

LiST GROUP
DATA EXCHANGE PROTOCOL

PART 1 OF 2

ELECTRONIC DISCLOSURE DOCUMENTS

v. 1.1

(Draft 04 May 2006)

Responses are invited to LiST's Data Exchange Working Group at

def@listgroup.org

by 29 September 2006

LiST

The Litigation Support Technology Group

<http://www.listgroup.org>

LiST Group

Data Exchange Protocol Part 1 of 2

ELECTRONIC DISCLOSURE DOCUMENTS

CONTENTS

1.	About this Protocol	1
1.1	Introduction	1
1.2	What this Protocol does	1
1.3	What this Protocol does not do	1
1.4	Target audience	1
1.5	Why this Protocol is necessary	1
1.6	How to use this Protocol	1
1.7	Who is behind this Protocol	2
2.	Potential core reference documents	2
3.	Details that may need to be provided	2
4.	Preparing to provide electronic Disclosure Documents	2
4.1	Disclosure Documents in Native Format	2
4.2	Electronic Images of Disclosure Documents	3
4.3	Exchange media	3
5.	General technical guidance	4
5.1	Naming files	4
5.2	Directory structure	5
6.	The authors	6
	Glossary and definitions	7

1. **About this Protocol**

1.1 **Introduction**

For simplicity LiST's Data Exchange Protocol is provided in two parts:

- (A) this part, Part 1, deals exclusively with electronic [Disclosure Documents](#);
- (B) Part 2 deals exclusively with their associated [Disclosure Data](#) and involves provision of a Data Exchange File.

Parties exchanging both electronic [Disclosure Documents](#) and [Disclosure Data](#) will therefore need to refer to both parts of this Protocol.

This Protocol will be updated by LiST from time to time as practices in this area change, the Civil Procedure Rules are updated and/or technology develops.

1.2 **What this Protocol does**

The LiST Data Exchange Protocol Part 1 for Electronic [Disclosure Documents](#) sets out best practice for providing copies of [Disclosure Documents](#) that are already held in electronic form. It is a fundamental assumption of this Protocol that parties wishing to receive orderly disclosure in a business-like manner must be prepared to provide orderly disclosure in a business-like manner.

Throughout this Protocol reference is made to electronic [Disclosure Documents](#) but the conventions and rules set out below apply equally to any exchange of electronic documents between any organisations (including litigation support bureaux).

1.3 **What this Protocol does not do**

This Protocol neither provides assistance with the creation of electronic documents (for instance, scanning paper documents to [Electronic Images](#)) nor suggests that this be a mandatory exercise on your particular case.

1.4 **Target audience**

This Protocol is aimed primarily at those who are using some form of litigation support system to manage their case. Specifically, it will be of benefit to specialists who have day-to-day responsibility for processing and managing electronic disclosure. Although drafted in relatively simple terms, this Protocol assumes that the reader has a level of technical knowledge sufficient for undertaking electronic disclosure.

1.5 **Why this Protocol is necessary**

The amount of original evidence ([Disclosure Documents](#)) either created electronically or converted to electronic format during the life of a case is increasing, yet the practicalities of dealing with it remain complicated, time-consuming and expensive. This Protocol seeks to simplify and speed up procedures and reduce cost by establishing sensible standards that all parties can follow to improve the provision or exchange of that evidence.

1.6 **How to use this Protocol**

This Protocol is a step-by-step process for the preparation of electronic [Disclosure Documents](#) for provision to, or exchange with, another party. It is intended to be used by agreement between the parties but it is possible that, in the absence of any agreement, an order or direction may be made to the effect that this Protocol be followed.

1.7 Who is behind this Protocol

This Data Exchange Protocol (Parts 1 and 2) has been drafted by a Working Group of litigation support specialists selected from the membership of LiST ([the Litigation Support Technology Group](#)) as representative users of the most common litigation support systems currently available in the UK. The individual authors are cited at the end of this document.

2. Potential core reference documents

- (A) List of [Documents](#)
- (B) [Civil Procedure Rules Part 31](#)
- (C) [Practice Direction 31](#)
- (D) [Draft Practice Direction for the Use of IT in Civil Proceedings](#)
- (E) LiST Data Exchange Protocol Part 2 for [Disclosure Data](#)¹

3. Details that may need to be provided

In order to follow this Protocol it may be necessary to provide:

- (A) for a [Disclosure Document](#) that was originally created electronically:
 - (1) [the Document in its Original Electronic Format](#) (or "[Native Format](#)"²); and/or
 - (2) [the Document converted to an Electronic Image](#)³; and/or
- (B) [the Electronic Image of a scanned paper Disclosure Document](#),
- (C) [on one or more agreed exchange media](#).

The following sections cover these in more detail.

4. Preparing to provide electronic Disclosure Documents

4.1 Disclosure Documents in Native Format

- (A) Copies of electronic [Disclosure Documents](#) will need to be renamed to match either their [Document ID](#) or [Disclosure List Number](#) (see [Naming files](#)⁴, below). An important part of their identity, and perhaps their evidential value, is what they were originally called and where they were originally stored. That information may need to be preserved and may also need to be provided with the [Disclosure](#)

¹ The LiST Data Exchange Protocol Part 2 is currently being drafted and will be released for public comment later in 2006.

² For convenience, we define this here as the electronic format in which the document was originally created (for example MS Word 97 or WordPerfect 5.1).

³ Parties may wish to consider providing an electronic Disclosure Document both as an image and in its native format, either as a matter of course or just for particular file types. Where parties choose to do this the standard naming conventions and procedures set out in this Protocol will accommodate.

⁴ Be aware that only copy electronic Disclosure Documents should be renamed, never original ones, as the process of renaming will irrecoverably change certain of each Document's metadata.

[Data](#) in the Data Exchange File (see Part 2 of LiST's Data Exchange Protocol).

- (B) If agreement has been reached to convert [Disclosure Documents](#) from their [Native Formats](#) to [Electronic Images](#), for instance if part of a document is to be redacted, this should be done electronically, usually by means of a TIFF printer driver, rather than by printing to paper and then scanning.⁵

4.2 Electronic Images of Disclosure Documents

- (A) [Electronic Images](#) of [Disclosure Documents](#) should be provided as single-page TIFF (tagged image file format) image files.⁶ A three page [Disclosure Document](#) will therefore correspond to three separate TIFFs, one for each page.
- (B) [Electronic Images](#) should always be in black and white, unless agreement has been reached otherwise in relation to specific [Disclosure Documents](#).⁷
- (C) [Electronic Images](#) should have a resolution of 300 dots per inch (dpi).⁸
- (D) The compression used should be CCIT Group 4 compression.
- (E) [Electronic Images](#) should be A4 in size.
- (F) [Electronic Images](#) should all be in portrait orientation (where the document itself is not in portrait orientation, it should be rotated so that the top of the document is on the left hand side when the [Electronic Image](#) is viewed).⁹

4.3 Exchange media

Parties will need to discuss and agree the appropriate medium for the provision and receipt of electronic [Disclosure Documents](#). Whatever exchange media agreed upon, no single electronic [Disclosure Document](#) may be split across two or more media if this would require specialist software to be used by the recipient in order to access that [Disclosure Document](#).

⁵ Page breaks, margin widths, etc. are static in images but can be fluid in, for example, Word or Excel documents where they are dependent on external factors such as the printers being used. The proposed method is also cheaper.

⁶ TIFF files contain limited metadata. Single-page TIFFs are used (or supported) by the majority of litigation support systems. Reasonable, free image viewer software is packaged with all Microsoft Windows operating systems.

⁷ Black and white images of the format specified in this Protocol achieve a balance between file size and image quality whilst colour scanning is slower, more expensive and produces larger file sizes.

⁸ We are aware that current practice is to scan paper documents to TIFF at a resolution of 300 dpi. We considered whether it might be sensible when converting electronic documents to TIFF to use a resolution of just 150 dpi. The rationale for this proposal is (a) converting electronic documents to TIFF at 150 dpi results in a TIFF that is sufficiently clear for the human eye, (b) a TIFF with a resolution of 150 dpi can be printed and photocopied with acceptable results, (c) if you have the original electronic document you can use that as the source of the text (which will be 100% perfect) rather than using any [Optical Character Recognition](#) (OCR) software and (d) greatly reduced file sizes (by up to 50%). The potential problem is that if text is not being exchanged, the reduced resolution might impair the later extraction of text by OCR software. We would welcome opinions on this, and on whether a sensible standard would be 300 dpi for paper documents scanned to TIFF and 150 dpi for electronic documents converted to TIFF. Although perhaps sensible, a single dpi for all documents is clearly more elegant.

⁹ This reflects the requirement in para 2(vi), Appendix 10, Admiralty & Commercial Courts Guide 2002.

If possible, exchange media should be provided in read-only form.

5. **General technical guidance**

Some general points apply and should be noted before parties begin to prepare their electronic [Disclosure Documents](#) for exchange.

5.1 **Naming files**

- (A) Every electronic [Disclosure Document](#) must have a file name which is unique within the case for clear, easy and unambiguous identification, notwithstanding the directory structure used ([see below](#)).
- (B) Single-page TIFFs:-
 - (1) The file name should be the [Disclosure Document's Document ID](#) or [Disclosure List Number](#)¹⁰ followed by an underscore, followed by a four digit number (padded with leading zeroes as necessary) corresponding to the page number within the document, followed by the normal ".tif" file extension (see the [examples below](#)).
 - (2) Where necessary, the portion of the file name corresponding to the [Document ID](#), [Disclosure List Number](#) or to the page number can be lengthened to contain additional digits by padding with additional zeroes.¹¹
- (C) For all other electronic [Disclosure Documents](#) (that is, any [Disclosure Documents](#) in their [Original Electronic Format](#)) the file name should be the [Document ID](#) or [Disclosure List Number](#) followed by the appropriate file extension for a file of that type.
- (D) When naming files, the following characters should be avoided as they are conventionally reserved for automated computer functions:
 - (1) Double quotation marks (")
 - (2) Pipe symbol (|)
 - (3) Tilde (~)
 - (4) Question mark (?)
 - (5) Ampersand (&)
 - (6) Comma (,)
 - (7) Forward slash (/)
 - (8) Back slash (\)

¹⁰ Basing the file name on the Disclosure List Number allows electronic Disclosure Documents to be directly linked back to the List of Documents; basing the file name on the Document ID will also achieve this provided that a cross-reference between the Disclosure List Number and the Document ID is supplied in a Data Exchange File (see Part 2 of LiST's Data Exchange Protocol, to be released later in 2006).

¹¹ This convention allows for missing pages to be inserted after the TIFF has been created and ensures that files can be sorted logically.

- (9) Greater than symbol (>)
 - (10) Less than symbol (<)
 - (11) Full stop (.)¹²
 - (12) Colon (:)
 - (13) Asterisk (*)
- (E) The naming convention should be consistent across all of the electronic [Disclosure Documents](#): use either the [Disclosure Documents' Document IDs](#) or [Disclosure List Numbers](#) throughout, not a combination. If [Document IDs](#) are used then it is assumed that a Data Exchange File will also be provided as governed by Part 2 of this Protocol.
- (F) The full file path and file name of the electronic Disclosure Documents being provided, when combined, must be fewer than 255 characters in total length to avoid incompatibilities with certain computer systems.¹³
- (G) Examples:
- ```
123456_0001.tif
Disc-0001-0003_0004.tif
123458.pdf
/Smith v Jones 2435-998/Images/123458/123458_0001.tif
```

## 5.2 Directory structure

- (A) The exchange medium should contain in its root a directory which is named after the case name and action number. All files and sub-directories relating to that case must then be placed within that master directory.
- (B) Where providing electronic [Disclosure Documents](#) in a combination of [Native Formats](#) and [Electronic Images](#), use a separate directory for each generic format. All native files must go into a directory called "Native", and all [Electronic Images](#) into a directory called "Images".
- (C) Where the [Native Format](#) of one of the electronic [Disclosure Documents](#) is itself an [Electronic Image](#), include it in the `Native` directory so that the `Images` directory contains only [Electronic Images](#) made of [Disclosure Documents](#) which have been converted from their [Original Electronic Format](#) or which have been created from paper [Disclosure Documents](#).
- (D) If the [Native Format](#) of any of the electronic [Disclosure Documents](#) is a multi-page image file there is no need to convert it to single-page TIFFs.
- (E) It is acceptable either to put all of the [Electronic Images](#) in the `Images` directory or to put all [Electronic Images](#) making up one document in to their own sub-directory (as shown by the example below). It is equally acceptable to divide the [Electronic Images](#) for [Disclosure Documents](#) into administratively sensible sub-

---

<sup>12</sup> Except when conventionally used as part of the file extension (e.g. `report.doc`)

<sup>13</sup> A full file path and file name in excess of 255 characters will cause some computer operating systems and litigation support systems to be unable to process the electronic Disclosure Documents.

directories (such as the [Electronic Images](#) for Documents 1 to 500 in Images/one and the [Electronic Images](#) for documents 501 to 1,000 in Images/two, and so on).

- (F) It is not acceptable to split the [Electronic Images](#) making up one document across multiple directories.
- (G) The directory structure to be used on the exchange media should therefore appear as follows:

```
/
/Smith v Jones 2435-998
/Smith v Jones 2435-998/Images
/Smith v Jones 2435-998/Images/123456
/Smith v Jones 2435-998/Images/123456/123456_0001.tif
/Smith v Jones 2435-998/Images/123456/123456_0002.tif
/Smith v Jones 2435-998/Images/123456/123456_0003.tif
/Smith v Jones 2435-998/Images/123457/123457_0001.tif
/Smith v Jones 2435-998/Images/123458/123458_0001.tif
/Smith v Jones 2435-998/Native
/Smith v Jones 2435-998/Native/123459.doc
/Smith v Jones 2435-998/Native/123460.xls
/Smith v Jones 2435-998/Native/123461.ppt
/Smith v Jones 2435-998/Native/123462.tif
```

## 6. **The authors**

Reza Alexander (DLA Piper Rudnick Gray Cary UK LLP)

Sandy Bunting (Freshfields Bruckhaus Deringer)

Mark Dingle (Simmons & Simmons)

Jonathan Maas (DLA Piper Rudnick Gray Cary UK LLP)

Vince Neicho (Allen & Overy)

Mark Simmons (Ashurst)

This Protocol does not necessarily reflect the views of the firms mentioned above or of partners within those firms.

## Glossary and definitions

|                                          |                                                                                                                                                                                                                                                                                                                                                 |
|------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Disclosure Data                          | Electronic data which identifies <a href="#">Disclosure Documents</a> , including for example the type of <a href="#">Document</a> , the date of the <a href="#">Document</a> , the names of the author/sender and the recipient, and the party disclosing the <a href="#">Document</a> .                                                       |
| Disclosure <a href="#">Documents</a>     | <a href="#">Documents</a> disclosed by a party in accordance with <a href="#">CPR Part 31</a> .                                                                                                                                                                                                                                                 |
| Disclosure List Number                   | The sequential number allocated to each <a href="#">Disclosure Document</a> within a List of <a href="#">Documents</a> in accordance with <a href="#">CPR Part 31</a> .                                                                                                                                                                         |
| Document                                 | Anything in which information of any description is recorded as defined in <a href="#">CPR Part 31</a> .                                                                                                                                                                                                                                        |
| <a href="#">Document</a> ID              | The unique identification automatically assigned to each of the <a href="#">Disclosure Documents</a> in a party's litigation support system.                                                                                                                                                                                                    |
| Electronic Image                         | An electronic representation of a document (such as a TIF, PDF, JPG, etc.), created either from a paper copy or from its <a href="#">Original Electronic Format</a> .                                                                                                                                                                           |
| Native Format/Original Electronic Format | The electronic format in which the document was originally created (for example MS Word 97 or WordPerfect 5.1). Otherwise referred to as "Original Electronic Format".                                                                                                                                                                          |
| Optical Character Recognition (OCR)      | Recognition of printed or written characters by computer. Each paper page of text is converted to an <a href="#">Electronic Image</a> using a scanner and OCR is then applied to this image to produce a text file. This involves complex image processing algorithms and rarely achieves 100% accuracy so manual proof-reading is recommended. |