

respect of each of the key witnesses, will give us the potential number of documents in respect of that as well. It is at that stage, when that crude way of finding out what documents might be in existence is completed, that a service provider will have to be agreed between the parties, and will have to be instructed to look at what the next stage of the exercise should involve and how much it is going to cost, in order to produce a corpus of documents which is reviewable by both parties.

27. At the moment we are just staring into open space as to what the volume of the documents produced by a search is going to be. I suspect that in the long run this crude search will not throw up more than a few hundred thousand documents. If that is the case, then this is a prime candidate for the application of software that providers now have, which can de-duplicate that material and render it down to a more sensible size and search it by computer to produce a manageable corpus for human review – which is of course the most expensive part of the exercise. Indeed, when it comes to review, I am aware of software that will effectively score each document as to its likely relevance and which will enable a prioritisation of categories within the entire document set.
28. It is also possibly going to be necessary to look at whether we should in fact be running all of these key words at this stage. In my judgment, that is what the exercise should be. There should be disclosure of electronically stored information. It is clear that documents created by these four witnesses exist which are likely to support the claimants' case and damage the defendant's. The only question is how we go about finding them. I think the proper thing to do is to start with a fairly crude search and then, if the numbers are within reason, to work with experts to render the corpus of documents down and de-duplicate them and then move on to the review stage. The parties should consider their obligations under PD 31 2A to "*... discuss any issues that may arise regarding searches for and the preservation of electronic documents. This may involve the parties providing information about the categories of electronic documents within their control, the computer systems, electronic devices and media on which any relevant documents may be held, the storage systems maintained by the parties and their document retention policies.*" And further discuss the defendants' proposed search methodology, once they have absorbed my views in this judgment and the defendants must also complete the questionnaire in the schedule below as a means of providing the claimant and the court with the necessary information in a structured manner, should there be any further application for directions on disclosure.

Master Whitaker

The Senior Master

SCHEDULE

ESI QUESTIONNAIRE^[1]

Part 1 – Your disclosure

Extent of a reasonable search

Date range and custodians

1. What date range do you consider that your searches for documents should cover ("the date range")?
2. Identify the custodians or creators of your documents whose repositories of documents you consider should be searched.^[2]

Communications

3. Which forms of electronic communication were in use during the date range (so far as is relevant to these proceedings)?

A Communication	B In use during the date range? (yes/no)	C Are you searching for relevant documents in this category? (yes/no)	D Where and on what type of software/equipment/media is this information stored?^[3]	E (a) Are back-ups or archives of this information available, and (b) if so, are you searching the back-ups or archives?
i) E-mail ^[4]				
ii) Other (provide details for each type ^[5])				

Electronic documents

4. Apart from attachments to e-mails, which of the following other forms of electronic documents were created or stored by you during the date range?

A Document Type	B In use during the date range? (yes/no)	C Are you searching for relevant documents in this category? (yes/no)	D Where and on what type of software/equipment/media is this information stored?^[6]	E (a) Are back-ups or archives of this information available, and (b) if so, are you searching the back-ups or archives?
i) Word (or equivalent – state which)				
ii) Excel (or equivalent - state which)				
iii) Imaged documents ^[7]				
iv) Other ^[8] (state which)				

Databases of electronic records

5. In the following table identify database systems, including document management systems, which may contain data which may be disclosable and which were used by you during the date range.

A Name	B Brief description	C Nature of data held	D Are you disclosing data held in this database? (yes/no)	E Proposals for provision of relevant data to or access by other parties to this litigation
1.				
2. (etc)				

Method of search

Key words

6. Do you consider that keyword searches should be used as part of the process of determining which

documents you should disclose?

If yes, provide details of

(1) the keywords used or to be used (by reference, if applicable, to individual custodians, creators, repositories, file types and/or date ranges),^[9]

(2) the extent to which the keyword searches have been or will be supplemented by a review of individual documents.

Other types of automated searches

7. Do you consider that automated searches or automated techniques other than keyword searches (eg concept searches or clustering) should be used as part of the process of determining which documents you should disclose?

If yes, provide details of

(1) the process(es) used or to be used (by reference, if applicable, to individual custodians, creators, repositories, file types and/or date ranges),

(2) the extent to which the processes have been or will be supplemented by a review of individual documents, and

(3) how the methodology of automated searches will be made available for consideration by other parties.

8. If the answer to Question 6 or 7 is yes, state whether attachments to (a) e-mails (b) compressed files (c) embedded files and (d) imaged text will respond to your keyword or other automated search.
9. Are you using or intending to use computer software for other purposes in relation to disclosure (eg de-duplication)? If so, provide details of the software, processes and methods to be used.^[10]

Potential problems with the extent of search and accessibility of documents

10. Do any of the sources and/or documents identified in this Questionnaire raise questions as to the reasonableness of the search which ought to be taken into account? ^[11] If so, give details.
11. Are any documents which may be disclosable encrypted, password-protected or for other reasons difficult to access, or do you have any reason to believe that they may be? ^[12] If so, state which of the categories identified at Questions 3, 4 and 5 above are affected, and your proposals for making them accessible.
12. Are you aware of any other points in relation to disclosure of your electronic documents which require discussion between the parties? If so, give details.

Preservation of ESI

13. Do you have a document retention policy?
14. Have you given an instruction to preserve ESI, and if so, when?

Inspection

15. Subject to re-consideration after receiving the responses of other parties to this Questionnaire, (a) in what format and (b) on what media do you intend to provide to other parties copies of disclosed

documents which are or will be available in electronic form?

- 16. Subject to re-consideration after receiving the responses of other parties to this Questionnaire, do you intend to provide other parties with disclosure data^[13] electronically, and if so, (a) in what format (eg CSV, MS Word, MS Excel) and (b) on what media?
- 17. Insofar as you have available or will have available searchable OCR versions of electronic documents, do you intend to provide the searchable OCR version to other parties?^[14] If not, why not?

Part 2 – The disclosure of other parties

The extent and content of their search

- 18. Do you at this stage have any proposals as to the date ranges which should be searched by other parties to the action? If so, provide details.
- 19. Do you at this stage have any proposals as to the custodians or creators whose repositories of documents should be searched for disclosable documents by other parties to the action? If so, provide details.^[15]
- 20. Do you consider that the other party(ies) should disclose all available metadata^[16] attaching to any documents? If yes, provide details of the documents or categories of documents.

Proposals for the method to be adopted for their searches

- 21. Do you at this stage have any proposals as to the keyword searches, or other automated searches, which should be applied by other parties to their document sets? If so, provide details.

Inspection

- 22. Subject to re-consideration after receiving the responses of other parties to this Questionnaire, (a) in what format and (b) on what media do you wish to receive copies of disclosed documents which are or will be available in electronic form?
- 23. Subject to re-consideration after receiving the responses of other parties to this Questionnaire, do you wish to receive disclosure data^[17] electronically, and if so, (a) in what format (eg .CSV, MS Word, MS Excel) and (b) on what media?

STATEMENT OF TRUTH

The [] believes that the facts stated in the answers to this questionnaire are true and I am duly authorised by the [] to sign this statement

Signed:

Name:

Position held: [Claimant] [Defendant] [Claimant's / Defendant's Solicitor] *delete as appropriate*

Date:

WARNING: Unless the Court makes some other order, the answers given in this document may only be used for the purposes of the proceedings in which the document is produced unless it has been read to or by the court or referred to at a hearing which has been held in public or the Court gives permission or the party who has completed this questionnaire agrees.

BAILII note: The Guidance Notes and Glossary prepared by the committee which produced the questionnaire set out in the Schedule to this judgment is [available here](#).

Note 1 See Schedule 3 of the Practice Direction Governing Disclosure Of Electronically Stored Information (CPR PD 31 B) for guidance on answering this Questionnaire and the Glossary of technical expressions in Schedule 1 of PD 31 B. [\[Back\]](#)

Note 2 Include names of all those who may have or have had custody of disclosable documents, including secretaries, personal assistants, former employees and/or former participants. It may be helpful to identify different dates for particular custodians. [\[Back\]](#)

Note 3 State the geographical location (if known). Consider (at least) servers, desktop PCs, laptops, notebooks, handheld devices, PDA devices, off-site storage, removable storage media (eg CD-ROMs, DVDs, USB drives, memory sticks) and databases. [\[Back\]](#)

Note 4 Consider all types of e-mail system (eg Outlook, Lotus Notes, web-based accounts), whether stored on personal computers/portable devices or in web-based accounts (e.g. Yahoo, Hotmail, Gmail etc.). [\[Back\]](#)

Note 5 For example, instant messaging, voicemail, VOIP (Voice Over Internet Protocol), recorded telephone lines, text messaging, audio files, video files etc. [\[Back\]](#)

Note 6 State the geographical location (if known). Consider (at least) servers, desktops and laptops. [\[Back\]](#)

Note 7 For example, .pdf, .tif, .jpg. [\[Back\]](#)

Note 8 For example, Powerpoint or equivalent, specialist documents (such as CAD Drawings) etc. [\[Back\]](#)

Note 9 Where keyword searches are used in order to identify irrelevant documents which are to be excluded from disclosure (for example a confidential name of a client or customer), a general description of the type of search may be given. [\[Back\]](#)

Note 10 Eg how de-duplication is to be carried out. [\[Back\]](#)

Note 11 See Practice Direction, which refers to the following matters which may be relevant: (a) the number of documents involved, (b) the nature and complexity of the proceedings, (c) the ease and expense of retrieval of any particular document, and (d) the significance of any document which is likely to be located during the search. [\[Back\]](#)

Note 12 Eg back-ups, archives, off-site or outsourced document storage, documents created by former employees, documents stored in other jurisdictions, documents in foreign languages. [\[Back\]](#)

Note 13 Disclosure data is data relating to disclosed documents, including for example the type of document, the date of the document, the names of the author/sender and the recipient, and the party disclosing the document. [\[Back\]](#)

Note 14 There is no requirement that you should obtain OCR versions of documents, and this question is directed only to OCR versions which you have available or expect to have available to you. If you do provide OCR versions to another party, they will be provided by you on an "as is" basis, with no assurance to the other party that the OCR versions are complete or accurate. You may wish to exclude provision of OCR versions of documents which have been redacted. [\[Back\]](#)

Note 15 Include names of all those who may have or have had custody of disclosable documents, including secretaries, personal assistants, former employees and/or former participants. It may be helpful to identify different dates for particular custodians. [\[Back\]](#)

Note 16 "Metadata" is information about the document or file which is recorded in the computer, such as the date and time of creation or modification of a word-processing file, or the author and the date and time of sending of an e-mail. The question is directed to the more extensive metadata which may be relevant where for example authenticity is disputed. [\[Back\]](#)

Note 17 See footnote to Question 16. [\[Back\]](#)

[BAILII note:

The Guidance Notes and Glossary prepared by the committee which

*produced the questionnaire set out in the Schedule to this judgment
are available at www.scl.org/site.aspx?i=xb537].*

BAILII: [Copyright Policy](#) | [Disclaimers](#) | [Privacy Policy](#) | [Feedback](#) | [Donate to BAILII](#)
URL: <http://www.bailii.org/ew/cases/EWHC/QB/2009/B41.html>