“Where Does a Wise Man Hide a Leaf?”¹:
Modernising the Laws of Disclosure in the
Information Age

Denise H. Wong
Assistant Professor of Law
Singapore Management University
denisewonghw@smu.edu.sg

Abstract. Litigation practice has been significantly altered by the advent of electronically stored information in daily corporate life. It is argued that the laws of disclosure should be updated to recognise that technology-assisted document review via keyword searching is crucial in ensuring that the costs of litigation are well managed. In order to facilitate keyword searching, a new legal concept of accuracy in the selection of keywords should be introduced into the laws of disclosure. At the same time, despite the adversarial nature of litigation, it is imperative that parties approach electronic disclosure with a spirit of collaboration in order to achieve collective savings of time and resources.

Keywords. Litigation, Disclosure, Keyword Searches, Electronic Documents, Technology

1. Introduction

“The interests of efficiency require that a case gets to trial as soon as possible with the best set of documents that can be amassed to assist in arriving at a decision on the merits… Efficiency seeks to cull the volume of documents to be disclosed and it employs the scythe of proportionality and economy… The Holy Grail is to arrive at a set of documents of the right size containing all relevant documents without expenditure of disproportionate costs.”

Global Yellow Pages Limited v Promedia
Directories Private Limited [2013] 3 SLR 758

A revolution has taken place in the world of litigation. The advent of the Internet and email technology have heralded a sea change in commercial habits, particularly in the manner in which people communicate and in the volume of electronic documents being produced and stored as part of daily corporate life. The mass of information that has been accumulated causes significant difficulties when the data has to be retrieved, sorted and reviewed as and when a dispute arises. This problem is exacerbated by globalization and the cross-storage of corporate information across jurisdictions. This

¹ Nichia Corporation v Argos Ltd [2007] EWCA Civ 741 (CA) (Jacob LJ).
phenomenon has necessitated a shift in the way that lawyers conduct and advise their clients on protocols for the preservation and retrieval of documentation in anticipation of or for the purpose of litigation.

This article examines the existing laws of civil procedure in the United Kingdom, particularly those relating to the pre-trial disclosure of documents, and seeks to determine whether those rules adequately cater for the modern day realities relating to the review of electronically stored information. The laws governing disclosure in most Commonwealth countries are similar in that they require all relevant documents and information to be disclosed to opposing parties at an early stage of proceedings. Yet, due to the large quantities of electronically stored information, it has become increasingly difficult for parties to litigation to manage the costs of document review while still complying with the requirements of the law. This electronic dimension has therefore added a layer of complexity to the fulfillment of the overriding objective enshrined in the Civil Procedure Rules, which is to “deal with cases justly and at proportionate cost”.

Just as technology has created the problem, technology also offers some solutions. This article explores some of the tools that parties may utilise in order to reduce the costs of document review that compliance with their disclosure obligations requires. In particular, the use of automated search technology, such as keyword searches and concept searches, may prove to be an efficacious and economical resource to aid such disclosure. In order to facilitate the development of automation as an effective tool in the process of disclosure, the law needs to keep up to date by prescribing new legal concepts and methodology, specifically tailored to such technological innovation. Using keyword searches as an illustration, it will be argued that the law of disclosure in the context of electronic documents should make a fundamental shift away from manual ocular review and endorse with greater confidence the disclosure of electronic documents primarily through the process of keyword searching. In this regard, the Singapore High Court in Global Yellow Pages Limited v Promedia Directories Private Limited [2013] 3 SLR 758 (“Global Yellow Pages”) has offered invaluable judicial guidance. This article will explore how some of the ideas espoused in Global Yellow Pages can be incorporated into the litigation landscape in the United Kingdom.

In addition to new legal tests, the mindset and approach of the litigating parties towards the process of disclosure must also change. Although perhaps counter-intuitive to the common law practitioner, the extensive and complex nature of information in the electronic realm demands that litigants and their lawyers adopt a more co-operative attitude and work with their adversaries in order to achieve the objectives of the disclosure process. The Court in Global Yellow Pages specifically endorsed collaboration amongst the parties to litigation as being an important means of managing the timeframe and expense of the disclosure process. It is argued in this article that such a sentiment is concordant with the principles of proportionality and agreement that are set out in the Civil Procedure Rules and the attendant Practice Directions. The concept of increased collaboration between the parties during the electronic disclosure process should therefore be adopted in the United Kingdom.

2. The Problem – A Deluge of Information

The explosive growth of electronically stored data has been attributed to “an evolutionary burst in writing technology”. The ease with which electronic information can be generated, together with the advancement of digital storage media, have made it far too easy for people to simultaneously generate and store multiple copies of documents across a variety of media, located across jurisdictions. Instant authorship of documents and information, in this day and age, is the norm rather than the exception.

---

2 The terms “documents” and “information” are used widely in this context to refer to emails, spreadsheets and other electronic documents that contain documents in the corporate context. These shall be referred to in this article as “electronically stored information”.

3 Civil Procedure Rules, Rule 1.1.


The statistics in support of these propositions are startling. It has been estimated that one hundred employees sending an average of 25 e-mails daily produce about 625,000 e-mails yearly. Other experts have estimated that about one-third of documents are solely in digital form. Elsewhere, it has been argued that “[t]he volume, number of locations, and data volatility” are significantly greater in electronic disclosure than in conventional disclosure of paper documents.

The traditional laws of disclosure, which were designed with paper documents in mind, have not coped well with this modern phenomenon. While the nuances of these laws differ slightly from jurisdiction to jurisdiction, the process of disclosure is generally the means by which one party obtains documents and other information relevant to the case from another party in advance of trial. The historical origins of disclosure can be traced to the Courts of Chancery and seem to date as far back as Henry VI in the 15th century. The aim of the disclosure process was to increase the likelihood of a fair trial by increasing the likelihood of the court resolving the dispute on the basis of facts which represent or which are reasonably proximate to the truth. The processes mandated by law are specifically designed to ensure that parties are able to successfully extract all relevant documents and other information from others and to thus find evidence supporting their own case and undermining their opponent’s. Other purposes of the disclosure process are to obtain the best evidence possible as well as to allow parties to evaluate the relative strengths and weaknesses of their case as well as that of their opponent’s. The removal of the element of surprise leads, in theory at least, to efficiency and savings of costs because parties may choose not to proceed with an action once they have a more complete understanding of the evidence against them.

While these ideals may well have been successfully achieved in respect of the disclosure of paper documents, it has been judicially observed that the traditional manner in which disclosure has been carried out is proving increasingly inefficient in achieving the purposes for which the disclosure process was developed. The problem has three facets. First, the sheer volume of electronic information means that the cost of review has increased exponentially as compared to the past. The time and labour costs of hiring junior associates and paralegals to look through documents become difficult to manage and can be disproportionate to the actual quantum of the claim. Second, the problem is exacerbated when disclosure, which is intended to be an equalizing force to prevent injustice, is used by the party with the deeper pockets as a tool to drag out litigation and confuse the opponent. Third, the quantity of information involved makes it difficult for the party giving disclosure to know what documents he has in his cupboards that he must disclose in order to fulfill his legal obligations.

---


The malaise described in the first and second facets was succinctly captured by Jacob LJ in *Nichia Corporation v Argos Ltd* [2007] EWCA Civ 741 in the following terms—

“46 ... It is wrong just to disclose a mass of background documents which do not really take the case one way or another. ...

47 Now it might be suggested that it is cheaper to make this sort of mass disclosure than to consider the documents with some care to decide whether they should be disclosed. And at that stage it might be cheaper—just run it all through the photocopier or CD maker—especially since doing so is an allowable cost. But that is not the point. For it is the downstream costs caused by over-disclosure which so often are so substantial and so pointless. It can even be said, in cases of massive over-disclosure, that there is a real risk that the really important documents will get overlooked—where does a wise man hide a leaf?”

As for the third facet, as highlighted by Senior Master Whitaker in *Gavin Goodale & Ors v The Ministry of Justice & Ors*, the problem in respect of the disclosure of electronic information is often for a party to gauge the scope of a reasonable search as required under Rule 31.7 of the Civil Procedure Rules of the United Kingdom and the accompanying Practice Direction 31B. The standard for what can and should be considered a reasonable search of paper documents cannot be easily translated to electronically stored information. This is acknowledged by Practice Direction 31B, which was specifically promulgated to “encourage and assist the parties to reach agreement in relation to the disclosure of Electronic Documents in a proportionate and cost-effective manner”. Paragraph 20 of Practice Direction 31B recognises that “[t]he extent of the reasonable search required by rule 31.7 for the purposes of standard disclosure is affected by the existence of Electronic Documents”. Rule 21 goes on to set out various factors that may affect the extent of the reasonable search, which include the accessibility of electronically stored information and the cost of recovering and disclosing such information.

It can thus be seen that the volume and nature of electronically stored information has altered the litigation landscape, particularly with regard to the manner in which parties fulfill their disclosure obligations in civil suits. While the law in the United Kingdom has adapted to specifically address this issue, uncertainty still exists as to the scope and extent of disclosure required as this is premised upon the amorphous concept of reasonableness, which becomes even more nebulous when applied to electronically stored information. Central to the equation is the idea of managing the costs of litigation to ensure that justice is economical, proportionate, expeditious and equally available to all parties to the action.

3. The Solution – Where Technology Meets Collaboration

3.1 The pros and cons of keyword searching

Keyword searching, a technique well familiar to the legal fraternity, has been defined as a software-aided search for words across the text of an electronic document. It is a fairly obvious proposition that such searches can be critical in aiding the disclosure process by quickly and effectively trawling through a mass of electronic documents to pinpoint potentially relevant ones. Less obviously, keywords searches are extremely significant in the disclosure process because parties may be allowed to discharge their legal obligations of disclosure by only turning over all documents that are identified by the keyword search. The keyword search thus obviates the need to conduct full ocular review. This is certainly a nod to the overarching concept of proportionate and economical justice, but it must be...
recognised that it does detract from the ‘no stone unturned’ conception of disclosure, with which lawyers are more familiar, and perhaps more comfortable.

Practice Direction 31B makes a concerted effort to take a nuanced position on the utility of keyword searches in the disclosure of electronically stored information –

“25 It may be reasonable to search for Electronic Documents by means of Keyword Searches or other automated methods of searching if a full review of each and every document would be unreasonable.

26 However, it will often be insufficient to use simple Keyword Searches or other automated methods of searching alone. The injudicious use of Keyword Searches and other automated search techniques –

(1) may result in failure to find important documents which ought to be disclosed, and/or
(2) may find excessive quantities of irrelevant documents, which if disclosed would place an excessive burden in time and cost on the party to whom disclosure is given.

27 The parties should consider supplementing Keyword Searches and other automated searches with additional techniques such as individually reviewing certain documents or categories of documents (for example important documents generated by key personnel) and taking such other steps as may be required in order to justify the selection to the court.”

The approach taken by Practice Direction 31B is finely tuned to recognize both the advantages and the potential dangers that are associated with keyword searches. It is laudable that the law recognizes that one can conduct a reasonable search (in the legal sense of the term) through the use of keyword searching. As highlighted above, this is in line with the objective of achieving justice in a proportionate and economical manner.

At the same time, the Practice Direction sounds a cautionary note as to over-relying on such searches to produce all the documents that are relevant to the suit. Interestingly, it states that the “injudicious use” of keyword searches can lead to over- or under-inclusion of documents in the disclosure process, but does not go on to explain what would constitute “injudicious use”. The Practice Direction then, unfortunately, takes a step backward by essentially recommending manual ocular review as a safety net to supplement keyword searching. Given that the aim of keyword searching is precisely to save the costs incurred by manual review, this recommendation is not only counter-intuitive, but also a potential avenue for abuse by errant parties (and their lawyers) who wish to drag out proceedings. The Practice Direction almost seems to conceive of disclosure as occurring in two stages – first via a keyword search, and then at a second stage, via a more targeted manual review. While this approach is certainly more comprehensive, it is also time-consuming and arguably, unnecessarily cautious.

Just as technological advancements have fundamentally altered the form and nature of documentation, so too must the law adapt and develop the corresponding principles that are tailored to the electronic realm. Specifically, in order for technology to truly deliver a solution to the problems created by the sheer magnitude of potentially discoverable electronic documents, there should be a decided move away from manual review as the fall-back option. A targeted legal framework should be put in place to facilitate the use of keyword searches as an effective mechanism for the review and identification of documents to be disclosed. The Singapore High Court took precisely such an approach to keyword searching in Global Yellow Pages. There, the Court firstly endorsed the use of technology, especially automated computer search tools, to cope with the burgeoning volume of discoverable documents. In particular, the use of keyword searches with Boolean operators was highlighted as a familiar and easily accessible tool that can aid in the identification and retrieval of relevant documents. However, the Court recognised that not all searches would result in retrieving relevant documents, since keywords could have different meanings depending on how the term is used. The main shortcoming of such searches was therefore that they were potentially over- or under-inclusive, with false positives and false negatives an inevitable consequence. As can be seen, the

In order to address this problem, the Court in *Global Yellow Pages* took the view that the selection of appropriate keywords was critical to the success of such searches being useful to the disclosure process.\(^{23}\) Significantly, the Court held in this regard that since any keyword or sets of keywords will produce false positives and false negatives, the traditional concept of relevance as understood in classical evidence law did not apply directly to the issue of which keywords should be used in searches. Keywords should instead be identified by reference to their *accuracy*. An accurate keyword was one that has a low proportion of false negatives and false positives to the total number of documents within the set of documents that has been identified using that keyword.\(^{24}\) This newly created legal concept of accuracy should henceforth be the touchstone of any collaboration between the parties in respect of keyword selection.\(^{25}\) Importantly, the Court should also adjudicate upon disputes between parties as to whether certain keywords should be used in the disclosure process by reference to their accuracy, rather than their relevance *per se*.

In the Court’s view, the concept of accuracy had two aspects.\(^{26}\) The inclusionary aspect related to the correlation between the keywords and the issues in dispute. The greater the correlation, the more likely it would be that the keywords would result in the relevant documents being included. Conversely, the exclusionary aspect dealt with the exclusion of irrelevant documents from the documents to be disclosed. Overall, the higher the accuracy of the keyword, the more likely it would be that the Court would grant an order that those keywords be used as part of the disclosure process.\(^{27}\)

### 3.2 Preliminary searches/data sampling

Having expounded upon accuracy as a novel legal concept, the Court in *Global Yellow Pages* went on to make clear that accuracy would not be the only factor which the Court would take into consideration when deciding whether to make an order relating to a particular keyword. The other factor was the size of the subset that was produced by a search using that keyword (i.e. the number of “hits”). A large number of hits would, in general, be a factor weighing against the inclusion of the keyword as a search term in the disclosure process.\(^{28}\)

The next logical question to be answered was how the parties and the Court should estimate the size of the subset that would be produced through the use of a particular keyword. In this regard, the Court resoundingly endorsed the use of preliminary searches and data sampling in order for the parties (and the Court) to get a better sense of the likely results from searches using keywords. Preliminary searches would be conducted by the parties against the universal set of documents for the purpose of determining the number of “hits” obtained by the relevant search term. The party running the search would not be permitted at the preliminary search stage to review the documents that had been revealed by the search. Such an approach has already received judicial endorsement in the United Kingdom in *Gavin Goodale & Ors v The Ministry of Justice & Ors.*\(^{29}\) Data sampling, on the other hand, involved the parties identifying the keyword search terms and the repositories that they wished to sample. Upon the search being run, parties would view the documents that had been revealed by the search in order to determine how useful the search term was. Hence, parties conducting data sampling would actually have sight of the documents revealed by the search, but only of a relatively small subset of the universe of potentially relevant documents.\(^{30}\)

---

\(^{23}\) *Global Yellow Pages Limited v Promedia Directories Private Limited* [2013] 3 SLR 758.

\(^{24}\) *Global Yellow Pages Limited v Promedia Directories Private Limited* [2013] 3 SLR 758 at para 53.

\(^{25}\) *Global Yellow Pages Limited v Promedia Directories Private Limited* [2013] 3 SLR 758 at para 54.

\(^{26}\) *Global Yellow Pages Limited v Promedia Directories Private Limited* [2013] 3 SLR 758 at para 54.

\(^{27}\) *Global Yellow Pages Limited v Promedia Directories Private Limited* [2013] 3 SLR 758 at para 54.

\(^{28}\) *Global Yellow Pages Limited v Promedia Directories Private Limited* [2013] 3 SLR 758 at para 57.


\(^{30}\) In Practice Direction 31, ‘Data Sampling’ means the process of checking data by identifying and checking representative individual documents.
3.3 Selection of keywords
Prior to Global Yellow Pages, another Singapore Court had set out an instructive guide on how to go about selecting keywords. In the first instance decision of Breezeway Overseas Ltd v UBS AG\(^31\), the Court set out as follows –

“First, commence with the specific before expansion to broader search terms. Specific search terms would include the following:

(a) Unique reference numbers. For example, bank account numbers or client account numbers where the context is in a banking relationship. In the context of other commercial transactions, if one party has in place a file reference number or account identification number, these may be used as well. This very closely approximates the traditional paper filing system.

(b) Names of specific projects. This can be an important keyword particularly where the dispute arises from a developmental project or commercial transaction which has been assigned a project name …

(c) Keywords which identify the key witnesses (or custodians). For example, e-mail addresses, contact numbers and names or initials. Search terms may be formulated based on such keywords. For example, e-mail addresses of two key witnesses appearing in the same e-mail in order to identify e-mail conversations between them…

(d) Significant events and locations. Depending on the facts, there may have been a significant meeting which took place. It has proven useful in some cases to make use of the meeting location or a short-hand reference to a key meeting as a search term. This may identify correspondence and documents that have been generated surrounding the event. However, care must be taken in selection of locations….”

It is suggested that if similarly clear judicial guidance is provided in the United Kingdom to set out the parameters of acceptable keyword searches, it would not be necessary for the Practice Direction to explicitly endorse manual ocular review as a safety net to supplement keyword searching. Perhaps more importantly however, the time has come for a paradigmatic shift towards the notion that even if disclosure via keyword searching is less comprehensive than manual review,\(^32\) the gains in efficiency and economy far outweigh the dangers of under- or over-inclusive document exchange. Such a recalibration would be entirely in line with the overriding objective of the Civil Procedure Rules.

3.4 Collaboration and good faith negotiations
It is clear that technology as a solution would be ineffective unless parties are willing to work together in a collaborative setting in order to identify and sieve through the relevant keywords. To this end, the Court in Global Yellow Pages highlighted that parties should work together in an iterative process whereby parties propose keywords to their adversaries and then endeavor to reach consensus on the terms to be used. Parties should negotiate and resolve any disagreements without reference to external adjudication. Such collaboration would result in “virtuous cycles” and “iterative feedback loops” whereby meet and confer sessions between the parties occur at regular intervals for information exchange and discussion.\(^33\)

While seemingly at odds with the adversarial nature of litigation, the Court highlighted as follows –

“It would be desirable for parties to propose keywords and to cooperate to arrive at an acceptable compromise by way of an agreed list of keywords. Such cooperation would be desirable because it reduces the cost of litigation for clients. Paul & Baron have suggested that it may in fact be in both parties’ interest for such collaboration and cooperation to occur:

---


\(^{32}\) Indeed, it has been argued that there is evidence that the use of technology-assisted processes in electronic disclosure is not only more efficient, it can also yield results superior to those of exhaustive manual review, see Maura R. Grossman and Gordon V. Cormack, ‘Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review’ (2011) XVII Richmond Journal of Law and Technology 11, <http://jolt.richmond.edu/v17i3/article11.pdf> accessed 24 October 2013.

\(^{33}\) Global Yellow Pages Limited v Promedia Directories Private Limited [2013] 3 SLR 758 at paras 44 – 45.
Quite simply, as courts and commentators have increasingly come to expressly recognize, the volume and complexity of electronically stored information demand new forms of collaboration. In turn, in many such instances, a tipping point can be said to have been reached where the game theoretical aspects of litigation practice, dictating what is in one’s self-interest, have necessarily changed. Without greater cooperation among adversaries, parties are doomed to any number of defeating consequences, not the least of which will be a real or perceived information “gap” in ferreting out evidence.”

Without the positive cooperation and regular collaboration that the Court in *Global Yellow Pages* has encouraged, it would be all too easy for electronic disclosure to be used as a means to stall or delay the litigation process. If parties were to quibble over the accuracy of each and every proposed keyword, the electronic disclosure process would become protracted, cumbersome and prohibitively expensive, especially if parties were to take out interlocutory applications for the Court to adjudicate upon issues of relevance or accuracy. The very electronic tools that were designed to facilitate document identification and exchange can easily be used as instruments of delay and oppression in the hands of an errant party with deep pockets.

As such, it is imperative that parties approach electronic disclosure with the correct mindset, bearing in mind the overriding objective of proportionate and economic justice. Lawyers must thus encourage their clients to overcome the combative mindset that is inherent in the adversarial nature of litigation and focus instead on collaborating with the opposing party, at least during the stage of disclosure, in order to achieve collective savings in terms of time and costs. As has already been highlighted, the general sentiment of proportionate and cost-effective justice has been well entrenched in the litigation landscape of the United Kingdom since the Woolf reforms and the advent of the Civil Procedure Rules. Specifically in the context of electronic disclosure, Practice Direction 31B features language that encourages agreement and cooperation between the parties. The positive collaboration that the Court in *Global Yellow Pages* envisages in relation to keyword searches would therefore not be out of place in the United Kingdom.

4. Conclusion

It is no longer sufficient in this day and age to rely on traditional principles of disclosure to contend with the unique legal issues that surround electronically stored information. It has been argued that, in the litigation context, it is necessary for a coherent framework to be in place wherein technology can assist to discharge the parties’ legal obligations relating to disclosure. In this regard, it has been argued that keyword searches are a valuable tool that parties have at their disposal to minimize the cost of document review. At the same time, the technology itself can be misused in order to delay the pre-trial process and increase the costs of litigation if parties do not approach electronic disclosure with the mindset of cooperation and collaboration. At the end of the day, it is the parties who must be invested in achieving the overriding objective of achieving justice in a fair, proportionate and cost-effective manner.

* * *

© 2014 Denise H. Wong. This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works.

35 Practice Direction 31B at para 2.