LJN LAW JOURNAL NEWSLETTERS

LJN's LEGAL TECH Newsletter®

An incisivemedia publication

Volume 27, Number 2 • May 2009

e-Discovery of Structured Data

When e-Mail Won't Suffice

By Erik Post and Jim Vint

How did an event happen? What proof do you have? Further, what proof do you need?

At the outset of an investigation, timelines and custodian-created information assets can all be garnered from extracts of accounting data, e-mails and documents through traditional discovery and disclosure. None of this tells you how, specifically, an event or series of events occurred. Were the alleged claims the result of system control errors, malicious employee action, fraud or simple misrepresentations by management as a means of covering up systemic losses by cooking the books to avoid audit concerns?

Finding Out How

The investigations into a number of high profile scandals that have begun in recent months highlight this exact issue. Fraud, accounting scandals and system control circumnavigation have plagued the economy over the past few quarters. While Bernie Madoff and Satyam Computer Services exemplify the more recent and prominent scandals, there are countless investigations that have arisen due to the economic conditions where understanding the means by which events unfolded is the primary focus for all the parties of interest. The only way to do this is by tracing the problem through financial systems, trading systems, e-mails and other bank records. e-Mail alone is no longer enough to solve the problem. So what do the other systems contain? Enterprise systems are created for the centralization of core business functions such as accounting, financial reporting, trading systems, customer relationship management and resource planning. The majority of the time these systems are intertwined and transfer data between them in one of two ways:

1. Through manual processes of extracting from one system and uploading into another, or;
2. Through automated processes of data migration and data workflow processing.

Regardless of how enterprise class or specialized computing systems interact, the number of facets of data analyses and volumes of calculations that occur during these "transfer" periods is of staggering proportion. For example, portions of invoices are split into subledgers and data is transformed and migrated into information for use by the business, investors, regulators and auditors. External market feeds from suppliers, vendors or financial institutions are uploaded along the way, further complicating and convoluting the data flow and sources. Dissecting these systems and rebuilding focused subsets or entire processes allow investigators to gain valuable insight into the details behind the spreadsheets and extracts contained within custodian information and file shares and can set the foundation for asset tracing and third-party reconciliation.

Unfortunately, enterprise systems are traditionally expansive, complicated and often antiquated. Due to the ebb and flow related to merger and acquisition activity over time, it is not uncommon to find multiple implementations of like-kind systems spanning multiple software packages. For example, recent mergers between financial institutions have resulted in redundant back-end systems from the legacy institutions for accounting, human resources reporting and supportive business function. While customers may see a unified interface — such as one online account access point — log-in identities and account numbers direct users to specific systems for retrieving account information, trade history and paperless statements.

It is certainly no easy task to demonstrate the functionality of these systems for regulators, oversight committees, or the courts — even sub-processes or components of these systems can be extremely complex. Imagine the consolidation of information required for recent mergers in the life sciences space from a financial and accounting perspective as well as a clinical trial and qualitative data standpoint. Gaining insight doesn't become any easier if and when system assets are seized or required to be produced as a part of an enforcement action or other discovery demand.

'Re-discovery'

e-Discovery of structured data has evolved to include the recreation/replication of these complicated systems ("re-discovery"). These mirror images of enterprise systems are being used as demonstratives and are built on common software platforms (such as SQL Server). This makes them easily hosted for review, shared in their entirety or passed along to other interested parties. Utilizing the code originally used to develop the systems in conjunction with extractions of pertinent data and institutional knowledge, these systems or fragments of systems can be rebuilt. As demonstratives, the mirrors provide parties with insight as to *how* trades were placed, embezzlement fraud occurred or system controls were circumnavigated in a workflow and by whom.

For example, a recent investigation into a financial services company required a rebuild of various systems and/or components of systems for a variety of reasons, the most prevalent of which was to understand how specific misrepresentations were made and who would have known of these actions.

Erik Post (Erik.Post@FTIConsulting.com) is a senior managing director in the FTI Forensic and Litigation Consulting practice, where he specializes in risk advisory services. **Jim Vint** (Jim.Vint@FTIConsulting.com), a Managing Director in the FTI Technology Consulting practice, has over 10 years' experience in the areas of information technology and financial systems development, management and implementation. The views expressed in the article are held by the authors and are not necessarily representative of FTI Consulting, Inc.

Given the effective "finger-pointing" of senior members of the organization and limited insight provided by the e-mail and other communication media, due to internal policies on e-mail usage, the financial and transactional systems were the primary focal point.

Why not go to the original system? Why not simply freeze the databases required and use those? Why does a "rebuild" need to occur?

One reason why a rebuild would most likely be required is that archaic and antiquated systems — especially mainframes — are not user-friendly and can be extremely difficult to interpret. (*See*, http://fuel-efficient-vehicles.org/pwsdb/pgm/RPG-ILE-AS400.php.)

However, when put into a relational database, where table constructs can depict events and processes and relationships are more clearly defined (usually through the system schema), the database can be displayed for understanding how the customer number is entered into a search form and the customer record (*i.e.*, all key data about the customer) is returned.

The advantages of using a relational database can be seen for financial statement construction, cash applications or disbursements or even trading systems. Regardless, the end result is a simplified exemplar of a complicated process.

Where can this lead the investigative team?

Take a hypothetical situation in which an executive at a company is masterminding a theft or skimming funds from his or her financial advisory firm. From the e-mails and documents collected, it appears the actions were of a single individual, effectively laundering cash through an off balance sheet vehicle and collecting "commissions" or "administrative fees" via a shell company bank account. In certain instances, transaction tracing would allow investigators to follow the flow of cash. Understanding the processes within an organization and reviewing the code, specifically the internal risk controls, can assist in a number of ways. Among them, it serves to:

- Trace cash outflow to determine if deposits were laundered through the any type of account;
- Track payment mechanisms that allocated cash amounts to the shell bank account;
- Uncover additional vehicles or accounts; and
- Identify and expand the custodian list.

Traditional structured data e-discovery would allow counsel to determine who approved transactions, who entered them and who may have modified them and when. This information is evident in most general ledger entry programs and resides throughout the system in various tables. Additionally, systems are created with the requirements to have risk controls and checks and balances in order to safeguard against malicious intent against a company by an insider. That said, many of these safeguards require personnel to approve or deny suspicious activity — whether that be payables, receivables or even trades. For instance, these work flow protocols can be based on expenditure size when dealing with accounts payable or lot size or volume for trading blocks.

It is often evident that an approver of specific transactions (*i.e.*, cash disbursements) is the very person of interest in the case. By understanding the system code, it is possible to demonstrate the work flow processes that pushed cash approvals to the desk of the key custodian(s), or the process through which key custodians were able to gain access to various parts of a system to facilitate alleged misappropriations.

Gather All Information

So what does this all mean when confronted with a new client matter?

It means that you need to think about the type of case with which you are confronted and scope out the non–e-mail systems that will be of issue at the outset of the case. Most of us have become familiar with the e-discovery methodology contained in the Electronic Discovery Reference Model ("EDRM"; www.edrm.net). However, when confronted with non–e-mail type of systems, modifications to this well-established process need to be considered and understood. The following are several of the concepts that change and therefore help to reshape the strategy and potential focus of information gathering and analysis that are required to solve the problem at hand.

The first is in the **preservations and collections** phase. It is within this stage of the EDRM that group shares for technical specifications, system schema and IT program files are being captured. This serves to ensure that preservation of these documents, program scripts and other system data-flow diagrams occurs. Provided re-discovery is deemed an appropriate portion of the solution, these pieces of information will be extremely relevant to the technology experts and will be required throughout the process. Additionally, in accordance with structured data preservation efforts, it is important to obtain backups surrrounding the date in question (or as close to this date as possible) to be loaded into the reconstructed platform.

The second is within the **processing/review and analysis** phase. Review of the code will likely need to be performed by IT experts at the direction of counsel. Since most code is structured (at least in and of itself), it can be deciphered using various resource keys and searched to create mappings of data flow and determine interrelations between various functions and tables. During this phase, the database administrators, IT support staff and users of these systems will need to be interviewed regarding the setup, usage and customization that was built into those applications. Without this knowledge it will be difficult to trace the problems.

The last concept is within the **production** phase. This is where the bulk of time on re-discovery efforts is spent and, as mentioned earlier, is driven by the scope and magnitude of the investigation. The production of structured data to other parties is still a nascent science and the standards are in development. However, .xml files, .sql tables or .dmp files are rapidly becoming the accepted norm as these files all preserve the table relationships that are required to understand the information contained therein.

Conclusion

The e-discovery industry will continue to adapt in order to accommodate the rising need for producing complex information in simplistic, efficient and effective solutions. Investigatory needs combined with the increase in sophistication of white collar crime and the ever growing complexity of the technology medium upon which business operates have pushed solutions to this point. As we adapt to the increasing demands for information from enterprise systems, more innovative and creative solutions will be required. The problems of today and the future are going to be solved by reading e-mails and by identifying how information was hidden, diverted or actually circumvented program controls within an organization. Therefore, the structured data systems are an important and necessary component of the data review, analysis and production efforts of any matter.