The Intersection of Data Privacy and E-Discovery

By Andrea Donovan Napp – December 17, 2014

In the wake of data breach announcements by several major retailers, data privacy and security are hot topics in the legal community. Similarly, most litigators are likely tired of being subjected to endless articles and presentations on the pitfalls of e-discovery. Seldom, however, does anyone discuss the intersection of these related issues, which have their roots in "big data." Although it is infrequently addressed, there is a significant nexus between the two concepts that grows more pronounced as the volume of data generated multiplies exponentially and the ability of e-discovery tools to collect and process the data grows increasingly sophisticated. Specifically, the e-discovery process presents a very real risk of unintentionally compromising personal identifying information (PII). While there is the real possibility that law firms might be subject to attack by hackers seeking to access what they perceive as vulnerable repositories of valuable data, there is a much more mundane, yet equally harmful threat: the inadvertent disclosure of PII through the routine document production process.

The Problem: PII Caught Up in the Collection of Documents

While this may seem like hyperbole, the following scenario is likely familiar to most litigators. A law firm is engaged to assist a large corporate client, ABC Corp., litigate an unfair trade practices claim related to the manner in which ABC uses its market power to negotiate deals with other commercial players in the market. In conjunction with that effort, the law firm collects electronically stored information (ESI) from 25 custodians, all current employees of ABC. Specifically, the firm uses keywords that it believes will be responsive to the document requests to collect emails and all loose electronic documents from the individual employees' computers. That ESI is then uploaded to a review platform, and a team of contract attorneys is dispatched to review the materials. Because of the large volume of data collected, the firm is instructed to review for privilege only, and the keywords are relied on to determine responsiveness. As a result of this instruction, the review team never notices that the data set included Jane Doe's personal tax returns, which she sent to her accountant using her work email address; Harry Smith's network password and log-on information, which he saved in a Word document; Joe Jones's Social Security number, which was on a payroll document sent to his supervisor; and Mary Murphy's health records, which she had scanned to herself to submit to her insurer. If this material is produced to opposing counsel, has a data breach occurred? The answer, quite possibly, is yes.

The Context: Data Privacy Protections in Civil Litigation

As many litigants know, there is an inherent tension between e-discovery and privacy. In fact, unlike the European Union, the U.S. has a long history of promoting freedom of information over privacy. In addition to the free exchange of information that is encouraged, if not required, by the Federal Rules of Civil Procedure during discovery, this tradition manifests itself in our Freedom of Information laws as well as our open court system. In part, because of this rift between access to information and privacy, there is no comprehensive statutory scheme or overarching policy shaping U.S. data breach laws. Instead, a patchwork of federal and state regulations have sprung up to address the need for privacy regulations. While some of the federal statutes are more complete, such as the Health Insurance Portability and Accountability Act (HIPAA), many of the state regulations contain somewhat ad hoc and wide-ranging privacy regulations, notification requirements, and enforcement provisions. Over the past several years, the confluence of the development of expansive state privacy statutes, the maturation of HIPAA, the exponential growth of data, and the increasing emphasis on e-discovery in civil litigation has brought the long-simmering tension between e-discovery and individual privacy to the boiling point.

Connecticut's data privacy statutes provide an example of the kind of far-reaching statutory language that can turn the routine production of documents in a business torts case into a technical data breach. Connecticut's Protection of Social Security Numbers and Personal Information Act, Conn. Gen. Stat. § 42-470 et seq., provides that "[a]ny person *in possession* of *personal information* of another person shall safeguard the data, computer files and documents containing the information from misuse by third parties, and shall destroy, erase or make unreadable such data, computer files and document prior to disclosure." (Emphasis added.) "Personal Information" is defined as information

> *capable* of being *associated* with a particular individual through one or more identifiers, including, but not limited to, a Social Security number, a driver license number, a state identification card number, a credit or debit card number, a passport number, an alien registration number or a health insurance identification number. . . .

Id. (emphasis added).

Accordingly, Connecticut protects from disclosure any information that is merely capable of being associated with an individual, regardless of whose possession it is in or how it came to be there. Notably, violators may be subject to civil monetary penalties.

The most famous of all data privacy statutes, HIPAA, protects personal health information (PHI). As of 2013, HIPAA applies with equal force to health care providers, as well as any downstream contractors that receive, access, maintain, or transmit PHI. Thus, attorneys who receive and produce PHI could run afoul of HIPAA's data breach provisions and be subject to civil monetary penalties as well a breach notification requirements.

These statutes are merely two examples of the types of privacy regulations that might apply to data under a law firm's control during discovery. Although there are currently few, if any, reported decisions applying data breach laws in the e-discovery context, running afoul of these regulations could cause unnecessary aggravation for counsel, subject an attorney to disciplinary action, and create needless professional liability exposure.

The Fix: Best Practices to Avoid Production of PII
Avoiding production of PII or PHI is possible if the review team (or review technology) is trained to spot categories of protected information. Although the definition of PII or protected information varies by jurisdiction, there are certain categories of information that are generally recognized as sensitive and that should be safeguarded from unnecessary dissemination in the discovery process. These categories include the following information:

- Social Security numbers
- driver's license, passport, or state identification numbers
- taxpayer identification numbers
- any financial account numbers, credit card numbers, or other personal financial information
- any log-in/password information
- PHI
- birthdays in conjunction with any other identifying information

Many of these categories are set forth in Federal Rule of Civil Procedure 5.2, which identifies the types of data that should not be included in materials filed with the court. Although Rule 5.2 does not specifically address discovery material, it is likely that a court would find it persuasive authority if a "discovery data breach" issue was brought before it. In business torts or unfair competition cases, attorneys are frequently focused on the commercial aspects between corporate parties and are not focused on the potential for PII in data sets. As illustrated in the hypothetical above, PII can often lurk in data sets culled for a specific commercial purpose. The following strategies can help minimize the potential for producing arguably protected data:

- Conduct targeted collections. Over-collection of ESI causes many problems, the inclusion of PII among them. The more information counsel learns from knowledgeable custodians prior to collection and the more targeted the collection is, the lower the likelihood of sweeping up personal identifying data.
- Have sensitive information highlighted. If you are using an electronic review platform, it likely has the ability to highlight terms that may be sensitive, including Social Security numbers and numbers that might appear to be credit card numbers. Have these types of terms appear in a bright color to call the reviewer's attention to them, as you would with privilege terms.
- Emphasize the importance. Frequently, when preparing training manuals or protocols for document reviews, attorneys focus intently on identifying responsive materials and protecting privileged information to the exclusion of other practical factors. When drafting your materials, devote an entire section to personal information and be very clear about how it should be treated.
- Offer your reviewers the right tools. Make sure the review team knows what to do when it comes across PII. The coding form should include a field for PII that allows a reviewer to indicate that PII is present and that a redaction is needed. If the ESI will be produced in TIFF, have the reviewers image and redact on the fly. Alternatively, include a text field that allows the reviewer to specify the nature and location of the redaction.
- Perform quality control searches. When you get ready to produce, perform some additional quality control searches to ensure that you are not letting any PII out the door unknowingly. Sample searches include "SSN," "Visa," and "passport."
- Establish appropriate protocols. Give appropriate consideration to the nature of your data. In commercial cases, we endeavor to collect primarily business data and often structure our reviews accordingly, sometimes not reviewing for responsiveness or relying principally on technology-assisted review. While this has its risks, protocols may make sense in that context. However, in a products liability case involving allegations of consumer injury, make sure the review team is on alert for consumer information and PHI.

Conclusion
Modern document reviews have a lot of moving parts. Nonetheless, litigators should be aware of the potential for inadvertent disclosure of PII through the routine document production process. As data volumes continue to increase, so does the potential for Social Security numbers, account information, and other PII to slip through the cracks. While protection of PII may seem a minor issue in the context of a major, bet-the-company

commercial dispute, production of such material to opposing counsel might very well be considered a data breach. Attorneys should therefore safeguard against producing this arguably protected data by implementing the strategies previously listed.

Keywords: litigation, business torts, privacy, discovery, personal identifying information, Social Security number, HIPAA, data breach

Andrea Donovan Napp is a business litigator with Robinson + Cole LLP in Hartford, Connecticut.

More Information
» Business Torts Home
» Practice Points
» Articles
» Case Notes
» Programs & Materials
» Business Torts Litigation Committee
- About
- Join

Back to Top