

## The Reliability Of Structured Data In Litigation

*Law360, New York (June 25, 2010, 12:39 PM ET)* -- Lawyers are familiar with the complexities of collecting and processing electronically stored information in response to discovery requests. The scope of such requests often requires parties to collect data from numerous sources, including enterprise resource planning (ERP) systems, hard drives, e-mail accounts, smartphones and other electronic storage media.

The burden and expenses associated with electronic discovery, as well as the potential legal pitfalls, have prompted many law firms to initiate practices which exclusively handle e-discovery issues. The focus of these practices generally has been “unstructured” data, such as internal memoranda and presentations, e-mails and scanned correspondence.

A second category of information is known as “structured” data. Structured data refers to databases maintained in programs such as Microsoft Excel and Access, Oracle, PeopleSoft and SAP, which contain records of sales, products, employees, prices, accounting data or financial statements. Issues with this type of data have been at the center of a number of recent legal challenges.

The use of structured data requires special attention beyond simple identification and production in response to interrogatories or discovery requests. A frequently underappreciated, yet vitally important element of structured data processing is conducting a review of the quality, consistency and reliability of the data.

This process often requires rigorous preparation and assessment beyond that which is needed for other types of document discovery. Legacy databases, mergers and the departure of IT personnel may present unique challenges in this process. Increasingly, courts are taking a very critical view of erroneous data. Mistakes in the production of structured data in a number of recent cases have had significant consequences, particularly for expert witnesses.

Structured data is frequently the foundation of expert analysis in litigation for a variety of damages and liability issues. Often, statistical analyses prepared by experts are based on structured data from the parties to the litigation. The Federal Rules of Evidence explicitly require that expert testimony must be “based upon sufficient facts or data.” The recent decision in *Kreidler v. Pixler* is the latest in a series of federal cases in which courts excluded expert testimony on the basis of “unreliable data.”

In *Kreidler*, plaintiffs claimed that defendants engaged in a breach of contract, where defendants “failed to comply with payment terms under the Workers Compensation Contract.” One of the defendants’ experts used accounting data to summarize payments and disbursements between plaintiff, Cascade National, and defendant, Midwest Merger Management. The defense expert’s testimony was based on

two sources of data: (1) a Midwest database exported from QuickBooks, a commonly used accounting program, and (2) disbursement data prepared by defense counsel.

The court found there was sufficient evidence that Midwest's accounting data was problematic because it "was not maintained in accordance with good accounting practices" and was "significantly revised years after the entries were made when Midwest's financial records were subpoenaed." Moreover, the disbursement data relied upon by the expert was created by counsel for the purpose of the litigation, and the court did not find "any evidence that [the expert] sought to verify the information presented to him."

Understanding how structured data is built and managed by the client can be an important part of an expert's assignment in litigation. These types of data mistakes can be avoided by experts by performing a data quality assessment before using the data in subsequent analyses.

In *Barron ex rel. U.S. v. Deloitte & Touche*, the damages estimate submitted by the plaintiff's expert in a dispute over Medicaid reimbursement was excluded due to the fact that his calculation was based upon "bad data." In this case, a sampling of relevant school districts was conducted and relied upon to calculate damages. In the random sample of districts chosen by the expert, a large proportion had missing observations.

The expert treated these missing observations as invalid claims, or "paid claims [that] were not made in accordance with Federal and State requirements and were false." The expert made this assumption despite the fact that many other explanations potentially existed, particularly limitations on document retention.

Here, the court objected to the fact that the expert relied upon a dataset that was necessarily incomplete and made no effort to rectify or account for the problem. The court subsequently excluded the expert's testimony because "the data upon which [the expert] relied in forming an opinion as to damages is so unreliable and lacking in probative force that no reasonable expert could base an opinion upon them."

In *Johnson Electric North America v. Mabuchi Motor America*, the plaintiff's motion in limine to preclude the defense damages expert's testimony was granted. In his report, the defense expert constructed a statistical analysis to "estimate how many of Johnson's allegedly infringing micro-motors were used in the United States" and calculated the damages due to Mabuchi. The two types of motors at issue were "100 series motors, which are typically used in small household appliances, and the 200 series motors, which are typically used in automobiles."

In this case, the defense damages expert constructed an econometric model to estimate price erosion damages in the market for mini-motors. Despite using a sound regression analysis, characterized by the court as having a "dazzling sheen of erudition," the defense expert's reliance on erroneous data was one of main factors that led to the exclusion of his testimony.

According to the court, the defense expert "assumed that all the Johnson micro-motors were destined for small household appliances when calculating lost profits, his inexplicable reliance on automotive data introduced an error with unknown but probably significant impact on his conclusions."

Further the court concluded "if Mabuchi wishes to reconstruct the micro-motor market, that reconstruction must be grounded on the most relevant and reliable data available." The expert's

reliance on third-party structured data that did not match the market dynamics and facts of the case was a critical mistake.

So, what lessons can be learned from these cases?

First, structured data can play a pivotal role in litigation. Expert analysis often hinges upon the accuracy and completeness of structured databases. Addressing data issues early in litigation is therefore critical to a successful outcome.

Second, the implications of data accuracy can reach far beyond discovery sanctions — it can ultimately affect expert testimony, settlements, damage claims and the credibility assessments in the case.

Third, when possible, an early assessment of the validity, credibility, and applicability of structured data is critical — this type of assessment can be made by a consulting expert separate from any potential testifying witnesses.

Finally, when providing and using structured data as a basis for an expert opinion, attorneys should be certain that testifying experts are fully engaged in the relevant details of the data sets upon which they will rely.

--By Matthew Milner, Edgeworth Economics LLC

*Matthew Milner is a senior vice president at Edgeworth Economics and chairman of the firm's data management practice in Washington, D.C. He can be reached at [mmilner@edgewortheconomics.com](mailto:mmilner@edgewortheconomics.com).*

*The opinions expressed are those of the author and do not necessarily reflect the views of Portfolio Media, publisher of Law360.*

---

All Content © 2003-2014, Portfolio Media, Inc.