

# MANAGING ENTERPRISE DATABASE DISCOVERY IN PHARMACEUTICAL AND MEDICAL DEVICE LITIGATION

**Conducting litigation in the pharmaceutical industry takes special efforts when it comes to both data discovery and intellectual property protection. Understanding the technical and business aspects of database systems, and the steps one must follow to properly undertake litigation and the accompanying database discovery in the pharmaceutical industry is critical to success.**

**P**harmaceutical product litigation in US courts involves highly complex data that pose unique challenges for discovery — the pretrial phase in lawsuits in which each party obtains evidence from the opposing party through various means, including requests for document production and depositions. Data for discovery in everyday litigation typically includes email and custodian documents, but in pharmaceutical litigation, material often includes unusual data types, such as enterprise databases that manage information in categories such as clinical trials, sales, adverse events, medical inquiry and FDA filings. This information is stored in intricate proprietary systems and off-the-shelf packages that include sophisticated workflows and connectivity capabilities.

Discovery of this information is a multistep process that involves understanding the technical and business aspects of each system, as well as the manner in which it is developed, maintained and operated. Protection of this information outside the corporate firewall is far more complex than for traditional forms of custodian data.

## Producing Enterprise Databases as Documents or Standard Reports

Requesting parties, provided with source database data in static reports, have argued that they should be afforded the same analytical capability available to defendants, such as the ability to statistically analyse adverse reactions to drugs. This raises several issues.

A traditional method of collecting from enterprise systems is to run ‘front end’ reports into a document format and process the reports with an eDiscovery vendor. Requesting parties, however, are now challenging the sufficiency of this method. Reports arguably do not give them the same analytical functionality that defendants have and, therefore, courts often mandate production “in the format in which it is maintained.” If enterprise systems have been identified as containing relevant information, the traditional industry response has been to

treat databases as exceptions. Resisting discovery with the argument that the “data is not reasonably accessible” is, however, not holding water in today’s environment.

This ‘native’ concept does not lend itself well to the consideration of complex enterprise systems, regardless of whether they are large-scale enterprise resource planning (ERP), financial, human resources, messaging or manufacturing systems. In these cases, the native repositories are rarely produced as a whole. Instead, relevant records and their metadata are extracted from the enterprise database, prepared, reviewed and produced in an agreed-upon format. Each system is unique, but all may be approached and assessed using a defined methodology, summarized below:

- Take stock of databases that contain relevant information. Identify systems, platforms and application functionality at both the operational and business levels, and document the nature of the information stored within each database. Be sure to include legacy and proprietary data sources.
- Understand how various systems inter-relate. Systems can contain derived data that is retrieved from other data sources. Identify where data redundancy exists and eliminate it by tying data to its original source.
- Narrow the number of unique databases. In pharmaceutical companies that are the result of mergers and acquisitions, multiple repositories may perform overlapping functions in slightly different ways. Furthermore, although systems may be migrated and replaced, legacy data may remain behind.
- Identify relevant data and the selection process during meet and confer sessions.
- Reach agreements concerning specific data elements to be retained in the database production, production formats, documentation and scheduling.
- Develop extraction methodologies and scripts, and quality-check the extracted data.

The following sections outline the major issues involved in producing these materials in native format.



### Interpreting the Request and Interacting with the Requesting Party

As a preliminary matter, the production request must be vetted to confirm that there is a need to take measures apart from traditional discovery. The requesting party may have awareness of certain systems as a result of a course of dealing, a meet and confer, interrogatories or the deposition of a Rule 30(b)(6) witness. As another example, because US-based manufacturers of medical devices are subject to FDA regulations for quality systems (CFR 21 Part 820), such recordkeeping systems are specifically mandated, and, therefore, are generally known to be targets in certain cases.

Further steps toward qualifying the request for production include an application of relevant time frames and related products. The relevant databases in many large corporations may include multiple versions of such systems as adverse events and sales data, stemming from retired systems and corporate acquisitions. Linked applications such as workflow tools and supporting reference tables must be evaluated for relevance as well.

Meet with both 'business' and 'technical' owners of these systems well in advance to allow the collection

professional to interpret the provenance of the data for later certification, assess volumes, obtain permissions, determine costs and time frames, address confidentiality and licensing issues, collect all applicable documentation, and understand the exact nature of how the system is used in practice. This will set the stage for negotiations with the requesting party. The goal of any production is to offer only the responsive, nonprivileged materials — but especially difficult in cases involving large productions of structured data is ensuring that the producing party does so without revealing intellectual property (IP). The meet and confer may include a 'quick peek' at the database schema to establish relevant elements. Database schemas may involve hundreds of columns, tables and millions of rows of data. Requesting parties have become increasingly sophisticated at understanding these underlying data structures and building their own electronic hosting systems for the benefit of plaintiffs acting in concert.

The legal team must prepare to estimate the cost of the effort and associated time frame, and determine who shall bear it. They must determine which portion of the database plaintiffs are actually entitled, what must be redacted, as well as which ancillary or third-party data must be included.

They must also determine whether third-party aggregate data (for example, industry analytics from IMS Health) must be produced. A complete and verifiable cost document becomes an important component of interactions with the requesting party and for motion practice.

### Managing and Documenting the Process within the Corporate Enterprise

In pharmaceutical eDiscovery, the collection team is rarely in a position to capture enterprise data directly. The process involves a co-ordinated effort with teams of internal IT, security and risk, legal, business, and records personnel to secure the information on the corporate side of the firewall, execute the collection whilst maintaining chain-of-custody and data integrity, and document the process defensibly. Business owners must often participate in the process of documenting their systems and workflows, and those materials must then be vetted with legal before becoming actionable. Typically an iterative process is required to create the correct queries for the desired extractions, determine the right data sets, and confirm time frames (including an examination of archival and backup practices).

Because of the project-orientation of the typical corporate IT department, discovery data collection may be subject to IT department priorities. Creating timeframes for these projects should inform negotiations with the requesting party. In certain instances, there may be alternative means to obtain the information (such as restoring a backup tape versus querying a data warehouse versus querying a live system) that may vastly affect the cost and time frames.

### Processing and Preparing Enterprise Data for Review, Redaction and Production

At this point, the data passes from the corporation to the law firm, vendor, or other eDiscovery professional. It should be accompanied by sufficient documentation to allow the preparer to understand the format and data elements, but lines of communication should remain open. The preparer must create an environment by which the material can be loaded for assessment, processing, review and production. If there are multiple versions of the repository, they must be aligned and consolidated; for example, an adverse event database from one supplier may include a column called CASE\_NOTE and another may include a column called CASE\_DESCRIPTION. The preparer must determine whether these columns refer to the same information and should be combined.

The preparer must also analyse the data for missing or incorrect information. Although it is not the preparer's responsibility to ensure the quality of the inbound data, he or she must be prepared to speak to the condition of the data. Ultimately, the data are processed into a format whereby it can serve the purposes of the discovery process, such as ad hoc

query and reporting by the producing party, review by the attorney teams, and production in the format mandated by the case order or negotiation.

The production plan must also take into account third-party sources of information and their licensing (such as built-in MedDRA dictionaries that help define coded entries), as well as confidential, proprietary or EU/HIPAA or otherwise protected information. The redaction process must consider not just obvious forms of patient identification (such as social security number), but also other elements that may readily identify a patient, such as the name of the doctor coupled with the related patient zip code.

Many of these functions, such as staging an inquiry platform on a secure Web server or creating a turnkey redaction system for the attorneys, require substantial custom programming and provision for thorough audit trails. Complex hierarchical systems such as the format for the electronic Common Technical Document used to transfer regulatory information can involve extensive programming to present the material to attorneys for review. Prior to production, the data sets must be subjected to a consistent and auditable quality control process. The creation of appropriate affidavits and other documentation by the preparer forms an important part of the production effort.

### Protecting Data, Confidentiality and Corporate Know-How Outside the Firewall

Inside the corporate firewall, enterprise data may be protected through a variety of means: physical security, database, network and application security, encryption, tightly controlled conditions for access including time constraints, audit trails, numerous supporting computer applications, business rules and policies, and compartmentalized job functions. Once the data is reduced to a transportable format for discovery, many of the safeguards are severed from the electronic information. Notwithstanding legal safeguards such as protective orders, nondisclosure agreements and confidentiality designations, the legal team must enact creative safeguards to protect the information beyond the life span of the case.

### Conclusion

The pharmaceutical industry is highly regulated and highly litigious, so it's in each company's best interest to learn as much as it can about discovery to ensure it is remaining in compliance with the rules governing the process. Although discovery within the pharmaceutical industry can pose a challenge, following the best practice recommendations set forth here can ensure the process progresses smoothly — from identifying and working with internal stakeholders to structuring data production formats to ensuring the security of data — and the company's IP — once it's been produced. **Pharma**



James Mittenthal

#### For more information

James Mittenthal  
Vice President  
Consulting Service  
Epiq Systems  
Tel. +1 212 710 6933  
jmittenthal@epiqsystems.com  
www.epiqsystems.com