BY DAVID FISHEL AND CAROLE LEVITT

New Tools in the Discovery of Sound Recordings

RECENT CHANGES to the Federal Rules of Civil Procedure are poised to move the discovery and analysis of sound recordings out of the realm of high-profile government investigations and into the everyday world of normal litigation. While only a few attorneys currently have experience with discovery of small amounts of recordings from voice mail, fewer still have engaged in discovery of hundreds or thousands of hours of voice mail or call center recordings. The proliferation of audio recordings created in the course of business will make the availability of this kind of evidence more common. Along with the new rules are new tools that make discovery and analysis of this evidence more manageable.

On the surface, the challenges presented by hundreds of hours of audio recordings are similar to those presented by any other large body of potentially relevant materials. Requesting parties must skillfully frame their requests. Responding parties will attempt to limit the scope of requests and to manage the production process as best they can. The tasks are similar to any e-discovery endeavor. Parties are faced with locating, extracting, searching, reviewing, culling, and producing responsive recordings, and doing so accurately, cost-effectively, and under short deadlines. The primary pitfalls stem from the complexity of recording technologies and the sheer size of recorded audio collections, combined with the difficulty of searching them.

Audio recordings have long been discoverable and admissible as evidence and have often provided compelling proof in court, but most lawyers have not yet conducted audio discovery on the massive scale now common for electronic and paper documents. That is about to change. Recordings can be very good evidence. They provide a person's words in his or her voice. In many cases this might include a tone of voice, verbal inflection, or snicker that conveys far more information than a written transcript ever could. Recordings can contain "smoking gun" statements and are very effective when used in conjunction with documentary evidence or witness testimony. With the right analytical tools, audio can be used to quantify evidence of specific behaviors, such as the number of times a company engaged in a specific prohibited activity.

Three factors are rapidly transforming discovery of audio evidence:

- Recent amendments to the Federal Rules of Civil Procedure (FRCP) expressly identify "sound recordings" as "electronically stored information" and impose new requirements for disclosure, case management, planning, and form of production of all electronically stored information.
- The quantity of sound recordings is increasing because of technology advances such as digital storage, regular office and cellular phone voice mail, VoIP telephony, and unified messaging systems that deliver voice mail as e-mail.
- New tools are making it feasible to find, extract, convert, search, and produce audio recordings.

Amendments to the FRCP (and similar provisions in state courts) will soon affect discovery requirements and practices regarding sound recordings. The new rules make it clear that judges and lawyers must understand and manage electronic discovery from the beginning of a case. Parties need to decide whether and how to ask for sound recordings. In addition, parties should have a reasonable understanding of recording and storage systems, put appropriate controls in place, formulate credible plans for how they will review and produce responsive recordings, and decide on the forms in which the recordings will be produced.

Phonetic audio search technology is based on breaking down audio recordings by analyzing the smallest components of human speech.

The newly revised Rule 34(a) specifically identifies "sound recordings" as "electronically stored information" (ESI). This new term clarifies and expands the "documents and data compilations" included in the earlier version of the rule. The new rule also allows the requesting party to specify the form in which it wants the information produced. The stated aim of this provision is production of ESI in a "reasonably usable" form. Audio recordings are captured in a wide variety of formats, many of which are proprietary to specific recording systems, so the question of whether a sound recording is reasonably usable can be expected to arise regularly. The cost and time required to convert a large collection of recordings from a proprietary format to a more "usable" one can be significant. It behooves attorneys to learn about the various audio formats and seek early expert help in making decisions about production issues so that they can be fully prepared to defend their preferred form of production.

Amended FRCP Deadline

Attorneys are required to do this very quickly. Amended FRCP 16(b) requires parties to hold a scheduling conference to consider electronic discovery plans within 120 days of the commencement of an action. Furthermore, at least 21 days before this scheduling conference, parties must meet to discuss and, if possible agree upon, electronic discovery procedures for the case.

Rule 26(f) specifically requires the parties to address "any issues relating to disclosure or discovery of electronically stored information, including the form or forms in which it should be produced." In relation to audio recordings, this means that parties must be knowledgeable about the relevant recording and storage systems, audio

David Fishel is a litigator and senior director and technology counsel for Nexidia. Carole Levitt is president of Internet For Lawyers.

formats, and other production issues, and must formulate their electronic discovery plans within the first 100 days of the life of a case. Following the Rule 16 scheduling conference, the court may issue a scheduling order that sets forth how electronic discovery will proceed. By the time of that order, parties should make certain that the court is fully informed of any issues regarding ESI, including audio evidence.

Discovering Audio

New digital recording technologies have rapidly increased the ability to record speech and to store vast amounts of those recordings. More recordings are being made, and since the new technologies are digital (that is, created and stored on computerized systems rather than analog tape), the growth is enabled by the same factors that brought about the recent explosion of electronic messaging and documents.

"Sound recordings" means voice mail to many attorneys, but voice mail is not the only variety of audio evidence. Collections of recordings may include audio archives from corporate call logging systems, Web conferences, or IP-based conference calls. Businesses that routinely record conversations with customers and business partners may have huge stores of potentially discoverable recordings. These business recordings are subject to the same requirements for retention, preservation, litigation holds, and production as any other potentially relevant information.

Digital recording technology spawns new opportunities and challenges for requesting and producing parties. Corporate call centers routinely record conversations between businesses and their customers for quality assurance. If the issues in a case include, for instance, the number and type of complaints about a certain product, when a company had notice of alleged problems with a product, or how a company treated its customers, that information may be found in the thousands (or millions) of calls the company received and recorded. Many businesses also record transactional conversations, such as trading or sales calls. This type of recording has proven critical in cases involving energy companies and commodities traders and is common in financial service businesses.

Voice mail itself is changing. Initially, voice mail technology moved slowly from the era of the cassette tape answering machine to the PBX. Now, businesses are rapidly moving to unified messaging systems in which all messages—voice, e-mail, or instant text—are delivered to the user's computer desktop and stored along with other message traffic on mail servers. In this environment, voice mail is converted into .wav format attachments to e-mail messages. Some routine e-mail discovery could now include voice mail attachments.

In the past, voice mail recordings were generally kept in a central location, with strict limits on the number of messages each user could save and how long a message could be stored. A single person's mailbox was easy to identify and could only contain a few minutes of recordings. As businesses switch to unified messaging, the same problems that have vexed e-mail discovery—duplicates, the propensity of users to retain e-mail, and mailbox archives that are not under the direct control of system administrators—will arise for voice mail. Voice mail was previously sought in hope that a particular message may have been saved, but now a company's entire voice mail system may be considered to contain potentially relevant information. For example, an employment suit may seek evidence of a pattern of behavior or practice in a company's voice mail archive.

Tools to Meet the Challenge

With the advent of e-discovery of large amounts of audio comes the need to fully process them as discovery materials, and that includes finding responsive recordings. It is not uncommon for a litigant's IT and legal staff to have a good grasp on the location of audio evidence but to have no idea of the content of the recordings. At present, there are three primary ways of reviewing and searching recordings: listening, manual transcription, and phonetic search. (A fourth option, speech-to-text transcription by computer software, is widely regarded in the legal field as currently not sufficiently accurate to produce reliably searchable transcripts.) Listening and manual transcription are the two most used at present and can be effective for small collections of recordings. However, both are very expensive, slow, and cannot scale up to economically handle the several hundred hours of recordings that a larger matter might produce. Even for small audio collections, the cost of transcribing or listening to recordings can quickly become prohibitive.

Listening suffers from a number of limitations, the greatest of which is the inability to search the audio content. Having once listened to recordings, if a new search is needed, the recordings must be listened to again in their entirety. Listening costs vary widely. Paralegals or contract attorneys are often used for the initial review, but attorneys generally listen to the potentially relevant recordings in order to make legal determinations about them. In addition to prohibitive cost and lack of searchability, effective listening takes much longer than the duration of the recordings themselves. Also, attentive listening is difficult to sustain for any long period of time, and it is difficult to get directly back

to relevant bits of conversation even after they have been located.

Manual transcription is similarly slow and expensive. On average, a law firm can expect to spend about $120 per hour of audio recording to produce a usable transcript of audio content. Economical transcription services can also take a long time, a problem that increases as the size of an audio collection grows. It may require transcribing more than 100 to 200 hours of recordings just to locate a few seconds of audio evidence.

Phonetic audio search technology is based on breaking down audio recordings by analyzing the smallest components of human speech, known as phonemes. (There are roughly 40 phonemes used in spoken North American English.) Since 2000, Nexidia (www.nexidia.com) has applied phonetic search technology in the government intelligence and commercial call center arenas. More recently, the company began applying the technology to the e-discovery and corporate compliance market.

The high-speed phonetic audio search approach has two phases: preprocessing and searching. The first phase preprocesses the sound recordings to break the words into their component phonemes. This step produces a phonetic search track and on one processor occurs roughly 60 times faster than real time. Thus, one hour of audio recording can be rendered searchable in about one minute. This means that it is feasible to handle collections of thousands of hours of recordings for discovery. Reviewers can run searches against the phonetic index for words and phrases. Search results are linked directly to the point in the original recording where the search term was found, allowing reviewers to jump directly to the point in the recording containing potentially relevant terms. This greatly increases the speed at which reviewers can pinpoint their listening to potentially relevant passages of the audio.

The phonetic approach can be up to 98 percent accurate on recordings with the best quality. Accuracy can be as low as 70 percent when speakers have regional accents or the recording is of poor quality, contains blended words, proper names, slang, code words, or ad hoc usage. Accuracy of other speech recognition methods, such as speech-to-text, can drop as low as 50 percent.

More lawyers are going to conduct discovery of sound recordings. Those who are going to be successful need to understand the technical aspects of audio recording systems, master the rules that govern discovery of audio materials, and use effective tools to find and organize the information. Whether they are requesting or producing sound recordings, attorneys who develop those skills can act with confidence. ■