CORPORATE
FINANCIAL FRAUD

Follow the trail

A recent high-value fraud claim gives an insight into the interrogation of ultra-high volume discovery and has shown how bespoke algorithms can be used for tracing exercises involving complex financial transactions

Technology has enabled 30 years of financial transactions and a vast, chaotic and un-coded dataset to be ordered, interrogated, and ultimately used to disprove a claimant's assertions to be the innocent victims of fraudulent business practices and defeat their attempts to trace purportedly misappropriated funds into the hands of the defendants.

On May 31 2018, the Grand Court of the Cayman Islands found in favour of the liquidators of the defendant companies in the Islands' longest ever trial in the matter of *Ahmad Hamad Algosaibi & Brothers Company (AHAB) v Saad Investment Company Limited and Others*. At 129 days, it is the Cayman Islands' longest-running trial and one of the longest fraud trials ever litigated worldwide.

AHAB claimed approximately $9 billion in damages. The Grand Court found, however, that AHAB's business practices were not only dishonest but that the AHAB Partners were complicit in a $330 billion Ponzi scheme, defrauding international lending banks over decades and specifically, in the period from 1980 until the global economic crash in 2008/9. In addition to being one of the most high-value fraud claims ever litigated, the sheer scale and complexity of the discovery exercise involving both documentary and financial data also make it one of largest fraud trials ever.

The history of AHAB and its claims

AHAB has its origins in a business began by Hamad Algosaibi in the 1940s. Hamad died in 1969 and was succeeded by his three sons, Ahmad, Abdulaziz and Suleiman. Together, they incorporated AHAB as a general partnership and successively chaired AHAB until Suleiman's death in February 2009. AHAB has since been chaired by Yousef, Ahmad's son. Yousef, Saud (Abdulaziz's son) and Dawood (Suleiman's son) are among the current AHAB Partners.

From 1980 onwards, AHAB strategically expanded into financial services and other related businesses. In 1981, its board of directors established the Money Exchange as an unincorporated division of AHAB. Maan Al Sanea, who had married Abdulaziz's daughter in

1 MINUTE READ

The Grand Court of the Cayman Islands has ruled against AHAB in one of the largest fraud cases ever litigated. The Chief Justice found that, far from being a victim of fraud (as AHAB alleged), they had been complicit in a $330 billion Ponzi Scheme defrauding international lending since the early 1980s.
The defendants' collaboration with analytics and discovery partners and their innovative use of technology allowed them to uncover these practices in a vast and chaotic document population; and disprove claims of misappropriation by means of an automated Tracing Tool which was able to cut through a transactional data base formed of over 30 years' worth of statements.

1980, was appointed its managing director. In the 1980s, AHAB incorporated Algosaibi Investment Holdings EC (AIH) and Algosaibi Trading Services (ATS) (originally incorporated as Algosaibi Investment Services (AIS)) and in 2003, AHAB incorporated a bank in Bahrain, The International Banking Corporation (TIBC) (collectively, the Financial Businesses).

From near the time of the establishment of the Money Exchange until its collapse in May 2009, financial statements, disseminated to in excess of 100 lending banks, understated the extent of the borrowings and true extent of AHAB indebtedness to the banks and its status as a borrower. By presenting them to the banks, the false financial statements became the central instrumentality of a fraud.

In 2009, AHAB defaulted on more than SAR 34 billion of debt (approximately $9.2 billion). Shortly after that default AHAB commenced proceedings against Al Sanea and the corporate defendants (established by Al Sanea in the Cayman Islands and currently in liquidation).

AHAB's claims were for alleged fraudulent breaches of fiduciary duties committed by Al Sanea and restitution, damages and compensation from the defendants – on the basis of their conspiracy with Al Sanea, their knowing assistance in his alleged fraud upon AHAB and ultimately, their knowing receipt of the massive proceeds of that fraud. AHAB also brought proprietary claims against the defendants on the basis that their assets represent AHAB's property – the proceeds of the fraud – which AHAB could trace into their bank accounts or other assets.

The challenges that the defendants faced

In disproving AHAB's case, the defendant's legal teams faced three specific challenges:

- first, a vast, chaotic and uncoded dataset (numbering 2.5 million documents, including documents that were typed, in manuscript and a combination of the two, and documents that were in English and Arabic) from which to try to establish that the AHAB partners knew about the huge borrowings incurred in AHAB's name;
- second, the 1.4 million rows of transactional data across 30 years of trading from which to establish that funds allegedly misappropriated from AHAB could not be traced into the accounts of the defendants (including which tracing rules were applicable, and whether such rules or equivalent principles existed in various non-common law jurisdictions in which accounts were located); and
- third, that, during the course of proceedings, AHAB changed its case no fewer than seven times, including on the second day of trial. Each time the defendants put forward evidence to disprove AHAB's case theory, the case theory would change, requiring the exercise to be undertaken again.

Dealing with high-volume raw data during discovery

Electronic discovery platforms are not new. Most practitioners will know that even with comparatively small datasets, the ability to sort, filter, search and code by multiple users simultaneously is cost-effective and time efficient. With a data set over 2.5 million documents (which comprised some documents that were originally in manuscript, some that were in Arabic, and all of which were provided without search functionality or having been pre-sifted for relevance), the platform (which, in this case, was Ringtail – FTI's proprietary platform) was essential. It simply was not feasible or proportionate to page turn every document. The tight trial timetable that the parties were working to made it impracticable in any event.

The defendant's legal teams faced a vast, chaotic and uncoded dataset numbering 2.5 million documents

The first step, therefore, was to structure and code the dataset to make it searchable. To do this, the legal team first had to identify the important characteristics. In this instance, the date, origin (including physical location), author and the extent of foreign language/manuscript content were all vital field codes.

Particularly vital in this case was the ability to remove exact duplicate documents, without the risk that ostensibly duplicate documents, but which emanated from different locations, would be discarded. Typically, the process of de-duplication in a discovery exercise can be a welcome means of reducing document numbers, in this case each duplicate told a different part of the story and its retention was crucial.

This was made possible by location coding, which was one of the very few pieces coding fields with which the documents were provided. It was ultimately the location coding of the documents that enabled the defendants to piece together a true and complete narrative and demonstrate that the claimants knew of, and authorised, the practices of which they subsequently claimed to be unwitting victims. It was only with knowledge of the provenance of each copy a document that the defendants could assert with confidence that it bore the fingerprints of an AHAB partner and, in turn, that that partner was imputed with the necessary knowledge.

The chaotic manner in which the vast document population was gathered and subsequently disclosed by the claimant, however, meant that the system and significance of the location coding was not immediately apparent and decipherable.

For the small volume of electronic data, the provenance was relatively clear by virtue of its inherent meta data. The vast majority of discovery, however, (almost two million documents) was in hard copy form and much was in Arabic (thus requiring translation). This was in keeping with the somewhat old fashioned and paper heavy way in which the AHAB businesses were run. Those documents were located, reviewed, catalogued and scanned over a period of at least two years by the claimant and its representatives, and allocated a code according to its provenance. Curiously, documents of the same provenance (ie from the same building, office, cupboard and file) were rarely disclosed by the claimant together, and often not even within the same tranche of discovery (of which there were in excess of 40). In short, the integrity of particular files was not retained so a process of reconstruction of files was required. This was possible by means of mining the document population not for particular document content but for specific coding,

and then pooling documents with that coding in common to recreate individual files, which resembled as closely as possible the physical files traceable to a particular location. Only then could the defendants properly understand the type and extent of documentation which the partners and their staff received, considered, and retained.

Once the dataset was coded and searchable, the next step was to use the latest available data analytics tools to develop and test the prevailing case theory. This was an iterative process, with both the case theory evolving according to the search results, and our search techniques becoming more refined (and ultimately including bespoke algorithms).

For example, seeding enables users to identify similar documents based on a key phrase or text structure. This is helpful for identifying groups of similar documents that vary slightly in format over time – like board minutes, financial accounts or other corporate documents, or further examples of a particular type of document that the partners had taken an interest in.

Clustering enables users to see which documents regularly appear together or are connected in the dataset, whether by common document content or coding. The review teams were able to consider the results of a cluster exercise and, in identifying from those results the documents which are the most fit for purpose, to re-focus and refine the system's subsequent attempts. It is an iterative process for both user and system. It enabled to uncover and better understand patterns of communication, in terms of which individuals were in regular contact, with what frequency, and on what topics. This again was important in establishing the level of knowledge of the AHAB partners.

In any case where knowledge of a party and specific individuals is relevant, access to this level of sophistication of data interrogation and manipulation is incredibly valuable. The techniques as described above ensured that the defendants could achieve a deeper understanding of the workings and practices of the AHAB business and the protagonists' particular level of knowledge and understanding of, and complicity in those practices.

Following the money – technology and tracing

In order to defeat AHAB's proprietary tracing claims it was critical for the defendants to be able to provide to the Court a robust and defensible explanation of the money flows.

FTI Consulting's Data & Analytics practice was instructed to work alongside RSM (one of the joint official liquidators) to facilitate analysis of the banking system data. The resulting web-based tool enabled their review with a rich set of analysis and visualisation technology. Further success by the team in bringing order to a messy million-plus record set of bank statement data led legal advisor Charles Russell Speechlys to enquire: could analysis of the claimant's proprietary claim be automated?

The highest purpose of data analytics is for thoughtful interrogation of information to enable and elevate expert insight. The combination of transaction volume with variety and complexity rendered analysis by accountants alone ineffective.

The team had the benefit of two rich sources of data representing the relevant financial transactions. The first, referred to as the transactional database (TD), exceeded 1.4 million rows of transactional information gleaned from over 10 years of PDF bank statements records produced by parties to the litigation and other relevant entities. The second, a banking ledger dataset, was essential to automating analysis of the proprietary claim and the creation of a tracing tool (TT).

A bank-validated transaction record set is often preferred for tracing. However, the TD suffered from several real-world limitations meaning that automated tracing would have been ineffective.

The second information source, the ledger, readily lent itself to systematic analysis of the universe of transactions. By virtue of being a cohesive double-entry recording system, the ledger met the pre-requisites of a data source for algorithmic analysis: 1) all in-flows and out-flows for each account were represented; 2) there were no conflicts in the ordering of transactions; and 3) the consistent identification of counterparty accounts enabled the automated mapping of out-flows in one account to in-flows in another.

To analyse AHAB's proprietary claim, FTI developed the TT in Python, a well-established programming or scripting language. Python is convenient and commonly used, facilitating team development and peer review. The legal team prepared the detailed objectives and established the tracing rules which required a first in, first out (FIFO) treatment of transactions. The tracing rules further specify treatment of overdraft situations. The TT applies a claimant-friendly transaction ordering approach (within a day, all in flows are considered to occur before any outflows) to minimise defendant-friendly conclusions of trace-ending overdrafts. In addition to this being the most credible methodology, it explicitly reinforced to the Court the independence of the data analytics expert's approach.

Refinement of the Python algorithms to create the Tracing Tool was performed in expert consultation with RSM whose deep familiarity with the ledger data set and accounting expertise enabled coding for the variety of transactions therein. The tool's modular design allowed for nuanced treatment of transaction types including: foreign currency exchanges; equity (share)

Figure 1: simple FIFO matching of inflows with multiple outflows

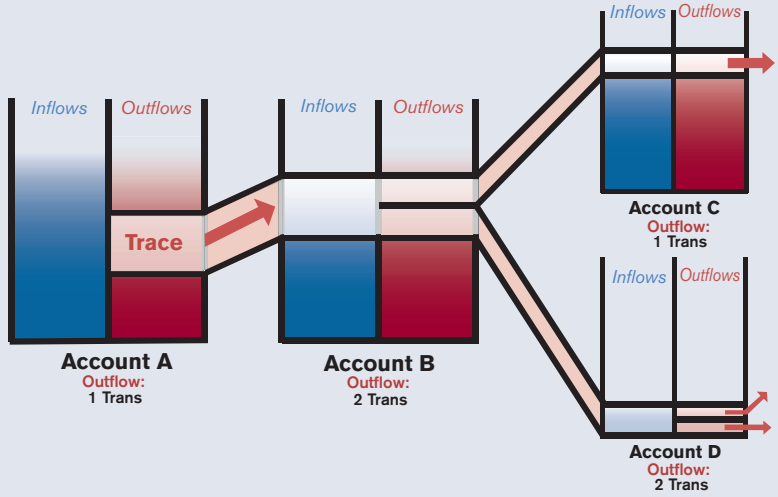
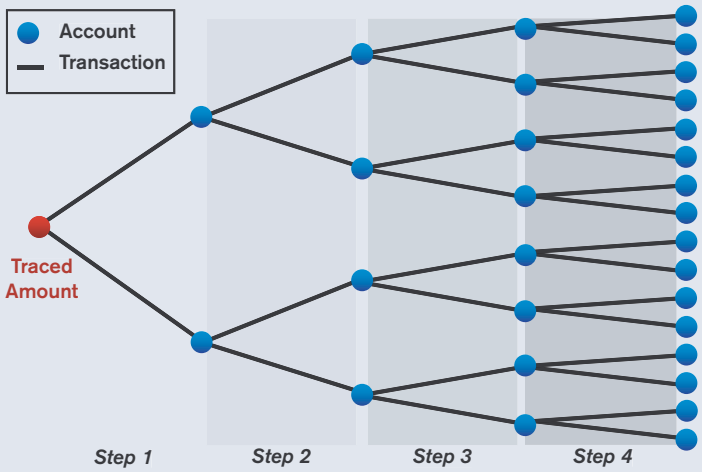


Figure 2: the single trace rapidly fragments into many different parts



transactions; and the frequent occurrence whereby a FIFO analysis identifies that a single outflow is related to multiple inflows requiring separate tracing pathways.

The TT was designed for a so-called reverse trace whereby the specific transactions alleged by claimants to support their proprietary claim were traced backwards to the original sources of funds. For simplicity's sake, figure 1 illustrates the bifurcation for a forward trace.

This nearly continuous bifurcation spread over several years, renders a manual tracing exercise inefficient for humans; this analysis would be nearly impossible in a spreadsheet-type tool. The majority of traces performed involve sequences of dozens of transactions whose bifurcations result in dozens of end points. The algorithm is designed to recognise each bifurcation as the creation of two new distinct trace pathways that need to be discretely sourced all the way from the original subject transaction through to the ultimate source of funds (see figure 2).

The TT is not a black box whereby a button push generates an answer. It was designed to enable the expertise of humans, not to override it. In fact, rather than attempting to engineer nuance for every situation, the tool was instead designed with stopping conditions whereby the algorithmic processing would halt, enable expert human intervention, and then return the recursive processing to the realm of computers.

An example of a stopping condition would be with equity sales: an offline process evaluates the appropriate allocation by referencing an external share dealing record. Once the amount is pro-rated (to account for value appreciation) through the equity account, the algorithmic trace would continue.

Criticisms of this analysis include that the ledger is not limited to cash transactions, and can incorporate accounting entries such as reversals. This is a reasonable criticism – taken in proportion. Reversals are also found among some cash transactions from bank statements in the TD. The extent to which reversal entries have any impact is entirely dependent on the account specifics.

Importantly, these two data sources were used to validate the analysis: ledger endpoints (sources of funds in the reverse traces) determined by the TT were reconciled to the TD. Bank statement evidence supports the substantial majority of ledger endpoints. For presentation to the Court, code was written to visualise the vast numbers of trace steps and end points.

While development of algorithmic models such as the TT enable delivery of important expert evidence, where technology is closely relied upon, rigour and a commitment to transparent declaration of assumptions is essential for the Court to appreciate the expert's independence.

Key takeaways

- Huge, chaotic datasets are not impenetrable and no longer require time-consuming and costly page-turn review. The use of technology early in the process can accelerate a merits analysis and allow decisions as to case management strategy, including offers to settle, to be made earlier.
- The potential to differentiate client cases by use of technology is substantial when advanced analytics are engaged in complex matters. In any case where assets have been purportedly misappropriated and dispersed through numerous channels, or where the knowledge of a party is in issue, technology can provide you and the court with clarity and certainty.
- The courts are increasingly receptive to technology. The Cayman Islands Court wholeheartedly embraced the use of cutting edge technology throughout (including entirely new, bespoke technology in this case).
- Cases that were previously considered commercially not viable to proceed to trial (due to the costs of disclosure or the inability to interrogate a dataset manually) may now, in fact, be brought to a successful resolution provided the legal team has the relevant technological experience and enthusiasm.
- The legal team must have the strategic vision to recognise what's possible. Analytics partners must be experienced with current technologies, and have the agility to select and develop in a platform that is suited to both the data and the analytical challenge. The collaboration model therefore requires a cross-disciplinary team where all participants have at least a basic appreciation for the others' capability and a shared vision of the ultimate objectives.

Charles Russell Speechlys advised Bruce Mackay, Geoff Carton-Kelly, Chris Johnson and Russell Homer – the Joint Official Liquidators of the six Awalco Defendants.

Charlotte Pender
Partner
Charles Russell Speechlys (London)

Nick Hourigan
Senior managing director, data & analytics
FTI Consulting (London)