The Legal Intelligencer

Law Topics ⌄   Surveys & Rankings ⌄   Cases & Verdicts ⌄   People & Community ⌄

Public Notice & Classifieds ⌄   All Sections ⌄

ALM | LAW.COM

Page Printed From:

NOT FOR REPRINT

COMMENTARY

Newer Data Sources Present Opportunities and Challenges for Litigation Support

Today, most documents—even if initially received in physical, hard-copy format—are stored electronically with sophisticated electronic document storage warehouses, which provide the ability to easily search document text and metadata. Despite these advancements in technology, the document review process may still be time-consuming.

October 23, 2023 at 12:56 PM

Litigation

By David H. Glusman, Neil Young and Jesse J. LaGrossa | October 23, 2023 at 12:56 PM

Litigators and litigation support professionals who have been practicing long enough recall the days when Bankers Boxes filled with original documents (or copies of originals) were the primary source of "data" in litigation. For such professionals, the tedious and time-consuming task of identifying relevant information in such documents was, and still is, a challenge in commercial and other categories of litigation. Today, most documents—even if initially received in physical, hard-copy format—are stored electronically with sophisticated electronic document storage warehouses, which provide the ability to easily search document text and metadata. Despite these advancements in technology, the document review process may still be time-consuming.

However, in addition to "traditional" hard-copy and electronic document review, newer categories of digital data and the expanded use of such data are creating new challenges—as well as opportunities—for litigators and litigation support professionals alike. As we progress further into the "big data" culture, the current focus appears to be on buzzwords (i.e., artificial intelligence), cloud-based software as a service (SaaS) applications that synchronize data from various sources, and new data analytical tools. This focus undoubtedly has a wide range of applicability for executive decision-making and management of business operations. However, the large volumes of data maintained by businesses and their SaaS providers may undoubtedly be used by attorneys and their experts in litigation.

Categories and Sources of Data

Aside from the financial transaction data obtainable from a company's general ledger accounting system, there are other broad categories and sources of [relatively] newer structured and unstructured data with potential uses in commercial civil and white-collar criminal litigation. Such categories, which are more operational versus financial in nature, include metadata (i.e., data about data), operations data (i.e., time and expense systems, building access logs), and task/industry (specific software application data).

Metadata—while not necessarily a "new" source, document metadata includes fields such as date/time stamps, GPS location information, user identification, software version information, and other fields. Such fields can be used as a cross-reference tool to overlay against other data sources, providing insight into events, timelines and individuals.

Operational data—in today's society, it seems there is little an individual can do in their day-to-day lives without leaving a proverbial digital "breadcrumb." Consider throughout the day how many times one swipes a building access card, opens an application on their work laptop or cellphone, interacts with a critical business software product, etc.

Task/industry-specific software application data—business operations software applications perform a wide range of industry-specific or task-specific functions that automate or streamline tasks. This includes applications for self-driving vehicles, medical monitoring devices, technology infrastructure monitoring, robotics automation, and business management applications (i.e., customer relationship management), among others. Such applications may use, produce, and store vast amounts of "real-time" data. If events occur that give rise to potential litigation, such "real-time" data can be overlayed with metadata, financial data and other operational data.

Discovery, Data Production and Data Management

Expanded data requires attorneys and litigation support experts to work together from the initial discovery phase to the trial phase. Increasingly, litigation support teams may comprise a digital forensics expert who may obtain and document custody of the data in coordination with one or more subject matter experts with business/industry knowledge. A statistician can also serve as a compliment to such a team depending upon the facts and circumstances of the matter and the data analysis approach (i.e., using a sampling of invoices to extrapolate damages). Regardless of the data source, possible tasks the digital forensics expert or a data analyst may perform include:

- Data Source Identification (Scoping)—at the onset of a potential matter, it is often unknown what data is needed, exists, and is available. Digital forensics experts can be involved in conversations to address all three topics.
- Data Collection and Preservation (Discovery)—once data sources are defined, the collection of identified data must be executed. If the data is

produced by opposing counsel, defining production parameters can be important to avoiding inefficient "back and forth."

- Document Control Totals and Validation—documenting information about the data is a critical initial part of validation. Often, this includes recording data sources, record counts, sum of number fields, date ranges, account numbers, and other relevant fields. At this stage, outliers and data exceptions are evaluated. For example, email records with sent dates of 1/1/1900 are clearly "dirty data" as email did not exist in the year 1900.
- Conversion and Normalization—data may be unstructured or in a nontabular format unsuitable for creating relationships with other data sets; think PDF bank statements, standard QuickBooks reports, media, etc. Conversion may be necessary to develop clean sets of data that eliminates redundancy and applies a uniform structure.
- Relational Database—as most matters involve data sources that do not have active linking or are dissimilar in context, relationships between the data may need to be manually built. Consider this the "overlay" process.
- Summarizing/Reporting—a clean dataset may be provided for the forensic investigators or developing analysis in collaboration with the investigators. Results of summary data may need to be traceable back to the source data for export reports.

Forensic Investigations and Commercial Damages

Fraud investigators and commercial damages experts may collaborate with data analysts to find meaning in the data and determine how it can be applied in the context of active or potential litigation. For example, is there a correlation between building and software access logs (operational data), metadata in certain allegedly altered documents, and the recordation of fictitious entries in a company's accounting system? Likewise, are vital events identifiable in data extracted from industry/task-specific SaaS applications that may be casually connected to alleged commercial damages? Such collaboration between fraud/damages experts and data analysts may be a "looping" process by continuously considering certain questions—i.e., what does the data mean? If meaning is found, is it useful? If useful, how can it be applied to this litigation?—then determining an actionable approach.

David H. Glusman is an advisory partner in Marcum LLP's Philadelphia office. He provides consulting services in the areas of forensic accounting, litigation support, health care fraud, and high net worth estates, trusts and taxation. Contact him at david.glusman@marcumllp.com. Neil Young *is an IT forensic analyst at the firm and* Jesse J. LaGrossa *is a manager in the valuation, forensic and litigation support services department at the firm.*