. **PD 9/32**
- Consider what is *reasonable search* for proceedings. **PD 20**

NB ED Questionnaire requires view on disclosure from other party

- Use of technology in management of ESI and proceedings **PD 8**
- Parties should consider Staged Disclosure. **CPR31.13**

If not addressed at PAP stage

Parties to attempt agreement on disclosure : application possible if no agreement. **PD 14-15/17-19**

Parties to Lodge with Court **CPR 31.5(4)**

Disclosure Report **CPR 31.5(3)**

- Identifies ESI (including custodians and how stored)
- Estimated costs for different levels of disclosure.
- Proposed directions.

ED Questionnaire **PD 10-13**

- Covers own disclosure and by other parties.
- Identifies parameters of "reasonable search" (ie date range, custodians, databases)
- Identifies methods of search (ie keywords, other automated techniques)
- Identifies potential problems.
- Considers presentation of documents.
- Identifies approach to inspections.
- Identifies your proposals for disclosure and inspection of other parties.

Agreed eDisclosure Protocol

Party/Party Discussions
Seek to agree proposal in relation to disclosure, court will approve or, if not proposed, give directions. **CPR31.5(5)**

Parties exchange lists and Disclosure statements explaining extent of search.
NB. Parties can agree to dispense with both. **CPR 31.10**

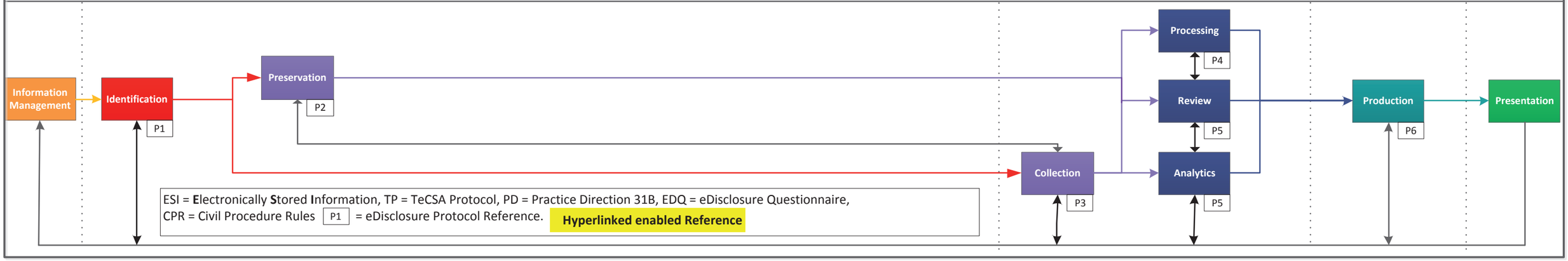
Consider Producing a "Data Map"

Start drafting/agreeing eDisclosure Protocol

Secure services of litigation support vendor?

Estimates for different Disclosure options

Collect | **Review** | **Produce/ Receive** | **Present**



ESI = **E**lectronically **S**tored **I**nformation, TP = TeCSA Protocol, PD = Practice Direction 31B, EDQ = eDisclosure Questionnaire,
CPR = Civil Procedure Rules P1 = eDisclosure Protocol Reference. **Hyperlinked enabled Reference**

APPENDIX B
Suggested Pathway to the First CMC