DOCUMENT MANAGEMENT SYSTEMS & PUBLIC INQUIRIES

Nicholas C. Walker, eDisclosure Project Manager

A white paper exploring how the use of eDiscovery document management systems in Public Inquiries can save the taxpayer money by providing a fast and efficient way of reviewing and disclosing documentary evidence for use in the hearings and beyond

"An inquiry should have sufficient funding to ensure that its conclusions are robust, and that its recommendations are valuable and based on an in-depth analysis of the issues involved. However, it should always be borne in mind that inquiries are funded ultimately by the taxpayer. There will be reasonable limits to how long an inquiry should last and how much the process should cost"[1]

It is true to say that whilst the public appetite for investigating instances of cover-ups, inexcusable behaviour, and scandalous inefficiencies of our public institutions and persons in positions of power is cyclopean, there is a general reluctance to want to release significant funds towards their discovery. A survey commissioned by the Centre for Effective Dispute Resolution revealed almost 60% of Britons polled believed that Public Inquiries are too costly[2]. There is, however, a need for public inquiries to fulfil a role in restoring public confidence. The Home Office states, "they are an appropriate and useful device for dealing with matters that have caused public concern[3]; "The tradition of the public inquiry has become a pivotal part of public life, and a major instrument of accountability in the United Kingdom"[4].

It is not however for this White Paper to concern itself with the merits of public inquiries and whether they benefit the public at large, but rather to consider how the use of technology can help to reduce the cost and time spent during review and additionally facilitate the disclosure process to core participants and, eventually, the wider public. In recent years there have been outcries over various public inquiries where the costs have been far higher than anticipated. A prime example is the Al-Sweady Public Inquiry[5], which cost the British taxpayer in the region of £25 million pounds[6]. For the British taxpayer this is a very large sum of money for investigating an incident that took place in a war zone thousands of miles away. The public understandably questions the overall costs and whether they are getting value for money. The Al-Sweady Public Inquiry was complex, requiring statements and interviews to be taken from around the world. By taking advantage of an eDiscovery platform in order to improve collaboration amongst their team and speed up their review, the overall cost saving was substantial compared to running the Inquiry on a traditional paper basis.

Firstly, it is important to consider where the costs are generated, and then consider the various eDiscovery tools that help to reduce these costs. The Inquiries Act 2005 was created in reaction to mounting public sector costs and law reforms such as Jackson[7] continue to be implemented in order to press upon the legal community the need to focus on controlling legal costs. Undoubtedly the largest cost for an Inquiry is the legal fees for the paralegals, solicitors and counsel in relation to the gathering and review of evidence. Legal costs are unavoidable – the review the documents will always be an essential component and preferably those people undertaking

---

1   11.24, Memorandum by the Department for Constitutional Affairs (GBI 09), Parliamentary Publication July 2004, http://www.publications.parliament.uk/pa/cm200304/cmselect/cmpubadm/606/4052502.htm
2   Opinium Survey, May 2012, commissioned by CEDR, http://www.cedr.com/news/?item=It-s-time-for-reform-of-the-UK-Public-Inquiries-System
3   13.1, Memorandum by the Department for Constitutional Affairs (GBI 09), Parliamentary Publication July 2004, http://www.publications.parliament.uk/pa/cm200304/cmselect/cmpubadm/606/4052502.htm
4   Public Inquiries, Jason Beer QC, OUP Oxford, July 2011, ISBN0199287775
5   http://www.itv.com/news/update/2014-12-17/anger-over-cost-of-31-millon-al-sweady-inquiry/
6   http://webarchive.nationalarchives.gov.uk/20150115114702/http://www.alsweadyinquiry.org/costs/index.htm
7   https://www.judiciary.gov.uk/wp-content/uploads/2014/05/impact-of-the-jackson-reforms.pdf

the review should have the requisite training in dealing with questions of relevancy and privilege. Sir Menzies Campbell has been quoted as saying "[lawyers]… have the skills which can ensure an inquiry is pursued expeditiously and comes to conclusions which are understandable and capable of being implemented"[8]. Most lawyers' costs are generated through billable time, and so the use of an eDiscovery document management system is important in helping to reduce the time that fee earners spend reviewing material by making the documents more accessible. Historically, review work was undertaken with the hard copy document in front of you, highlighter pen at the ready, and a mountain of coloured reference tabs in the top desk drawer. With the use of online eDiscovery platforms that can quickly search, display the most relevant documents and highlight key words automatically, a lot of that preparation time can be reduced. For example, in recent Inquiries the establishment of keywords early in the process has enabled the legal review team to prioritise documents for review and disclosure without prior need to read documents for content to establish a core group. The use of standard litigation tools such as keyword searches, date culling, persistent highlights and search term reports aid the Inquiry team to determine which documents are potentially relevant.

Online eDiscovery document management systems also allow for greater collaboration between legal teams, regardless of where they may be based. With dynamic coding tags, automated batching of documents and reviewer statistical analysis scripts it is possible to report in real-time the pace of the review and which areas of review may require more or less resources. This allows Inquiry team managers to allocate certain resources appropriately, reducing time and cost. The effort to collate this information without eDiscovery tools is staggering, and arguably could not be achieved with the same accuracy if undertaken manually.

From a document review perspective, historically it has been a pre-requisite for an Inquiry that almost every document needs to be preliminarily reviewed to determine relevance. Typically this is because the documents will come from various different sources and the Inquiry team will have little to no knowledge of the contents of those documents prior to arrival. Ideally the Inquiry team would require the providing organisations to create documentation logs of all material before delivering it to the Inquiry team, although this is not always possible and the resulting data, if not objectively coded consistently and precisely, can be misleading. Regardless of how painful a task this would appear to be, the value of this work cannot be underestimated. Objective coding becomes an incredibly important tool, especially when dealing with hard copy documents, as this metadata is vital in order to efficiently and speedily understand the potential importance of a document. Using filters and pivot charts on this data helps to visually present the information clearly. Qualitative and quantitative analysis of this metadata information through reporting mechanisms within the eDiscovery platform enables Inquiry teams to prioritise their review, and ultimately allows all parties to the Inquiry the ability to quickly locate documents of significant value.

A key challenge that concerns historic Inquiries relates to the amount of material received in hard copy format. There is a much greater expense scanning documents than there is processing documents electronically. The quality and quantity of information that can be extracted from hard copy documents also presents itself as an issue - in order to be able to search within an eDiscovery platform, hard copy documents must have an optical character recognition (OCR) program run over them in order to recognise and extract textual content. Whilst OCR is a powerful and cost effective method compared to manually transcribing documents, limitations exist relating to the accuracy of the process. OCR success rates are dependent on the quality of the scanned material, ranging greatly between 60% and 98%, and the presence of handwriting may not yield any results[9]. Consideration should be given to the quality of the hard copy material before processing and it would be advisable in some instances to look at the cost of objectively coding these documents, using this in tandem with the OCR. Electronic documents in comparison do not have OCR issues as their text is extractable through computer processing (unless the electronic file is a photo or a flat image).

Compared to some wide ranging regulatory investigations or complex civil commercial litigation, historically Inquiries have not generally been provided with large volumes of documents. This has led their review to be more linear and reactionary. The relatively smaller volumes of documentation are due in part to a historic lack of written records and inadequate procedures to preserve data

8   Big Question: are public inquiries a waste of time?, Prospect Team, Prospect Magazine, 23 January 2015; http://www.prospectmagazine.co.uk/blogs/prospector-blog/big-question-chilcot-inquiry-child-sex-abuse-tony-blair-theresa-may-are-public-inquiries-a-waste-of-time

9   http://www.ipcgroup.co.uk/information-management-blog/items/how-well-does-ocr-work-with-scanned-documents.html

DTI

being observed. Whilst document volumes and the requirements for evidence are dependent on the type and era of the event being investigated, the importance of having a secure, collaborative platform for review is paramount. Whilst Inquiries all vary in amounts of documentation, whether an Inquiry has 1,000 documents or 100,000 documents, the benefits of using an eDiscovery platform are equally applicable. An example of a mid-range sized Inquiry is the Leveson Inquiry, which lasted just over a year, and "took three months to start receiving the evidence (although the process continued for a year)"[10]. The Inquiry team reviewed approximately 10,000 documents (over 110,000 pages) within this time. Therefore having five to six solicitors on staff meant a structured linear review was understandably acceptable. Conversely the Mid Staffordshire Public Inquiry reviewed over 60,000 documents with more than 1,600,000 pages. Sir Robert Frances QC set the aggressive deadline of a year to conduct this Inquiry, and they were required to draft in more paralegals to review the documents in order to complete in time. Without the eDiscovery tool to provide the benefits of collaboration and speed, the additional legal resources required would not have integrated so easily, however could the eDiscovery platform have gone further by offering a tool that would have prevented the cost of the additional legal resources being incurred?

It is inevitable that with mobile technology and social media being more widely used, much larger amounts of data concerning the incidents in the public interest will be available for collection. The 1989 Hillsborough tragedy has been described as the most well documented events in history, and this was before email and social media websites were even available. With the onset of Twitter, Facebook and online real-time news outlets it is not unreasonable to imagine that the document count, if the event had taken place today, could be several times that of its current numbers. It is impossible to imagine a paper based review of this material being time and cost effective.

To help deal with the challenges presented by the increase of evidence from modern technology, eDiscovery platforms use new industry applications designed to reduce the number of documents relevant to a review. One of the most talked about tools is Technology Assisted Review, where an eDiscovery platform "electronically classifies documents based on input from expert reviewers, in an effort to expedite the organization and prioritization of the document collection."[11] Not only designed to reduce the amount of human review time, it saves costs through automating decisions of relevancy based on logical fact interpretations of a document's content. It requires an initial small review input from the Inquiry team, as a way for the platform to better understand which documents are responsive, and then regular review of those results. Through the rounds of the assisted review, the platform increases its understanding of which documents are most relevant and helps to reprioritise responsive documentation. DTI currently have an Inquiry team engaging with this technology to assist them with their document review. To demonstrate the effectiveness of this technology and using this Inquiry team as an example, they started their review with over 2,000,000 documents. After an initial review of 2,500 documents to help train the platform, it produced an initial set of decision based coding for the entire data set – 35% categorised as responsive and 27% categorised as non-responsive. There were

Diagram 1: Computer Assisted Review – Training & 3 QC rounds of document review

---
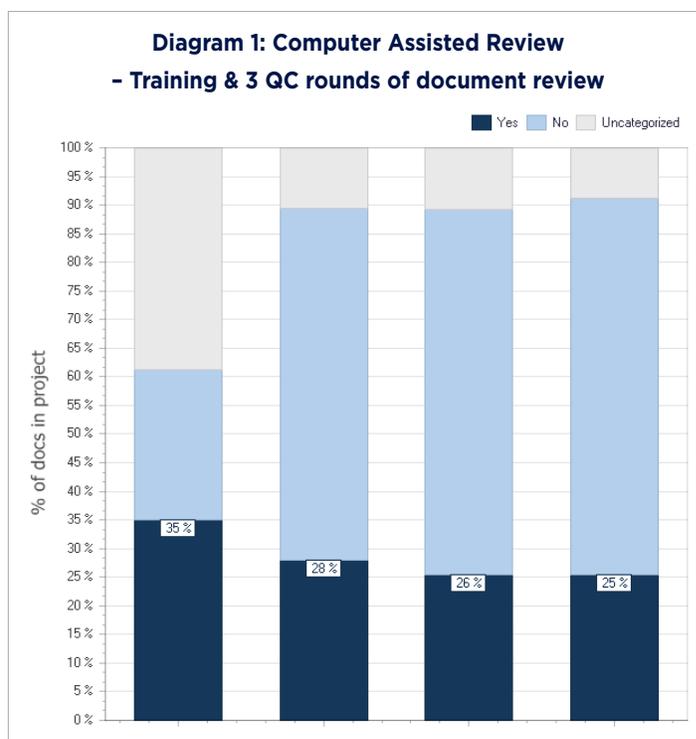
10   https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/229039/0779.pdf, page 4
11   http://www.edrm.net/resources/carrm

World-Class Service. Local Commitment.
Visit DTIglobal.com to learn more.

a proportion of documents (38%) that were uncategorised from the training set. This is due to the platform being unable to determine their classification based on the review team's decisions. Subsequent rounds of human review were then required to teach the platform which documents from that uncategorised set were responsive, thereby increasing the platform's knowledge basis and ability to identify responsive categorisation. After four rounds of human review the system significantly reduced the number of uncategorised documents to less than 10%. Diagram 1 visually explains the evolution of the responsively coded documents through those sets; the first column shows the platform's initial response to running the assisted review algorithm on the data set, the second column shows the changes after the first quality checking rounds. The third and fourth columns show the further rounds decreasing both relevant and uncategorised material as the system sharpens its knowledge of the types of document that the Inquiry team deem to be relevant.

The end result is a vast saving of cost and time in helping the Inquiry team to access relevant information with the wider document collection and prepare for disclosure. With the recall rate and uncategorised portions of the data set reduced further and further, for the small task of re-reviewing a stratified sample set of 1,500 per round to help re-teach the platform, the Inquiry team are potentially saving almost 35,000 hours of reviewer time by using Technology Assisted Review[12]. The saving for the tax payer is quite considerable – in the case above potentially £385,598[13].

Faced with a gargantuan amount of documentation, and extensive terms of reference, it is entirely plausible that Inquiries will not know their keywords, or even which documents are most important in the context of the review. Inquiries may struggle to engage with the documentation if they are unaware of what identifies relevancy and will spend more time, and ultimately more cost, searching for documents in the hope that they match a criteria of interest. When discussing Technology Assisted Review, the process is reliant on the Inquiry team being able to identify relevant documents from the outset, however there is an analytic tool that can effectively help to identify documents of most relevance without prior knowledge of the keywords and issues – cluster analysis.

Cluster analysis, or Clustering, is a process by which the eDiscovery tool runs an algorithm across the documents and groups them in such a way that documents in the same group are more similar to each other not just in content, but conceptually. This tool is incredibly useful because it does not require the Inquiry team to provide much input, and does not require any example documentation in order to find matches, "Depending on the conceptual similarity, the [system] identifies the most logical groupings of documents and places them into clusters."[14] For example, running cluster analysis over the documents has identified sausage, bacon and eggs as a group of documents where these words frequently appear in the content. However, it has also brought in documents relating to orange juice, beans and toast, because the conceptual algorithm identified and understands that the grouping refers to breakfast items. This means that documents that would not have ordinarily been captured by searching mechanisms have been identified as potentially relevant. In a recent case DTI have been engaged with, the cluster analysis tool was able to identify that an event that occurred on the same evening as the main incident. The system was therefore able to recall documentation which provided more information about the main incident through conceptual likenesses in content, which would otherwise have been missed. Used in conjunction with Technology Assisted Review, an Inquiry team can identify the key concepts of their investigation and prioritise the most relevant documents for review, saving vast amounts of time and cost.

Undoubtedly another main cost factor for Inquiries is their obligation to provide relevant material to core participants. The Bloody Sunday Inquiry, which did not use an eDiscovery document management system, "received 2,500 witness statements and compiled some 160 volumes of evidence, 13 volumes of photographs, 121 audiotapes and 110 videotapes, all of which had to be photocopied and sent to representatives of the 'interested parties'... when the ten-volume report was published twelve years later it had cost the taxpayer £195 million"[15]. The concept of photocopying and producing endless bundles of evidence for multiple parties, with the advent of modern technology, seems incredibly inefficient. EDiscovery platforms can provide separate core participant databases that all parties can

---

12   Figures based on irrelevant documents percentage in the fourth (67%), and on-sight relevancy review of documents at 25 documents per hour
13   Total cost saving generated based on median statistics for a mid-career paralegal's hourly pay rate (current as of 31 March 2015): http://www.payscale.com/research/UK/Job=Paralegal/Salary/39f07c00/Mid-Career
14   https://help.kcura.com/8.2/Content/Relativity/Analytics/Clustering.htm
15   http://ukconstitutionallaw.org/2013/02/28/carol-harlow-what-price-inquiries/

DTI

World-Class Service. Local Commitment.
Visit DTIglobal.com to learn more.

access and review the documents online, without the need to print out the documents or have to store them in multiple locations for different teams. The online databases can be accessed anywhere, anytime, and products such as Relativity Binders® can push witness bundles to counsel's tablets to review before and during the hearing quickly, efficiently, and easily. Instead of wasting money on printing and courier costs (which can be sizable depending on geographic location and time to deliver), large numbers of documents can be transferred in a relatively short space of time. Rather than using complicated printed indices and tables of contents at the start of each bundle (which often get revised on a constant basis), the use of bookmarked PDFs and hyperlinked spreadsheets enables all parties to navigate to sections in the most efficient way. The cost and time saving on preparing bundles for each participant of the hearing is huge – in the case of *Jackson v Thompsons and others* [2015], this amounted to nearly an 85% cost saving[16].

Other additions such as unique document reference numbers also allow for speedier document recall, and facilitate a greater transparency amongst legal teams of what has been disclosed. This was important for the Hillsborough Inquests, which were borne out of the report by the Independent Panel. The panel team created a website and published all the relevant documents from its review in order for the public to be able to see the documents themselves. The website contained a unique reference number for each document which was subsequently carried through to the Inquests, so that anybody could reference a document and everyone would know what was being viewed. This transparency gave the families confidence that the openness of the Independent Panel would be reflected in the Inquests.

"Sir, the really crucial feature of the Lextranet system as I understand it is that because documents are coded in a particular way…it really is a useful searchable database, which is not only of importance between now and the inquests, but of course when we actually start the inquests, which are likely on any analysis to go on for many, many, many weeks, all of the interested persons and their representatives will have access to a system which will enable them to put in a search code and usefully turn up the relevant documents. That's its beauty…so its magic, if I can put it that way"[17]

Once an Inquiry is ready to sit and hear the evidence, the value of using an eDiscovery platform is further increased. Courtrooms across the UK embrace trial support technology and Inquiries are no different. The document management system is made available on every desk in the hearing room, rather than a pile of lever arch files which saves a great deal of space. Documents can be brought up on screen quickly without delay, saving time. Counsel enjoys the freedom they now have of having the documents on easy to use tablet devices which connect and integrate with their work computers seamlessly. Furthermore, through producing the documents electronically, they can be provided to an in-court presentation team for use with systems such as Trial Director® or OnCue®. This enables the Inquiry team and all Core Participants, using the unique reference number, to call any document almost instantaneously within the hearing room.

After an Inquiry has finished sitting and the report has been written, the final benefit of an eDiscovery document management system is the ability to archive the material accurately and safely. Most eDiscovery document management systems can export the documentation in a format that is universal for preservation, and the documents can be produced in various ways, for example with or without redactions, for the purposes of freedom of information requests. Through hashing methods, used to identify a document's unique fingerprint, documents can be transferred to archive systems such as the National Archives without danger of missing any information. Schedules of metadata relating to the documents can be provided ensuring that searching for evidence years in the future is as effortless as possible. This process in hard copy would be fraught with potential inaccuracies and potentially omitted documentation. It would also go against the digitisation direction archivists are now moving.

EDiscovery platforms are the best way to review Inquiry documentation quickly, efficiently, and in the most cost effective manner. As more and more court rooms embrace electronic technology, online databases offering collaboration from people all over the world are essential in order to minimise cost and reduce the length of hearings. The transparency of electronic data being made available to all parties helps to restore confidence in the Inquiry process. The savings made can help to fund other public institutions and in a time of

16   *Jackson v Thompsons and others* [2015] EWHC 218 (QB), as referenced at http://www.lawgazette.co.uk/analysis/comment-and-opinion/a-judicial-eye-on-bundles/5046776.fullarticle
17   Ms Christina Lambert QC, Transcript of initial hearing, Hillsborough Inquests; June 5, 2013; pages 74-75

World-Class Service. Local Commitment.
Visit DTIglobal.com to learn more.

austerity can the Government really afford not to use such a cost saving tool? Ultimately, an Inquiry that runs on time and delivers under budget through the use of eDiscovery technology can provide not only the catharsis sought by the individuals who are party to the hearings, but a financial relief for the British taxpayer.

About the Author

Nick Walker is an expert in the electronic document management of Public Inquiries and Inquests. Mr. Walker works closely with public sector departments such as the Revenue and Customs Prosecutions Office, the Home Office, the Serious Fraud Office, and the Crown Prosecution Service, having developed numerous e-Browser applications for a variety of cases. Project managing a mixture of public, civil litigation and arbitration work using document management systems, key projects Mr. Walker has managed and consulted on include the Inquest into the death of Jean Charles de Menezes, Baha Mousa Public Inquiry, Al-Sweady Public Inquiry, Potters Bar Inquest, 7 July Bombings Inquest, Vale of Leven Hospital Inquiry, Mid Staffordshire NHS Public Inquiry, Hillsborough Independent Panel Review, Leveson Inquiry, Litvinenko Inquest & Inquiry, and the Hillsborough Inquests. Mr. Walker graduated from Kent Law School and is an eDisclosure Project Manager with DTI.

About DTI

DTI is a leading legal process outsourcing (LPO) company serving law firms and corporations around the globe. DTI helps its clients accelerate the changes they must make to remain competitive. DTI is a preeminent leader in the management of information and processes. The company manages risks and minimizes costs associated with complex litigation and compliance functions. The extensive experience in eDiscovery, managed services, litigation support and court reporting is unmatched in the industry. To learn more about DTI's global footprint, flexibility, capacity and world-class project management, visit www.DTIGlobal.com.

World-Class Service. Local Commitment.
Visit DTIglobal.com to learn more.