

# Straight to the point

Intelligent new technology can streamline the e-disclosure process, says **Martin Bonney**

As lawyers know, in a typical disclosure exercise less than 20% of the documents reviewed are responsive to the case. The standard approach has been to apply a few basic search terms and then review everything in a linear fashion, slowly weeding the pertinent material from the largely unresponsive whole. The enormous growth of electronic business documentation, especially e-mail, has made this approach increasingly untenable, on the basis of cost alone. Accordingly, the courts have begun to stress the importance of proportionality when it comes to the scope and cost of the disclosure exercise. For lawyers, this brings its own risks—where should the line be drawn? How can lawyers demonstrate that their approach in limiting a review is defensible and has not unduly disadvantaged the other side?

## From evolution...

It is here that technology has an increasingly important role to play. Recent years have seen the introduction of a number of new technologies to deal with the growing volume of documents. Three of the most useful have been email threading, near duplication and clustering. E-mail threading is the process where groups e-mail conversations together so that only the latest, most complete member of a thread needs to be reviewed. This can significantly speed up review times, and has the advantage of allowing a lawyer to see each e-mail in context. Near duplication allows similar documents to be compared one with another, even scanned versions against the original. This is especially helpful when looking at variations on a draft. Clustering allows you to bring order to unstructured data, grouping like documents together. With further analysis it is possible to establish patterns in the data and quickly exclude unresponsive documents at the outset of the case.

Further tools are available for specific types of case. For instance, in international cases language identification is a process which identifies a document's predominant language and groups it accordingly. This greatly streamlines the workflow, allowing documents to be apportioned to the reviewers who can best handle those languages. Alternatively, in some cases a detailed analysis of a document's provenance may be necessary. Forensic techniques can find hidden links between documents and uncover deleted or older versions, or even evident discrepancies in dates. It is always interesting when a document has been printed before it was created, for instance.

## ...to revolution

The development of "prioritisation" technology adds perhaps the most powerful new method of getting directly to the responsive documents. Prioritisation goes way beyond keyword searching by using statistical algorithms to identify the likely responsiveness of a document. Using a small sample set, a senior expert in the matter "trains" software at the outset of a

review to recognise patterns in responsive and non-responsive documents within the greater collection. The software then uses statistical sampling algorithms to rate for responsiveness every document from 0% to 100%, before documents are sent to the wider review team.

No system is infallible, which is why these systems "rate" documents for their potential responsiveness rather than making a definitive judgement on their content. By reporting on the responsive terms, such software can even suggest directions in which an investigation should head. Obviously time and cost savings are substantial.

## Benefits

For lawyers and their clients, these new search technologies have two key benefits. First, they force lawyers to define the parameters of a document search at the beginning of the process and provide an audit trail of compliance. Being able to define and demonstrate the parameters of a search in this way is increasingly important as seen in *Digicell vs Cable & Wireless* [2008] All ER (D) 226 (Oct). The court ordered one party to revisit its disclosure exercise because of a failure to agree in advance the search methodology with the other side. Second, applying these in conjunction with workflow management processes and applications enables a significant reduction in the time spent identifying key documents. This is important because the courts are paying much closer attention to the cost of the discovery process. The preliminary report of the review of litigation costs being undertaken by Jackson LJ identified the disclosure process as one of the major contributors to the cost of civil litigation and makes clear that this is one of the key issues that Jackson LJ intends to address when his final recommendations are published at the end of the year. It is likely that financial penalties will be imposed on parties that fail to keep the cost of a disclosure exercise in proportion, and the scrutiny of the process post-Digicell means that lawyers will have

## “The courts are paying much closer attention to the cost of the discovery process”

to pay great attention to defending their methodology.

While such technologies cannot replace the human element in a document review, they will become an increasingly important part of it. In future, running disclosure exercises in a low-cost, high-scrutiny environment will require a combination of cutting-edge technology, efficient workflow, and lawyers who are willing and able to marry their legal expertise to this new approach.

NLJ

**Martin Bonney**, international consulting director at Epiq Systems.  
Website: [www.epiqsystems.co.uk](http://www.epiqsystems.co.uk)