The disclosure iceberg

Jonathan Maas outlines the factors that influence the cost of disclosure

Most people involved in litigation are familiar with the notion of submitting costs budgets to the courts via the uniquely succinct Precedent H. Above the waterline, at the tip of the disclosure iceberg, Precedent H allows for the following main categories of cost for disclosure:

1. Fee-earners' time costs (note assumption of multiple fee-earners);
2. Expert's costs (note assumption of single expert);
3. Counsel's fees (note assumption of single barrister);
4. Court fees; and
5. Other disbursements.

Under 'Disclosure' in the new (June 2015) guidance notes to that precedent, the parties are reminded to include costs likely to be incurred by:

1. Obtaining documents from the client and advising on disclosure obligations;
2. Reviewing documents for disclosure, preparing a disclosure report or questionnaire response and list;
3. Inspection;
4. Reviewing an opponent's list and documents, undertaking any appropriate investigations;
5. Correspondence between parties about the scope of disclosure and queries arising; and
6. Consulting counsel, so far as appropriate, in relation to disclosure.

But these specifically enumerated expense categories are only the tip of the disclosure iceberg in many cases. The sort of civil matters with which I have been involved over the last three decades or more have largely shown the same attributes: big organisations in acrimonious disputes over huge amounts of money after the failure of long-running contracts. Standard disclosure is the norm; the concept of proportionality is more honoured in the breach; and there is likely to be an enormous amount of disclosure. Matters of this scale may be the minority in terms of numbers of disputes going through our courts, but they support a massive global industry – and are not going away.

For these major disputes, most disclosure costs relate to activities that occur before actual disclosure – the underlying mechanics. While Precedent H allows a party to set out a budget for the costs of the lawyers, experts and court, it provides only a simple, one-line 'other disbursements' catch-all for any other costs, including those related to the practical elements that allow one to comply with disclosure obligations.

WHAT LURKS BENEATH

Receiving an itemised bill with 50% of the costs neatly set out in four clear categories, and the other half hidden under 'other disbursements', is akin to seeing the tip of an iceberg with a serious hint of the extra danger beneath. Yet that is all that Precedent H requires parties to do. Litigation funders may be forgiven if they feel rather like the captain of the Titanic at this point. The cost of the lawyers, experts and court is usually based on the passage of time, with the only variable being the length of that passage. In the mechanics of disclosure there are many, many unknowns, any one of which can wreak significant damage when (not if) one sails into it. By their very nature, these variables only reveal themselves as you hit them.

One could usefully here distinguish between a 'budget' and an 'estimate'. According to my *Penguin English Dictionary*, a budget is 'the amount of money available for or required for a particular purpose'; while an estimate is 'a statement of the expected cost of a job'. Precedent H refers to these things as budgets, which do not change. In my world, we estimate the likely cost of doing something based on assumptions that always change.

When dealing with things that change, a logical solution is to factor in a contingency sum and, at the appropriate point, revise one's estimate appropriately as the way ahead becomes clearer. But this article will make it clear that it is extremely difficult to estimate a realistic contingency sum because of the many variables in play. Furthermore, in the litigation context, parties seem to be generally reluctant to make applications to court to amend their so-called 'budgets', perhaps fearing the wrath of a judge unfamiliar with modern disclosure. Parties are obviously even more reluctant to amend their 'budgets' a second, third or fourth time – yet there are many ways in which an honest estimate can change over time, as new information comes to light.

It is worth noting that the court requires the budget before the first case management conference, so that it can decide how much of what is budgeted for could and should be done proportionately. One wonders about the validity of decisions made at this point, if the costs cannot be adequately estimated at this stage.

Continued on page 18

Continued from page 17

CHARTED AND UNCHARTED WATERS

Before setting sail on a discussion of the variables that can sink the budget, it is worth giving a broad overview of the disclosure process and mechanics, for those less familiar with it.

As soon as a party anticipates that a dispute may occur, they must preserve any potentially relevant documents or the court may draw an 'adverse inference' about why they did not do so. As trial approaches, each party provides the other with all their relevant documents, whether harmful or supportive, after having removed any for which they claim solicitor/client confidentiality, or that are truly irrelevant. The process is governed by formal rules and procedures, giving the parties a level playing field from which to prepare their respective cases for trial.

THE DISCLOSURE JOURNEY

Nowadays documents can be in paper or electronic form, or a mixture – and it can include 'anything in which information of any description is recorded' (part 31.4 of the Civil Procedure Rules). So, in its electronic form, 'document' includes video and audio recordings, mobile phone triangulation data, instant messaging, Bloomberg chat, Morse code, Flash animation: anything in which information is recorded.

For the purposes of this article, we follow a mixture of paper and electronic documents on their journey through the modern disclosure process, followed by some common examples of the variables that will affect its ultimate cost.

The first step is to alert as many staff as appropriate, in both front- and back-office functions, that there is the possibility of a dispute, and instruct them immediately to cease destroying/deleting any documents (this includes breaking the recycling of back-up media). Usually this is deliberately a very broad instruction, as there is unlikely to be much focus at this stage – because the statement of claim has probably not yet been served. As the issues in dispute become clearer, parts of the business can be freed from what can rapidly become an onerous task, especially in the world of email.

To comply with one's disclosure obligations, one must first find out where potentially relevant documents are stored or filed. This involves sharing the likely thrust of the dispute with the key people who may hold, or be responsible for, these potentially relevant documents, to allow them to think about where one might find what is sought. This will usually include a series of interviews with those key people and the IT staff, to draw up a 'map' of where to find everything. All too often, where the business thinks it stores things is entirely different from where IT actually stores them. It also quickly becomes clear that not *all* company policies are followed.

Particularly nowadays, when job security is perhaps not what it used to be, clear communication from the business to the employees about this process is crucial. Without this, some needlessly fear for their future in the face of what is a very comprehensive audit, while others fear for their privacy. Either way, they will act accordingly: being open can allay many fears and reduce overall cost.

Once the information and its location are identified, someone has to go and get it. This is where paper and electronic data diverges in this example (they will come back together again later).

For paper, where possible – and after consultation with the legal team – staff will have first weeded out obviously irrelevant cabinets/boxes/files/bags/piles to avoid needlessly incurring cost. This will usually be based on their intimate knowledge of the matters in dispute. The rest of the documents are securely sent to a scanning and coding company that specialises in this activity within a legal environment (the documents are now legal evidence). 'Scanning' is the process of first separating the paper item (it could be a lab notebook, a document with enclosures, a bound report, etc.) into its individual parts, and then taking an electronic 'picture' of each part (or page) of each paper

item. Following this process, the original paper item can be replicated electronically. As the paper is now a series of pictures, it cannot be sorted or managed in any useful way, so each item is coded. 'Coding' is the manual process of recording basic information (date, author, title, etc) about the item on an associated database record, that can be used for basic searching and sorting of the associated image. It is possible to attempt to add a layer of searchable text onto the pictures by having the computer try to read any typed characters and convert them into electronic words (optical character recognition or OCR), although the success rate of OCR is typically low.

Where electronic documents are concerned, the collection process is obviously different. In this example, we are going to visit the business and borrow key people's (the 'custodians') computers and take a forensically sound and exact copy of their hard drives. In addition, local IT staff will provide any back-up media spanning the relevant period, as well as physical and technical access to all relevant mail and file servers. We shall take forensically sound copies of the mail servers in their entirety. We are doing this because we are still unfocused on the actual matters in dispute, and who all the key people will be: it is foolhardy to restrict collection at this point because that would result in further disruption to the business if we need to return. The price difference can be minimal. For file servers we would endeavour to target just the potentially relevant areas.

All of this electronic data now needs to be unpacked and the chaff removed, so that all that is left is potentially relevant material of use to a human being. This is largely an automatic process, and is the first time that anyone can start to make any useful sense of the numbers of items and the volume of space they take up (these are a couple of the variables that were completely unknown until now). It is also the point from which humans can start to access the information, to begin to understand the supportive or harmful nature of what has been collected.

Before beginning the review, it is common to cull the document population to a set more likely to be relevant, once more is known about the legal focus. Some culling techniques include applying suitable date ranges to the documents, or reducing the final population to only those documents found by searching for certain key terms or phrases. The removal of exactly duplicate documents is also a common way to reduce numbers.

Review is the stage when costs really start to hurt. Prior to actual disclosure, someone usually has to read everything that is likely to be disclosed. So the scanned and coded paper and processed electronic documents are now combined and loaded into a 'review platform' (of which there are many varieties). This is a system that securely allows authorised users to search and sort all of the documents and share thoughts about them. For the purpose of disclosure, this is needed for three main reasons: to test for relevance; to test for privilege; to be aware of what supportive (to them) and harmful (to you) evidence is being provided to the other party.

Once the process has yielded a relevant, unprivileged disclosure set, this is exchanged with the other side, usually as data and images on one or more hard drives. But the game is still not over. One must now bring the jigsaw pieces together by combining each party's disclosure in order to see the bigger picture of what really happened. A review of the incoming disclosure will be needed to get the full facts in context.

While that process may seem simple, there are ample variables that make it hard to provide an accurate estimate up front, regardless of how experienced one is with disclosure. As mentioned earlier, it is not until the data processing stage is complete that one has any idea how many electronic documents there are, and how much space they take up. These very elementary variables have a tremendous impact on cost, yet we must still provide firm estimates quite some time before we have reached this level of awareness of what lies ahead.

The global disclosure industry has a number of ways of charging for what it does in support of the legal profession, and it seems no two disclosure companies do it the same way (which, from experience, makes comparing budgets a nightmare). Estimates are based on assumptions, and the ability to make good assumptions varies widely. The only assumption that is always right is that some assumptions will be wrong. Unfortunately, Precedent H does not really allow for that.

INFLUENCING FACTORS

So what are the factors that can hole you beneath the water line? In the world of electronic documents. There are a number of examples.

Compressed files: One would think that if you have collected data that eventually takes up X amount of space or contains Y amount of items, you would be safe to provide a cost estimate for subsequent activities relating to that data based on the figures that are X and/or Y. But it is common practice to shrink large but infrequently accessed data into smaller units so that they take up less space (the electronic equivalent of using vacuum compressed bags). Similarly, it is equally common to roll up multiple bits of data into a single unit which is smaller than the sum of its contents. In each instance, you need to expand these back into useable units before you can do anything more with them. Thus, the X or Y you collected will be bigger than X and more than Y after processing, because all the supposedly small things are now actual size. But you will not know that until you have processed it all.

Encrypted files: Businesses encrypt data to keep it safe from prying eyes, but during data collection they often forget that the data is encrypted, or are unaware that older files are encrypted. Data processing will either be diverted or halted while the key encryption key is found; and the data will remain inaccessible if the key cannot be found.

> Clients can directly influence the cost of some of the disclosure activities. Generally, the more support they provide the lower the overall costs

Password-protected files: Password-protected files cannot be accessed without their passwords. But unlike encryption, password protection is usually an ad hoc process applied by the creator of that file. What this means is that the file is likely to contain data of importance to the creator but not necessarily of importance to the business. But without seeing it, who knows? If unavailable, passwords can still usually be cracked, but it can be a time-consuming and expensive exercise.

Embedded objects: These are 'hidden' pieces of data that only reveal themselves during processing, and so are not apparent when considering Y amount of items (but could be included in the X amount of space). They are not hidden as such, they are just referred to in a

Continued on page 20

Continued from page 19

way that does not make them immediately apparent in a data 'head count.' Examples are a link to a spreadsheet from within a Word document, or a picture inserted in (rather than attached to) an email.

Non-standard/bespoke software: There is often data that requires some other software in order to read it. One example is an Access database that is effectively unusable without also having the original database structure written specifically to interrogate and report on just that data population. It is impossible to make any sense of the data without that bespoke wrap-around.

Most of us may be comfortable with 'traditional' office software (word processing, spreadsheets, etc) but occasionally a business will have something that is more rarefied, such as a computer-aided design (CAD) tool. If software like that is encountered and it is confirmed that it may be relevant, then traditional workflows and processes will need to be altered or new workflows created in order to accommodate their throughput. Basically, anything out of the norm usually costs more money to manage.

Incomplete data sources: Too often, there are incomplete sets of back-up tapes. It is nearly impossible to effect any form of data restoration without the complete set of tapes to hand. Many hours can be spent in discussion back and forth with a client's IT team trying to find a good set of back-up tapes. One cannot sensibly allow a sum in the estimate for that.

> Last-minute discoveries almost always happen... someone remembers the storage facility down the road

Foreign language: Often, and especially in Europe, one can encounter unexpected tranches of content in a foreign language. Although this has no impact on processing them, it can have a significant time and cost impact on the review. Multi-lingual reviewers are needed, and it is necessary to adapt the review manuals to reflect the different language requirements and terms of art being sought.

In the world of document scanning, there are a few very important elements that have an enormous impact on the cost.

Pages per document: The number of pages scanned and the number of documents coded will drive the final cost. At the point of estimating, one would usually make an assumption about the number of documents (based on the reported number of filing cabinets/drawers/boxes/bags/piles) and number of pages per document. That would lead to an estimate based on multiples of documents and pages. If the ratio proves to be wrong (and it usually is, as there is no universal document size) then the final price will change (usually upwards). Unfortunately, it is impossible to know what the page count per document is until everything has been scanned.

Non-standard size: The cheapest way of scanning is to do it in bulk using automation: drop the pages into the automatic document handler and press 'Go'. This works well for pages of a 'standard' and/or consistent size, and a robust physical quality. But other documents need to be handled manually – which increases cost. Non-standard binding can also create extra expense.

HUMAN ERROR

One should never overlook the extent to which human beings can have an unexpected impact on the overall cost of disclosure.

Custodians: As described earlier, custodians are those who 'own' potentially relevant documents. They may go on to become witnesses, but the terms are not synonymous. Custodians' data will need to be collected once they have been interviewed.

The variables in play here are: the number of custodians (this inevitably increases from the estimated number as the matter develops); their locations; their availability; their helpfulness (c.f. good communications); the amount of data they truly 'own.'

Assumptions about all of these are made when a cost estimate is prepared, but they always change.

Names: This one is often overlooked: sometimes there is no person who owns the documents, just an office, job role or function (such as 'cabinet secretary'). That entity can be represented by any number of people over time, and can often evolve into a multi-headed Hydra during the identification process. Some executives are supported by a PA who may correspond in their boss's names but, behind the electronic scenes, it is actually the PA's name over everything not their boss's (who is the 'real' custodian). It is also not unknown for custodians to change their names (through marriage, divorce or deed poll) during the period in which the matters in dispute arose.

DISCLOSURE COSTS

Planning and preparation: Clients can directly influence the cost of some of the disclosure activities. Generally, the more support they provide, the lower the overall cost. By example, great savings can be shown by having key people co-ordinated so they are available when needed, their computers available to have data collected, facilities available to use, senior managers aware of the processes and why they are necessary, and all privacy and confidentiality issues dealt with up front.

But beware: while a low-cost option where the client intends to take a very active role may make economic sense, it can cause issues if the intent and respective roles are not clearly communicated to everyone involved.

Last-minute discoveries: This almost always happens. While interviewing people and poking around in server and storage rooms, someone inevitably remembers the storage facility down the road that contains heaps of on-point paper documents, or the out-of-commission servers kept stacked in a cupboard that still have potentially relevant data on them. These all add to the final cost.

So, too, does a custodian innocently referring to their Gmail account or a thumb drive, each of which they have used for business purposes. Not to mention the home computer they share with their spouse, a doctor, and which both of them also use for business purposes.

The opposition: Co-operation between the parties is required by the rules, and is a Good Thing if both parties play nicely. But no one can ever allow for the time 'invested' in dealing with a party who thinks that being co-operative is an excuse to waste time 'trying' to agree key words, date ranges, collection methodologies, protocols and so on. Co-operation doesn't half cost a lot if not done properly.

Finally, I have yet to see one party provide the court with an estimate for the cost of analysing and reviewing the other party's disclosure, but that is as necessary an activity as getting one's own disclosure out of the door.

I have yet to see one party provide the court with a cost estimate for analysing the other's party disclosure

This article provides some indications as to why disclosure estimates can change so rapidly through no one's fault. There are no magic solutions to any of these issues; but there are some actions an organisation can take to manage some of the costs to the extent possible.

When a dispute arises, retain an experienced disclosure consultant who can advise regarding the issues and the best workflows and technologies to use based on what is known. Carefully review their pricing structure to determine whether it makes sense for the organisation and the dispute. Use a legal team with an understanding of the disclosure process, and insist on open communications among the component parts of team assembled to manage the disclosure process. Finally, do not be resistant to change: expect to move with the ebb and flow of the shared journey. Taking these steps can give you the lifejacket you need to prevent drowning when you are driven against the inevitable iceberg.

Jonathan Maas is senior director at Huron Legal; www.huronconsultinggroup.com