JS|HELD

PERSPECTIVES

Integrating Data Analytics into a Financial Investigation

Our perspectives feature the viewpoints of our subject matter experts on current topics and emerging trends.

INTRODUCTION

Structured data and information systems are increasingly utilized to improve the efficiency and efficacy of financial investigations. Forensic accountants and investigation experts have more frequently engaged skilled analysts to perform certain data extraction and complex technical analyses. However, certain matters force a bifurcated approach where the analytics team is left struggling to understand which results to produce, and the accounting team is left with insights that may not effectively move the investigation forward. While accounting and investigation experts have become more familiar with data analytics capabilities (i.e. what is possible) through a generation of integrated investigative experience, technical and financial expertise can appear as two different languages frequently translated back and forth. This often leads to inefficiencies and misinterpretations along the way. By fusing the languages of analytics and accounting together in a unified voice, the results are undeniable.

The following case study walks through the process of a financial investigation in which the data extraction, analysis, and forensic reconciliation are integrated seamlessly by accounting and data experts who understand both sides of an investigation.

SAMPLE CASE: BACKGROUND

This case study will navigate the investigation of a fictitious retail firm, ABC, Inc. The case will follow the key steps of a financial investigation and address the advantages of a data-centric approach at each stage.

Like many firms, ABC's online accounting system lacked any set procurement, HR, or A/P processes, and the small accounting team operated under minimal oversight and structure. While such conditions aren't rare, they leave firms vulnerable to fraudulent activity. In ABC's case, fraud went unnoticed for years, with suspicions arising only when discrepancies or irregularities were discovered by an inquisitive vendor seeking Accounts Receivable confirmations from their long-time customer, ABC.

Upon the realization of potential improprieties, ABC hired a third-party investigations firm to determine the source, as well as the depth and breadth of the fraud. In such cases where there is minimal structure and supervision of the accounting activities, a data-forward approach proves especially powerful as it utilizes technology to maximize the value of available information provided by the client company.

DATA COLLECTION

Identifying and extracting source data from an accounting system is typically one of the first – and most challenging – steps in an investigation. Although QuickBooks is highlighted in this case study, similar processes exist for most accounting information systems, and a "back door" is generally available when the data is structured in a way that allows for a straightforward import into analytical software such as SQL or Python.

ABC provided the investigating firm with four years of QuickBooks backup files to allow access to historical accounting entries dating back to when they suspected this fraud began. Each year contained several million records of accounting entries, as well as customer, vendor, and employee details. While popular accounting systems offer user-friendly front-end capabilities, they can be finnicky and troublesome to work with on the investigation side. Traditionally, forensic accountants may spend hours piecing together fractions of the QuickBooks data through QuickBooks reports in search of significant outputs. However, the investigating firm hired by ABC uses forensic data analysis tools to extract data from the backup files in just a few clicks. Within hours of receiving the data, the investigators were able to extract all underlying data sets from QuickBooks such as vendor lists, the chart of accounts, general ledger data, audit logs, and more.

METADATA ANALYSIS

Before performing any analysis on the records themselves, investigators can find clues in the metadata of the files received from the subject of

the investigation. The term 'metadata' refers to data points about a file or data set, such as the name, size, and date of creation. The investigators in this case analyzed the metadata of the QuickBooks backups from ABC. This quickly revealed several red flags.

Exhibit A: QuickBooks Backup Metadata

Backup Date	File Size	Generated By
2021	300 MB	Bookkeeper
2022	350 MB	Bookkeeper
2023	400 MB	Bookkeeper
2024	350 MB	Controller

One would expect the backup to increase in size every year as more transactions occur. Such is not the case in 2024 – the size decreases, which suggests that transactions may have been deleted. Furthermore, the backup was generated by a user who is not typically involved in the backup process. Therefore, the metadata indicates that the controller may be an accomplice to the suspected misappropriation of funds.

This straightforward example of metadata analysis demonstrates how clues hidden within data files themselves can help guide an investigation in the right direction. However, while metadata analysis can provide valuable information and can clue investigators on where to dig deeper, it does not singlehandedly prove wrongdoing. Definitive proof is uncovered during the next steps of the investigation: in-depth analysis. When additional discernment is required, a competent computer forensics team can use their suite of tools and expertise to perform detailed analyses of files.

QUICKBOOKS DATA ANALYSIS

Investigators will often begin with a comparative analysis of General Ledger (G/L) data within the accounting system – in this case, QuickBooks backup files. While necessary, this process can be cumbersome and slow in Excel, especially when working with large data files or a large quantity of files. However, many database analytics tools can perform this comparative analysis between the data files quickly and efficiently. With a few lines of SQL code, the investigators can perform a JOIN analysis across QuickBooks G/L transaction tables to reveal which rows existed in previous backups but were absent from the latest backup.

To perform the JOIN analysis, the investigators will first identify the data's "Primary Key." A primary key is a column of a data set which serves as a unique identifier for each row. Although any of the relevant information such as debits, credits, vendor, date, etc. can be changed, the Primary Key never changes. As with most enterprise resource planning (ERP) systems, QuickBooks maintains Primary Keys that are only visible when using an extraction tool such as QODBC and not when reports are run within the QuickBooks application itself.

Using this Primary Key as the basis of a JOIN analysis, the investigators can easily match transactions between files, even if underlying values were modified. In a case such as the ABC investigation, where there exists concern about transactions being modified or deleted, investigators can set criteria to only return instances where values differed between the four backup files. In a matter of seconds, the JOIN statement identified nearly 500 modified or deleted transactions from ABC's historical records. A particular transaction, JE #150, immediately stuck out:

Exhibit B: Sample of SQL JOIN

Backup Date	Transaction ID	Transaction Date	Vendor	G/L Account	Amount
2021	JE #150	6/30/2021	Supplier A	COGS	$5,450
2022	JE #150	6/30/2021	Supplier A	COGS	$5,450
2023	JE #150	6/30/2021	Supplier A	COGS	$5,450
2024	JE #150	6/30/2021	Supplier B	COGS	$5,450

As shown above, an entry booked in 2021 had its vendor changed sometime between the creation of the 2023 and 2024 backup file. This activity suggests that COGS may be an account where suspicious activity is being booked. In addition, the change from Supplier A to Supplier B may suggest that fraudulent activity is being booked under Supplier A. The investigators therefore lend focus to Supplier A and the COGS account. Because they had already extracted all QuickBooks data into an easily accessible format, within minutes the investigators can pull up a list of all transactions in each backup file associated with Supplier A.

Using a GROUP BY statement, investigators can quickly extract the amount paid to Supplier A across all backup files for analysis.

Backup File	Transaction Year	Vendor	Total Amount
2021	2021	Supplier A	$500,000
2022	2021	Supplier A	$500,000
2023	2021	Supplier A	$500,000
2024	2021	Supplier A	$28,000

As shown above, it appears that sometime between the 2023 and 2024 backup, a considerable number of payments were reclassified from Supplier A or deleted for transactions booked in 2021. As the data was already accessible in the investigators' SQL database, investigators were able to quickly pull all the transactions altered in the 2024 backup file.

Within hours, investigators performed a complete comparative analysis between four different QuickBooks files and extracted a complete listing of these differences. Looking at this data, investigators identified a suspicious vendor and an account where fraudulent activity could be hidden. Without approaching investigations from a data standpoint, these findings could take days given the large number of files and transactions in QuickBooks.

CONCLUSION

Centering data and technology in financial investigations can mitigate many of the common challenges that investigators face, and it can result in a more prompt and efficient delivery of results to clients.

While the ABC case is fictitious, it is based on real cases that forensic accountants work to solve. Although many investigations can be done without leveraging advanced technological tools, learning these technologies and applying them to engagements can help investigators uncover and protect against fraud. It is in the best interests of forensic accountants to embrace the technical tools of an investigation.

ACKNOWLEDGMENTS

We would like to thank our colleagues, Ken Feinstein, Matthew Cordell, Hannah Siegel, and Shane Jaeger for their insights and expertise that greatly assisted this research.

Ken Feinstein is a Senior Managing Director in the Digital Investigations & Discovery service line within the Global Investigations practice at J.S. Held. He specializes in investigative data analytics and provides investigations, regulatory risk and litigation support solutions spanning multiple sectors, including retail and consumer products, life sciences, technology, financial services, industrial products, and government agencies. His clients include law firms and Fortune 500 legal and compliance teams for whom he delivers large scale, complex investigations, regulatory response matters, proactive anti-fraud efforts, and compliance programs. He is a member of the American Institute of Certified Public Accountants and the Association of Certified Fraud Examiners.

Ken can be reached at ken.feinstein@jsheld.com or +1 917 277 7868.

Matthew Cordell is a Senior Director in the Digital Investigations & Discovery service line within the Global Investigations practice at J.S. Held. He helps analyze structured and unstructured data along with other forensic accounting and investigative methods to help clients achieve their goals in reducing fraud and regulatory risk. His work includes data modeling and analysis, database management, programming, and implementing finance software spanning multiple industries, including financial services, technology, industrials, life sciences, and retail. His clients include law firms and Fortune 500 multinational companies. He is a member of the American Institute of Certified Public Accountants and the Association of Certified Fraud Examiners.

Matthew can be reached at matthew.cordell@jsheld.com or +1 214 216 4960.

Hannah Siegel is a Senior Consultant in the Digital Investigations & Discovery service line within the Global Investigations practice at J.S. Held. Hannah leverages her expertise in data analytics and collaboration to deliver innovative solutions to clients. She contributes to a diverse range of projects across multiple business groups, undertaking a wide variety of tasks beyond traditional data analysis.

Hannah can be reached at hannah.siegel@jsheld.com or +1 646 551 4066.

Shane Jaeger is a Consultant in the Digital Investigations & Discovery service line within the Global Investigations practice at J.S. Held. She joined the firm in September 2024. Shane utilizes analytical tools and skills to derive history and meaning from large datasets.

Shane can be reached at shane.jaeger@jsheld.com or +1 203 505 9474.