E-disclosure in disputes

FW moderates a discussion on e-disclosure in disputes between Robert Jones and Tracey Stretton at Ankura, David Webb at DLA Piper, Ben Sigler at Stephenson Harwood and Mark Surguy at Weightmans.
THE PANELLISTS

Robert Jones  
Managing Director  
Ankura  
T: +44 (0)20 7015 8807  
E: robert.jones@ankura.com  
Robert Jones is a managing director at Ankura, based in London. He is responsible for developing and expanding the firm’s business in the UK and overseas. Over the last 19 years, he has worked on a wide range of disputes, investigations and proactive compliance initiatives and has led successful engagements with law firms, corporations, small to medium enterprises (SMEs), government agencies and individuals.

Tracey Stretton  
Managing Director  
Ankura  
T: +44 (0)20 7015 2338  
E: tracey.stretton@ankura.com  
Tracey Stretton is a managing director at Ankura with 20 years’ experience advising lawyers and their clients on the use of technology in litigation and investigations. Her experience in legal technologies is based on exposure to its use as a lawyer and consultant on a large number of complex cases in a variety of international jurisdictions. Ms Stretton is a published specialist on the impact of technology on law and business.

David Webb  
Head of Electronic Data Review  
DLA Piper  
T: +44 (0)20 7796 6820  
E: david.webb@dlapiper.com  
David Webb is head of electronic data review at DLA Piper (International), based in London. He provides leadership and development, while advising internal and external clients in disclosure and electronically stored information best practices. He develops process improvements, protocol development and large-scale data analytics, as well as collaborating on cutting edge, firm-wide initiatives. This includes, but is not limited to, dynamic litigation solutions, e-discovery strategies, and cost-saving enterprise-wide software innovation.

Ben Sigler  
Partner  
Stephenson Harwood  
T: +44 (0)20 7809 2919  
E: ben.sigler@shlegal.com  
A partner with extensive experience in all aspects of dispute resolution including arbitration, litigation and regulatory investigations, Ben Sigler primarily focuses on the financial and sports sectors, white-collar crime and reputation management. He manages the firm’s sports practice. Mr Sigler has experience of dealing with issues arising from allegations of fraud and insolvency claims, including a significant number of applications for interim relief.

Mark Surguy  
Partner  
Weightmans  
T: +44 (0)121 616 6587  
E: mark.surguy@weightmans.com  
Mark Surguy has over 30 years’ private practice experience of advising and representing companies, institutions and individuals in domestic and international commercial litigation, fraud and insolvency cases across multiple sectors. His practice encompasses commercial misconduct, corporate investigations, freezing and search orders, asset tracing and other remedies arising out of suspected fraud, cyber fraud and civil and criminal wrongdoing.
FW: To what extent are regulatory and legislative developments shaping e-disclosure requirements and data-gathering processes?

Webb: A growing development is the requirement of proportionality when conducting an e-disclosure exercise, and this is shaping requirements and data gathering processes. Civil procedure rules requiring a proportionate approach to disclosure were introduced in 2015 in the US and the current pilot scheme in relation to disclosure in England and Wales is concerned with managing costs. The need to deal with ever growing volumes and sources of data has led to the English courts driving an issue by issue approach to disclosure. This impacts on the data collection process as the options range from no disclosure at all to a full train of enquiry.

Stretton: Regulatory scrutiny of corporate behaviour is intensifying and anti-competitive behaviour, privacy breaches and financial crime, such as bribery and money laundering, are all under the spotlight. Regulators cooperate internationally and launch specialised task forces, and enforcement is escalating as they take a much more aggressive stance against corporate misconduct. Regulatory investigations and requests for information typically require data to be searched for and produced in extremely short time periods, which has increased the demand for discovery solutions. Organisations are also responding by carrying out proactive and reactive reviews and investigations to identify and contain exposure. Added to this, legal mechanisms, such as the European Damages Directive, have been introduced. Discovery in these class actions is complex, requiring careful planning to contain costs.

Jones: The way that businesses create and manage documents or data responsibly, in an age of heightened accountability, has had a profound effect on the way that data is created, managed, stored and used. Large organisations are increasingly called to account for abusing their dominant positions and acting unfairly in the markets in which they operate. In pursuing these companies, the regulators have set the standards by which all of us must conduct our business, and meeting data requests is a key part of demonstrating cooperation and avoiding unnecessary fines. Today, the regulatory telescope has zoomed out to an expanded view, looking beyond market distortion, at the use of data and how the rights of individuals are protected. The General Data Protection Regulation (GDPR) requires companies to have a firm grip on the data they manage. All of this is in addition to the various pieces of legislation that have some bearing on the retention of documents in specific circumstances, and the highly developed court and arbitration rules that drive the disclosure of information in disputes. One of the biggest commercial risks to any business is the possession of data, so the way that corporations choose to de-risk themselves is also a factor in the development of e-disclosure best practices. Performing e-disclosure from scratch can be expensive and fraught with danger, such as the inadvertent production of trade secrets, and privileged or otherwise private data. Having a data discovery plan is critical, and while some still perceive such plans as being optional, those that do have arrangements in place tend to perform much better.

Sigler: The New Disclosure Pilot in the Business and Property Courts has radically changed the approach to disclosure, frontloading a lot of costs and giving rise to new tactical nuances which litigators can exploit. There are a number of core changes. First, an enhanced obligation “to take reasonable steps so that agents or third parties who may hold documents on the party’s behalf do not delete or destroy documents that may be relevant to an issue in the proceedings” and to provide confirmation of the same before issuing proceedings. The full extent of this obligation is currently uncertain – for example, in circumstances where a party does not want to tip-off former employees regarding a dispute. Second, an obligation to provide Initial Disclosure when issuing proceedings. Third, the introduction of a disclosure review document to be filed before the first case management conference (CMC), at significant cost. Fourth, an obligation on parties’ solicitors to certify that any claims to privilege are properly made. Finally, an explicit adoption of a range of different approaches to Extended Disclosure – which is made following the close of pleadings – which has moved away from “standard disclosure” as the default position, as was formerly the case, notwithstanding the wide range of options which a court was able to order, but generally did not order, in the alternative.
Surguy: Regulators have the power to require those firms they supervise to submit information to them. In the financial services sector, requests for disclosure to the Financial Conduct Authority (FCA) have created enormous challenges. The FCA has developed its own in-house e-disclosure capability and is knowledgeable about data. Often, the volume of data and the timescale for its production to the regulator require significant resources to be deployed to collect and review the data in question before it is produced. Regulators expect the regulated entity to know and understand its own data. Law enforcement authorities, including the Serious Fraud Office (SFO), also have data production powers granted to them. The 1987 Criminal Justice Act (CJA) provides the SFO with its powers. These not only confer the power to require anyone under investigation to produce data, they permit the right of entry and search enabling unannounced visits to premises where data and information is believed to be stored. The CJA permits the SFO to seize any documents it considers relevant to its investigation. The Competition and Markets Authority (CMA), which is a government department, has similar powers of production, search and seizure, as does Her Majesty’s Revenue and Customs (HMRC) under tax legislation. The existence of these powers to demand data be produced or to insist on searching for it itself on site is creating the need to address how these requests or ‘dawn raids’ will be managed. Those sectors that are subject to regulation, most notably financial services firms, have created specialist teams to respond to the challenges of multiple requests.

FW: Based on your experience, what are some of the common challenges and pitfalls associated with e-disclosure? How can parties mitigate e-disclosure risk?

Stretton: Most of the challenges around e-discovery flow from a failure to plan, allocate enough time for the exercise and scope the project properly before starting. The net effect puts pressure on the process, increases costs and compromises the quality of the end result. It is essential to start early, work with experts to assess the data universe and scope the project carefully. You need to work out how to extract the data needed to support key issues in the case and how to best use the advanced technology available. Given enough time, it is possible to provide the legal team with visibility of the types of data and documents likely to be relevant to the issues in the matter, define a targeted and efficient data collection approach and assess the effectiveness of predictive coding and analytics review technology. All of this makes it possible to project the costs of collection, processing and review, more accurately.

Jones: For most businesses, e-disclosure is something that happens quite infrequently, which means that relatively few people have experience of conducting searches for information. Another complication is that looking for documents means getting to grips with data and perhaps, learning about technologies used for storage and searching, concepts which are alien to some people and can make the business of e-disclosure somewhat confusing. Add large data sets and complex legal issues to the mix and you could find that some serious mistakes are made during the e-disclosure process. There are three major risk areas to consider when performing an e-disclosure exercise. Arguably, the most dangerous of these is over-disclosure – disclosing more data than is necessary, especially legally privileged documents. The second type of risk is under-disclosure. Disclosing too few documents can lead to serious questions and criticisms about the way a party has conducted itself in the litigation process and this may lead to additional costs, reputational harm and, in extreme cases, the loss of the case. The third type of risk is financial risk, caused by excessive or disproportionate spending on one or more elements of the search. By itself, e-disclosure technology is insufficient to reduce these risks. It has become the norm to deploy tools which help to keyword search large sets of data and automatically sift through emails to identify relevant topics. This is a great start, but it is a common misconception that these tools can deliver perfect results with little effort in a short space of time. They are tools which are equally capable of being used improperly and there is a wide variety of technology suited to different tasks. It is more important to choose the right data and technologies and plan an approach for performing the search in a realistic time frame. Businesses that face a higher volume of litigation might consider how they can create their own e-disclosure strategy for avoiding

MARK SURGUY
Weightmans

THE VOLUME OF DATA CREATED AND STORED BY FIRMS, INDIVIDUALS AND COMPANIES IS ALWAYS GROWING. THE CHALLENGE WHEN REVIEWING SUCH LARGE DATA VOLUMES IS TO SEPARATE THE RELEVANT FROM THE IRRELEVANT AND TO MAKE SENSE OUT OF A VAST COLLECTION OF DATA.

www.financierworldwide.com | FINANCIER WORLDWIDE | NOVEMBER 2019 | REPRINT
these risks and might seek assistance from experts with expertise in handling e-disclosure searches.

**Sigler:** Cross-border disclosure exercises require compliance with relevant local data protection legislation and, on occasion, in-country review. These types of disclosure exercises, along with other disclosure exercises, often involve the consideration of large numbers of foreign language documents, which can also be problematic. Other common disclosure challenges include, in relation to the harvesting phase, dealing with large numbers of technical documents or unusual file types, understanding the repositories in which documents are stored, collecting hard copy documents and the additional time it takes for them to be processed for uploading, collecting documents from custodians that are no longer employees, and collecting electronic data which is not held on a local network. A further challenge is agreeing the scope of disclosure and the mechanism of exchange between parties as early as possible, which the New Disclosure Pilot requires parties to engage in. From the perspective of reviewing harvested documents, key problems include the extent to which the review team has a full understanding of the issues in dispute before the review process commences, often the disclosure process reveals issues which the parties or their legal representatives were not aware of at the outset of the dispute, and which may evolve as the review process is undertaken, ensuring that the entire review team has an adequate understanding of the issues in dispute, and therefore what they are looking for, and also ensuring that the review process is managed in such a way as to ensure effective workflow with regard to subsequent stages of litigation – the preparation of witness statements – by documents being tagged accordingly.

**Surgery:** The greatest challenge is locating and collecting the data. Once this has been done, the interrogation of the data is considerably more manageable.

The cooperation of clients and individuals within complex organisations, particularly multinationals where data is stored in many different physical locations and in geographies that are subject to differing legal systems, is essential. Not all organisations are regulated, and those that are not are less conscious of the risk of demands for data production requests by external bodies. Resources are not always available to respond. Knowing where your data is and how to access it are issues which all well-governed enterprises should be cognisant of. Regulated entities have probably been exposed to the challenges more than any others and have been forced into developing their skills and capabilities. However, the volume of data created and stored by firms, individuals and companies is always growing. The challenge when reviewing such large data volumes is to separate the relevant from the irrelevant and to make sense out of a vast collection of data. The best form of mitigation is to anticipate the possibility of such a request coming out of the blue and assessing how the business would comply. A pre-planned protocol on what to do is highly advisable. Conducting a review against an externally imposed deadline requires the processing of the data into an environment suitable for conducting the review. This can, in some cases, demand 24/7 working across different time zones. Inevitably, significant pressure on resources is created.

**Webb:** Common challenges are the location, jurisdiction and volume of data encountered. A related challenge is that it is vital to handle data in such a way as to maintain its original integrity and it is very easy to alter metadata associated with a document. Ways of mitigating this risk are to have a clear understanding and application of processes and policies associated with data management and retention within the business. Records of where data resides and what happens to data from legacy IT systems when a business migrates over to a new system or when an employee leaves a business, are vital. A business should know the what, where and when in relation to the electronic data it holds and understand what data it has under its control.

**FW:** Once a dispute has arisen, what triggers the need to preserve data for e-disclosure purposes and what practical steps need to be taken? How important is electronically stored information (ESI) in this context?

**Jones:** It is useful to think of data as your evidential insurance policy, in case of a dispute. This means that you do not necessarily need to be in dispute with a party or even be contemplating some form of action, to warrant the retention of data. As we know, lots of cases develop years after the fact and the lack of available data is sometimes a problem for those
types of matters. Forward thinking is helpful for a business in a dispute, or even a crisis, because it can save important time during a critical moment. For those who are inclined to wait until a dispute arises, there is plenty of guidance to be found in the relevant court rules, such as the Civil Procedure Rules in England and Wales and the Federal Rules of Civil Procedure in the US. The general rule is that you should be taking steps to preserve evidence as soon as you are in contemplation of legal proceedings. Just beware that in certain types of cases, preserving data at this stage may mean that you have already missed the chance to obtain critical evidence.

**Sigler:** Much depends on the nature of the trigger point giving rise to the need to consider preserving relevant data and the seriousness of the issues in dispute. The approach to be adopted in relation to a minor complaint by contrast to an allegation of serious wrongdoing, is naturally different. The obligation to preserve documents for the purposes of civil proceedings is triggered by a party knowing “that it is or may become a party to proceedings that have been commenced... that it may become a party to proceedings that may be commenced”. As previously noted, pursuant to the New Disclosure Pilot, the obligation to preserve documents extends to parties, agents or third parties who may hold documents on the party’s behalf, it also extends to their legal representatives. Given that the vast majority of documents relevant to any dispute are now held electronically, electronically stored information (ESI) now represents the key focus of any document preservation exercise.

**Webb:** Civil procedure rules include a requirement to preserve data. Civil procedure rules (CPR) in England and Wales, for instance, place a duty on a party to preserve data, including an express duty on the solicitor instructed to advise the party. This requirement extends to ex-employees who potentially hold data pertinent to a dispute. Practical measures extend to implementing litigation holds within technology, collecting data from servers and imagining laptops and mobile phones. ESI is important in this context as it now forms most of the information to be considered in a disclosure exercise. Preservation of data while maintaining the original metadata is fundamental to establishing a defensible approach to conducting an e-disclosure exercise.

**Surguy:** Once a dispute has arisen it is generally sensible to assume the requirement to preserve has arisen. Strictly speaking, preservation obligations are driven by the likelihood of legal proceedings. While not all disputes are going to end in legal proceedings, the danger is that if data is not preserved at the right time and is then lost or altered, the capability to prove or disprove important details of fact may be impaired. An opponent will not be slow to argue that preservation has not been done properly with potential adverse consequences, not only in costs but also in terms of the merits because certain facts can be taken as established against the party in default. The case can be weakened as a result, and an advantage can be served up to the opposing party unnecessarily. ESI is particularly important in this context because it is easy to alter or delete and may not be recoverable once deleted. Practically, a well-thought through document retention policy should be in place with a scheme to suspend any routine destruction in the event of a dispute. The IT department is a crucial player in being able to centralise non-destruction. Information and communication technologies (ICT) infrastructure can be programmed to put a preservation process in place automatically. More organisations are turning to this kind of automated approach. Mapping the company’s IT systems and understanding the location of, and access to, stored data are essential.

**Stretton:** The need to preserve evidence is legally defined either in common law or in the rules governing litigation. For example, in civil litigation in England and Wales, a party that knows that it is or may become a party to proceedings that have or may be commenced, is under a duty to preserve documents. This includes “documents which might otherwise be deleted or destroyed in accordance with a document retention policy or in the ordinary course of business”. The rules go on to state that the duty to preserve includes the obligation to suspend relevant document deletion or destruction processes for the duration of the proceedings and to send out a notification to all relevant employees and former employees to hold relevant documents. When a litigation hold is issued by a legal team, it is important to consider that electronic data sources need to be carefully considered. For example,
automated data destruction processes need to be halted.

**FW:** What methods are available to collect data in an efficient, forensic manner, bearing in mind that data is dispersed across multiple locations, including social media and the cloud?

**Sigler:** There are various methods of collection available, including in-person collection by an e-disclosure specialist, collection overseen in person or remotely by an e-disclosure specialist, collection conducted remotely by an e-disclosure specialist, and collection fully conducted by a client’s in-house team, potentially having had a conversation with an e-disclosure specialist. In this case, one would need to be comfortable with both the capability of the client’s in-house team, and the risks or drawbacks involved. This can be an easy point for the other party to exploit in the event that any errors in the disclosure process emerge, and the judiciary are unlikely to be sympathetic to such errors, where they could have been avoided by the party’s legal representatives taking the lead.

**Surgery:** If data is to be collected in a way that does not involve the compromise of the metadata, then the correct IT tools and data collection skills are needed on site. Some tools may enable online collection. The nature of the case and the need for a ‘forensic’ collection should be discussed and appreciated at the outset. Where data does not need to be collected forensically, it can be copied to a hard drive or uploaded to a file transfer protocol (FTP) site. A planning exercise to map where the data is stored, how to gain access and what form of download, upload or copying is appropriate should be carried out in consultation with a specialist service provider, if one is not available in-house, and in collaboration with the legal team before any collection takes place. Coordinated collections may need to be taken simultaneously across different time zones and in multiple locations.

**Stretton:** Efficient and defensible collection of data in an e-discovery exercise begins with proper scoping to identify the various data sources and custodians involved in a case. Experts are able to collect from various sources such as laptops, desktops, tablets, mobile devices, servers, network shares, MacBooks, backup tapes, exchange or domino servers and Office 365. External experts can design a collection approach and can either assist a client with self-collection or go onsite to do the data collection for them. Onsite collection is often needed in engagements such as dawn raids, where data needs to be collected quickly or in overnight collections when business interruption needs to be avoided. Forensic data collections which require independence and deep data investigation, typically to uncover evidence of misconduct, are also well suited to onsite collection by experts.

**Webb:** The best approach is to conduct a scoping call including client’s IT and legal teams, an outside legal representative and third-party forensic data collection experts. This can establish the ‘what’, ‘when’ and ‘where’ and consider the ‘how’ of collecting data. The collection process has to fit the requirements of a particular matter. There are basics which should be observed, however, and it is not as simple as selecting copy and paste. The Association of Chief Police Officers (ACPO) guidelines for computer-based evidence in the UK, for instance, offer a sound approach to collection and handling data, whatever the reason for its collection. Tools are available for collection from social media. There is often potentially interesting data cached on a person’s phone.

**Jones:** The data mining industry has received huge investment over the last 20 years. Lots of new products specifically designed to capture, process, review and produce electronic data for court proceedings have been developed. Over the last 10 years, artificial intelligence (AI) has risen to the surface of the e-disclosure market and product innovation has generally kept pace with the challenges of the current world. Automated products have been created to serve companies that may have large numbers of subsidiaries, employees and electronic devices, that may operate in dangerous territories, and that may have a need to interrogate public sources of information for evidential purposes. Almost at the touch of a button, a business with 100 evidential custodians spread across four continents can see exactly how much data it may have stored on its devices.

ROBERT JONES
Ankura
may indicate that it is time for a change in the way that you manage your litigation risk. Many of these products make ‘DIY data collections’ possible. Before rushing to use one of these solutions, it is worth remembering the basics of forensics – the process must be repeatable and achieve the same results. A mistake could be the undoing of your case. This means you should always consider whether you have the competence and capacity to perform an evidential collection of data in the time required. If in doubt, hire an expert.

FW: What innovative technologies are available to help parties interrogate and review their data? To what extent do these rely on artificial intelligence (AI)?

Stretton: AI and machine learning (ML) technologies, like Relativity and Reveal and a plethora of others, are being used routinely by lawyers to streamline the review of documents in cases, whether that be for case building or disclosure. Typically, lawyers will read a very small percentage of the document population online and the software algorithm learns from the lawyers and then automatically reviews and categories the rest automatically. Image classification software is also now available and can classify scanned images of documents in a similar way, which is useful in cases such as construction litigation with many handwritten documents. Early case assessment technologies, like Brainspace, for example, can be used to explore the key issues early on and help shape case strategy. This is an example of Big Data analytics, where AI is used to extract insights from a set of data in discovery. You can see 5 million documents on your screen and the software suggests what the key issues are. Solutions like NextLP offer sentiment analysis on communications such as email, meaning that the software can purportedly work out the emotions conveyed in emails. Taking this a step further, we are starting to see Big Data analytics being used to mine data and come up with suggested case strategies and predictions on case outcomes.

Webb: There are a number of unguided and guided analytics tools which can assist the interrogation and review of data. Unguided analytics clusters data together by themes allowing a doorway into the data. Guided analytics use the input of a reviewer to drive the tool. Approaches coming to the fore include the application of natural language processing, support vector machine and neural networks. All these are forms of active ML which assist in prioritising data and use statistical analysis to justify the decisions made by the technology. This assists bubbling relevant documents to the top. It also assists the application of appropriate resources and costs for review, based on the likely relevancy of the datasets.

Jones: There are multiple technologies designed to assist with e-disclosure. Each has certain characteristics which make them suitable for one or more of the stages involved in the process. Some tools are designed to provide an overview of the source data at the broadest level. These help you understand the size of the ‘haystack’, while others carve the data into different dimensions, helping you to find the ‘needle’. The technologies available today are distantly removed from the original platforms that enabled parties to read each and every document, one after the other, in a linear date order. Those types of technologies made it easier for teams of reviewers to read the documents but did not address the growing problem of finding relevant data quickly. In the present day, the courts expect parties to be using advanced legal technologies, which make it possible for millions of documents to be accurately reviewed, in a matter of days, without enlisting armies of staff to read the documents. Instead of reading documents one by one, early data assessment presents the data in visual terms. This means that a lawyer with the best knowledge of the case can run keyword searches and evaluate the results by comparing them to timelines, clustered topics and communication charts. These types of tools help lawyers to narrow and broaden the scope of the data as needed, and generally select proportionate amounts of data for a finer legal review. By way of example, a collection of 900,000 documents could be dramatically reduced to 100,000 documents within a matter of hours or days. The detailed legal review may also be enhanced using a form of AI known as technology assisted review (TAR). This type of technology observes the decisions made by the human reviewer and simultaneously puts forward documents that are more likely to be relevant, based on those which the reviewer has confirmed to be relevant. This type of technology empowers the review team to find key evidence at the beginning of the review and to choose to cease the review when the return of relevant documents diminishes to an acceptable level.

Surguy: Online review platforms are commonplace. The software used is constantly changing. It is the use of this software that gives rise to the possibility of innovation. So-called predictive coding is one form of AI. The software, much like Google and Amazon searching, which are widely encountered outside a legal context, recognises what the professional teams consider to be relevant by reference to relevancy criteria identified by the lawyers, law enforcement agency or the regulator. These criteria can be programmed into the software and developed as the understanding of the case develops. The software can then prioritise the most relevant data in the data set and present it to the reviewers to be looked at first. The software learns the criteria and automatically searches for all data matching the pre-programmed parameters. Through negotiation and agreement with the regulator or opposing legal team or with the law enforcement agency it may be possible to avoid looking at all the data. These tools are beginning to replace the more traditional keyword searches that have been used as relevancy criteria. AI is less blunt and can take account not only of the presence or absence of particular words in a dataset, but can examine contexts and subject matter in a broader manner in a way that avoids relevant documents being missed.
This lack of precision occurs because the chosen keyword may not be present in a document that is relevant, for example. Or irrelevant documents are produced because they do contain the chosen keyword but the context of the document makes it otherwise irrelevant.

**Sigler:** Various innovative technologies are available. These include TAR 1.0 and 2.0, near de-duplication, and email threading and tools such as Brainspace or NexLP which allow unstructured data to be interrogated, filtered and reviewed in innovative ways. Other software offers anomaly detection, where the system identifies documents that are outliers from ‘the norm’, and these can be very useful. For reviews where a population of documents contains a significant number of structurally similar documents, specific types of contracts or leases, specialist tools can assist with the identification, extraction and review of specific clauses or data points.

**FW: To what extent is good data governance impacted by data privacy and cyber risk?**

**Surguy:** Having a procedure for the retention of business information is double-edged. On the one hand, a business might be looking to retain data for as short a time as possible, particularly where the data is classified as personal data for the purposes of the GDPR. The GDPR generally encourages organisations not to store personal data longer than is necessary for specified purposes. The less data that is stored, the less it is exposed to unauthorised access from within or without an organisation. A good procedure may also mean not having to answer data subject access requests: if the data has been legitimately deleted, it does not have to be provided even if a right to request the data exists. If the data no longer exists, the right of access to it is redundant. However, if data is deleted too early and information is not stored for the future, it can be more difficult to prove or disprove relevant facts in the context of a legal dispute or regulatory intervention. A well-planned scheme of data governance will seek to balance the need to minimise data storage to reduce the risk of having to give access or access being gained unlawfully and to maximise the availability of valuable or useful business information. Legal requirements in relation to the retention of some classes of data will dictate what cannot be deleted. Understanding how these inform a data governance strategy is a key function of the strategy itself.

**Jones:** Good data governance is only strengthened by rules concerning data privacy and trends in cyber risk – or at least it should be. Think of a business as a kitchen and its data as its work surfaces. If you do not keep your kitchen work surfaces clean and tidy then, sooner or later, you should expect to be sick. In the same way, if you do not know what data you have, or where it is kept, then you will be in the unenviable position of being unable to produce evidence when you may need it, and being unable to protect it from internal or external threat actors. Maintaining good records about your data and where it is stored and how it is protected is a healthy position to be in. Acting in a way that is both compliant with the GDPR and mindful of the wide variety of cyber risks that could impact your business should assist your data governance initiatives in a positive way.

**Sigler:** One only need bear in mind the financial and reputational repercussions of the recent data breaches involving British Airways, Marriott and Tesco Bank as examples of the potentially existential impact of data breaches on organisations outside the legal sector. These impacts could easily be mirrored in the context of a data breach arising out of a disclosure exercise, where a third party accesses personal data, sensitive personal data or other confidential data, such as trade secrets. Indeed, the reputational impact of a data breach for a law firm is likely to be even more acute than that suffered by other businesses, given the sensitivity of the data which law firms hold. From a law firm’s perspective, in order to discharge its obligations under the GDPR it is important that its position is appropriately protected in relation to data processors, including e-disclosure providers, with whom it works. Amongst other things, law firms will need to consider ensuring appropriate contractual protections, such as indemnities, are in place and undertaking due diligence into the IT security measures of data processors appointed by them both at the outset of their engagement, and, if necessary, on an ongoing basis.

**Stretton:** The UK Information Commissioner’s Office earlier this year announced its intention to impose a £183m fine on companies for a data breach. The starting point for a company wanting to enhance its privacy and security programme is to look at it through the lens of ‘data governance’. Privacy, cyber risk and data management need to be viewed and managed holistically through an integrated programme. Often, a traditionally deployed approach is siloed, with no input from technical security experts and privacy officers. It will prove beneficial to replace such an approach with one involving constant communication, collaboration and an overarching strategy designed to ensure compliance with privacy laws, reduce cyber risk and unlock data-related value. A standard data governance framework used at a mature organisation typically includes a central data risk steering committee, a data governance steering committee or a GDPR steering committee. All data stakeholders should be represented, including operations, finance, HR and marketing. Such a committee ensures that policies and business practices are aligned, while taking responsibility for implementing policy and periodically auditing the organisation. At a secondary level, mature organisations will also have data security and privacy champions who act as brand ambassadors for responsible data governance and as liaisons for all business units. The heavy lifting, when it comes to ensuring compliance and protecting data, is carried out at the third...
tier of the data governance framework by dedicated privacy and cyber teams.

FW: What advice would you offer to companies on implementing best practices for data retention across their organisation, to assist with e-disclosure demands in the event of a dispute?

Jones: Be lean, be smart and be ready. The biggest challenge in the face of disclosure, whether it is for litigation, or to satisfy a regulator or a data subject making an access request or for any other purpose, is the task of sifting through large data volumes in short order. Tackling a comprehensive search of any magnitude under a time pressure is very likely to be expensive and risky, so embedding good data governance principles into your business as usual strategy will be helpful. Such principles would drive your reasons for keeping data and the amount you retain, plus a workable strategy for the deployment of teams and technology when the time comes to respond to an information request.

Surguy: The key is to involve the relevant stakeholders in the planning process. The legal function, IT team, technology services consultants, divisional directors and the board should take advice from each other collaboratively in order to identify the risks and pitfalls which will be driven largely by the nature of the business, where it operates globally and whether it is regulated. Those that are regulated or litigate regularly – for example, patents disputes in the pharmaceuticals sector – may be advised to create in-house teams with their own in-house e-disclosure systems. Others may need to have a list of preferred suppliers with pre-vetted capability to provide services on demand when the need arises. Best practice is learned through experience. That experience can be built up internally or can be bought in from those who practice in this field daily.

Sigler: Given the need to respond to data subject access requests in a timely and cost-effective manner, and their obligations under GDPR more generally, one expects that companies should now be better placed to understand the location and sources of their data – including archive repositories and backups – than prior to GDPR coming into force. In any event, all companies should have in place policies and procedures specifying how to proceed in the event that a dispute arises. At a minimum, this should prescribe points of contact who will be responsible for dealing with sending out litigation hold notices and ensuring automatic deletion processes are suspended, who will manage the collection of relevant data so that it can be secured or harvested, and the approach to creating any further documentation, particularly given issues relating to privilege. It should also include technologies which enable companies to manage, preserve and extract their data on an enterprise level.

Stretton: The biggest challenge with e-discovery is the vast volume of data that companies store that needs to be searched across to find what is needed for a particular case. Litigious companies need to consider the electronic discovery reference model (EDRM) and the guidance offered to manage data upstream of litigation in a way that reduces discovery downstream. Data retention policies that take into account litigation risk are essential. Procedures to put in place litigation holds are also critical to avoid court-imposed sanctions for failing to preserve data that needs to be produced in litigation. Beyond legal obligations to retain data for regulatory purposes and for litigation, the best advice has to be to delete data that has no value to the business. Data minimisation is a strong theme in global data protection law regimes and also reduces cyber risk. In addition, having well-organised and sleek data stores reduces the e-discovery burden significantly. Companies that are serial litigators should also put in place internal discovery procedures to prevent producing the same data repeatedly.

Webb: An organisation should retain data which is necessary for running the business and to comply with any regulatory obligations. Outside of this scope, an organisation should question why particular data needs to be retained. A sound knowledge of the data map of a company can assist with the identification and collection of potentially relevant data caught by disclosure or regulatory requirements in the event of a dispute arising.

FW: What are the key global variances in e-disclosure requirements and best practice?

Stretton: In civil law systems, the concept of discovery is not recognised and parties to the case produce the documents they rely on in court and if any further are needed the presiding judicial office will call for them. The terms ‘discovery’ and ‘disclosure’ are commonly used in both England and the US to describe the process of pre-trial evidence collection and production. In both the rules governing civil litigation in England and Wales, the Civil Procedure Rules (CPR), and in the Federal Rules of Civil Procedure (FRCP), the term ‘disclosure’ refers to each party’s duty to provide other parties with certain categories of information. In the US, ‘disclosure’ is followed by ‘discovery’, whereby parties have an opportunity to seek additional information, from each other and other sources, through several avenues, including specific document requests, depositions, interrogatories and onsite inspections. Where information is properly requested by one party during US discovery, the responding party is generally under a duty to produce them, unless the producing party can raise convincing arguments to the contrary. In England, until recently, each party had to disclose documents on which it relies and which support or adversely affect either its case or another party’s case. This therefore included adverse and damaging documents and was known as ‘standard disclosure’ and replaced the former, and wider, definition of ‘relevance’ as the basis of disclosure. As of 1 January 2019, a new disclosure regime is being piloted in the UK. The pilot is the product of a two-year
review led by the judiciary, in partnership with the legal profession and court users, to combat the increasingly burdensome time and cost associated with disclosure in commercial and property cases. The hope is that it will ensure that disclosure in commercial property and business cases is tailored to the specific issues to which disclosure relates. The pilot strives to achieve a mini disclosure to start with. Anything greater needs agreement from a judge, has to relate to the issues to be decided and must be proportionate in cost terms. What it requires is proactive conduct on the part of judges. Extended disclosure is not being thrown out but is often not necessary in the vast majority of cases and does not need to be the default norm. Early indications are that the Pilot scheme is working.

**Sigler:** The extent of parties’ disclosure obligations, and the relevant regulatory environment governing the disclosure, vary from jurisdiction to jurisdiction. An issue of particular importance is ensuring that GDPR is properly complied with in the course of a cross-border disclosure exercise. For example, one exercise which we conducted prior to GDPR coming into force, involved undertaking a first pass review of documents collected in the UK for use in US proceedings, to ensure that any documents containing irrelevant personal data were identified and removed from the data to be exported – though that approach might need further adaptation post-GDPR given the onerous restrictions on data transfer therein. It is worth noting, in this regard, that such exercises are likely to become more commonplace assuming that Brexit occurs, as the UK will become a ‘third country’ from a GDPR perspective.

**Webb:** The key global variiances in relation to e-disclosure, or e-discovery as it is referred to in the US, are based on the differing disclosure obligations between common law and civil law jurisdictions. Common law jurisdictions bear the burden of more onerous disclosure obligations, including disclosing documents in their possession which may not assist their own case. Data protection privacy and state secret considerations vary globally and need to be taken into account in all matters. Considerations vary depending on where the matter is situated and especially for those involving cross-border transfer of data. It may be necessary to conduct an initial review onsite, for instance, in order to satisfy requirements and before data can be transferred. In terms of best practice, the principles remain the same regardless of where the electronic data is being handled.

**Surguy:** There is no single rule or practice which applies universally. The common law jurisdictions have developed rules and guidance to manage large volumes of data in legal contexts including law enforcement scenarios, the exercise of regulatory powers and litigation. There are different approaches to what might be considered ‘relevant’ and therefore disclosable. There may also be differences in the understanding of the discloseability of deleted data or data that is not immediately accessible, such as that stored on legacy systems that are no longer in use. UK practice has tended to follow US practice. In non-common law jurisdictions, practice may vary even more because there is no legal framework or no highly developed framework governing the circumstances in which data must be produced. In the European Union (EU) the GDPR provides a framework in relation to personal data. This regime may extend to corporate groups outside the EU or to organisations that trade with the EU. Non-EU member states may have less restriction or governance requirements in relation to the preservation or production of data.

**Jones:** E-disclosure is a procedure of the common law system of England and Wales. Other common law systems, such as the US, have similar procedures for the exchange of documents that are relevant to the litigation at hand. In civil law jurisdictions, such as Germany, France and the Netherlands, the procedure does not exist, however companies that are domiciled in those jurisdictions and involved in cross-border litigation, where disclosure or discovery is a requirement, must be prepared to comply with searches that may be ordered. The overriding best practice for any business should be to remove geographic borders from the equation and form a disclosure strategy that adds value to your business. An assessment of need does not look solely at whether the business is likely to be engaged in proceedings before the High Court. A proper evaluation might also consider the future growth potential of the business. If disclosure of documents is not realistic today, might it be on the horizon, as the business attracts new customers and begins trading in new countries? There is also a regulatory and compliance angle to e-disclosure best practice. Being ready to investigate and cooperate with trade-related investigations is a huge advantage for any business. These types of investigations involve wide-reaching powers of search and hefty financial penalties for wrongdoing. The risk of business interruption and potentially crippling fines should be a sobering incentive for most business leaders to establish a proper internal system for cataloguing and searching data when needed. A general approach can always be tailored for the specific regional needs of different courts and agencies that may request information.

**FW:** What models can companies adopt to manage e-disclosure efficiently?

**Sigler:** Ideally, if they are to undertake document reviews in-house, companies should have at their disposal a number of options in terms of technology. For example, it would be preferable to use a cheaper, lightweight solution for smaller cases, but a more expensive, feature-rich solution for cases where the additional functionality justifies the additional costs entailed. They should also consider having a number of pre-agreed options in terms of pricing models with e-disclosure providers, such as with regard to processing, an ‘in-out’ model, a ‘tiered per GB’ model or a ‘pure GB model’. Clear policies and procedures around the
preservation and extraction of data – such as ensuring employees have consented to the processing of their personal data to this end – are also essential. Finally, companies should ensure that progress and costs are carefully monitored on in-house reviews.

**Surgery:** A tailor-made approach is best. Within the confines of bespoke, case-led data management, however, data governance protocols and response mechanisms are best put in place on a risk-based basis. A particular software package could be identified as being suitable for the anticipated tasks, although with rapid changes continuously taking place, the most important element will be to have access to data consultants who can advise on what software and processes are best deployed in the particular circumstances of the case.

**Webb:** The EDRM is an effective framework for companies to adopt to manage e-disclosure efficiently. It is designed to cover the whole process from information governance right through to presentation at any hearing. The EDRM has been amended to accommodate active ML, however the fundamentals of reducing the corpus of non-relevant documents through to uncovering and reliance on relevant documents remains the same. Project management of all the stages of the EDRM is essential to manage e-disclosure efficiently.

**Jones:** Broadly speaking, companies should base their e-disclosure plans around the long-established EDRM. At the core of this model are the steps required to identify, preserve, collect, process, review and produce electronic documents as evidence in any proceedings. Each of these are responses to a trigger incident and the EDRM is a useful step-by-step guide to getting organised for the day when you might need it. Although presented in summary form, the stages of the EDRM have been the subject of much thought, and various sub-models have been developed off the back of it. For example, there has been much debate around the proper way to use TAR and it now has its own framework under the EDRM. Similar think tanks, such as the International Legal Technology Association (ILTA), have developed a document exchange protocol, which contains detailed guidelines to produce documents for inspection in UK proceedings. A lot can go wrong when working to produce information to another party. Privileged documents can be inadvertently disclosed. Relevant documents can be inadvertently concealed. Due to the adversarial nature of disputes and investigations, a company’s opponents will likely question the integrity of any disclosure if anything looks out of place, so it pays to be familiar with all the current thinking around disclosure. No discussion of the EDRM is complete without mentioning the preliminary stage concerning information management. This is often an afterthought because it is based on all the best lessons learned from e-disclosure projects that could have been managed better. It is a guide toward the effective management of data before a dispute arises. Anyone actively thinking about how business processes feed into e-disclosure workflows can be considered a highly mature thought leader.

**Stretton:** Companies need to have in place a cross-functional team including legal and IT that is responsible for handling discovery requests in litigation and investigations. They can reduce the burden and cost of e-discovery by having defined processes in place for collecting, processing and reviewing data in order to respond to e-discovery requests. Organisations like banks, that litigate often, often deploy their own technology and internal experts to manage e-discovery. Others opt for a managed service offering in terms of which the e-discovery process is managed by an external e-discovery provider. Those companies that have a low litigation risk profile are still advised to form relationships with external experts and to keep a close eye on emerging technologies, so that they can respond quickly if the need arises to search for and produce data.

**FW:** How do you expect e-disclosure requirements and processes to evolve in the years ahead? Are any particular trends likely to impact the process?

**Webb:** Trends will likely involve greater application of analytical AI technology to datasets. More tools will have elements of this type of functionality, allowing this type of analysis to be carried out in-house. The use of the cloud will likely increase, allowing flexibility and processing power to be dialled up or down depending on requirements. The particular trend for courts to apply proportionality considerations to the costs associated with e-disclosure is unlikely to go away. The drive for this is the ever-growing sources of potential data with advances, such as driverless cars, all on the horizon.

**Jones:** The rules governing e-disclosure have been revisited several times since they first appeared in 2005. Each revision seems to be aimed at reducing the burden and costs of e-disclosure and radical changes are currently being piloted in the Business and Property Courts under Practice Direction 51U. As we anticipate the outcome of the pilot, at the end of December 2020, instinctively it feels as though the rules will have to change again. New technologies are constantly emerging and touch all aspects of the document lifecycle, from its creation to its disclosure if it becomes evidence. Businesses and employees are widely using emojis, memes and videos to communicate in a different way. Companies are widely seeking to streamline manual processes with automation. Facebook recently announced it would acquire a startup that is developing ‘neural wristbands’ – wearables that will enable users to control computers using their minds. All of these are new forms of evidence that will define the way that investigations will be managed in the future. At the evidence-gathering end of the lifecycle, we already have technologies that can perform sentiment analysis, to help us seek out
aggressive, uncertain or unhappy tones in emails. Emojis and memes can be detected but are difficult to interpret. In time, the courts may be asked to give direction on the proper use of such technologies. And the costs debate is never likely to go away. There may be further calls to abolish disclosure altogether. However, dispensing the need to disclose documents, if that ever happens, is very unlikely to obviate the need for a review of the documents, since recorded information has and always will be a useful method of corroborating oral evidence.

**Stretton:** Given the sheer volume of data and documentation that needs to be considered in disputes, we can expect to see a narrower approach being adopted to disclosure. In arbitration, for example, International Bar Association (IBA) rules and the Prague rules are focused on getting to the heart of a matter without causing the parties to spend small fortunes on disclosure doing so. The need for a narrower scope has already been recognised by UK courts in litigation too in the new Pilot scheme. We can expect enterprise systems like Office 365 to include discovery modules, making it easier for companies to carry out disclosure themselves without exporting their data to specialist third-party providers of e-discovery technology. Law and AI are increasingly becoming intertwined. In the future, experts predict that we will see AI judges determining the outcome of cases. That might seem far-fetched, and algorithms cannot yet apply subjective judgement or exercise discretion, but online dispute resolution is being trialled in UK courts. More and more online services are emerging to help litigants resolve disputes without the cost and burden associated with the current rule-heavy systems. These DIY systems might bypass disclosure as we know it. In 2016, a UK teenager designed an artificial intelligence lawyer called DoNotPay to provide free advice to people on how to contest parking fines. This is obviously for simple claims, but it gives us a hint of what is on the horizon and is being extended to handle claims against companies and may have a role to play in class actions.

**Surguy:** E-disclosure will be driven by the use of AI and the nature of the evolution of the software and technology. Less human intervention and the greater use of automated, ‘intelligent’ processes will grow. Large review teams will tend to be used less. Armies of paralegals reviewing millions of documents will be replaced by a much smaller corpus of experienced legal practitioners who will work closely with the appropriate software to use new techniques to identify relevant material and exclude irrelevant material.

**Sigler:** There will be changes both to the law governing disclosure reviews and to the available technology. It remains to be seen what impact the New Disclosure Pilot ultimately has; at present, the full extent of its impact is unclear, although initial indications suggest that there seems to be a broad range of different disclosure models being ordered, and Model D – which is most equivalent to standard disclosure under CPR 31 – is not the dominant model. Model C is being used quite a lot, with some Model B too. As would be expected, there is not much evidence of Models A or E being used. TAR is now well established as an industry standard, and this is reflected in the New Disclosure Pilot. The technology around TAR will continue to develop, facilitating further efficiencies which, in turn, will lead to greater focus on proportionality of costs. We also expect more consolidation in the marketplace with respect to technology that sits in different areas of the EDRM lifecycle. There may also be an increase in service providers building their own ‘end-to-end’ e-discovery platforms and law firms and in-house legal teams continuing to build in-house capability, and procuring technology on a managed service or SaaS basis.

**FW:** In your opinion, should parties wait for a dispute before using e-disclosure technology?

**Jones:** Parties should not wait for a dispute before using e-disclosure technology. The technology is highly effective for fact investigations and intelligence gathering in many scenarios that do not necessarily involve or lead to a dispute. Businesses have used e-disclosure technology for internal investigations into employee conduct, compliance risk assessments, case preparation and due diligence around the acquisition of a target. Some companies that have a higher litigation risk also use e-disclosure technology to curate databases of documents as part of their ongoing litigation management strategies. Optimising the disclosure process transforms the experience and reduces the sting of having to organise a response under pressure.

**Stretton:** Disclosure technology can be used very effectively by companies to support compliance efforts. If a company needs to check on whether or not its anti-bribery, anti-money laundering or antitrust policies are being properly implemented, it can use e-discovery technology to carry out internal health checks. By reviewing a few mailboxes of key individuals, you can see what is really going on and if there is anything the company needs to worry about. E-discovery technology can also be used effectively to respond to data subject access requests, which require a company to search for the personal information of data subjects that it might hold and produce it in a short time period of a month. AI and ML technologies which automate the review of documents can also be used for contract review.

**Surguy:** The risks posed by storing data are not confined to the possibility of legal disputes. Security of personal data, the possible involvement of the industry regulator, where there is one, the risk of law enforcement activity from abroad as well as at home, all create the potential for data having to be collected, analysed and produced outside the organisation. Data subject access requests are not dependent on the existence of any kind of dispute. Disclosure technology can be used to
assess ongoing risk, for audit purposes and to identify what data exists within a business. Data governance routines will usually involve the use of some e-disclosure related technology. Good data governance is not dependent on the existence of litigation.

**Sigler:** In short, a trigger point is required for there to be any need to use e-disclosure technology, whether this is a complaint giving rise to the need to conduct an internal investigation, or the threat of potential litigation. Ideally, parties should have access to the necessary technology as soon as possible. However, this depends on the extent to which the costs of using such technology are justified in all the circumstances. It should be borne in mind that, given the obligation under the New Disclosure Pilot on parties to disclose known adverse documents, there is a risk that undertaking a wide harvesting and review exercise early may result in an obligation to disclose documents which might otherwise not have been within a party’s knowledge.

**Webb:** There are countless applications of e-disclosure technology which fall outside of litigation. Fundamentally, the technology allows a party to search across various file types, apply AI and analytics tools, tag documents and produce those documents in an array of format options. The technology is ideal for targeted pre-action searches. Caution, however, should be taken should litigation be in contemplation and the application of the tools should be made in conjunction with overall legal advice.