AVANSIC
E-Discovery & Digital Forensics

Concept, Privilege and Plaintiff Review: Oh My!
By Dr. Gavin W. Manes, CEO

Every document review project has a specific goal and therefore every project has a distinctive workflow. Producing party and plaintiff document reviews are two common projects with different goals, review phases, type of reviewers, and ideal types of software. Clustering and technology assisted review can be used in any review but they can be particularly powerful when applied

Software Requirements
Although there are many tools that perform both e-discovery processing and review capabilities, this discussion focuses only on the review phase of the process.

All review projects require a system that can easily present email for logical review; this is the most common type of document reviewed during discovery. When email is presented in conversation threads or grouped as a family with its attachments, it saves significant review time and effort.

Privilege review can be accomplished with an online review tool that allows coding for privilege and sub-coding for the type of privilege. A tool with fairly simple administrative capabilities is typically sufficient in all but large privilege review projects. However, it is critically important that the software be capable of automatically generating a privilege log based on the coding applied.

Plaintiff, hot document or production reviews can often benefit from more sophisticated software. Any software tool must be able to handle large number of folders and subfolders and they must be easy for administrators or certain users to create and organize at will. Given the broad net of many of these cases, such as class action suits, there can be dozens of areas of relevance. This tends to be a multi-pass type of review, where documents may be revisited many times. Therefore, the tool must have a very robust way to search, folder, and tag information. Reviewers must be able to assign multiple tags to a single document, as well as assign documents to multiple groups.

Concept Searching
Concept searching is the process of finding documents that are alike and assigning them to buckets based on those concepts. Attorneys can then pick which of these buckets to include or exclude for their particular project. There may be entire categories of documents that don't need review, and that is where concept searching can be a more powerful tool than keyword searching for large data sets.

Concept searching can be equally useful in hot doc and production reviews, but is generally not used in privilege review situations since those decisions are generally context-based.

Plaintiffs have recently seen great success in using concept searching; document set size and the broad scope of documents sought are most likely the driving factors. The

power of concept searching lies in the initial sort of documents into categories that project managers can then assign based on their level of relevance. Likewise producing parties have used concept searching to assist with batching review documents. This batching technique is superior to grouping by custodian, date or email thread since it groups conceptually similar documents together.

Predictive Coding or Near-Dupe Coding

Predictive coding, also known as technology assisted review, is the process of populating coding from a sample set to the rest of a document set. This is a complex process but in general, the initial coding is performed by a human and the extrapolation of that coding is projected to the remaining documents by a computer. It results in coding assigned to documents that no human has actually seen. Although this may seem a frightening prospect quality checks and error/recall rates are part of the process. These quality control steps may require the process to begin again. There are many different descriptions and implementations of this process; some are automated, some use Excel, some are built into tools. Thus, predictive coding is not just something to be tacked on to an existing process. When using any type of technology assisted review, the risks and benefits should be contemplated from the beginning of an e-discovery project.

For a hot document review, predictive coding has similar benefits to concept searching because they both rely on the same underlying machine learning techniques. Similarly, predictive coding does not provide major assistance when performing a privilege review. Plaintiffs could benefit from predictive coding to project coding on reviewed documents, but it is more common to use clustering and concept searching.

Conclusion

Document review technology is converging and processes are improving.
Investments in these three key areas are different; buy technology, train people, and develop methods. Good analysis of those three factors for the particular type of review will lead you down the yellow brick road.

About the Author

Dr. Gavin Manes is a nationally recognized expert in e-discovery and digital forensics. He is currently the CEO of Avansic: E-Discovery & Digital Forensics, which provides e-discovery services to law firms, litigation support departments, and corporations across the country.

In the development of Avansic, Dr. Manes has made a commitment to high-technology innovation, research, and mentorship; he currently has several patents pending. Avansic's scientific approach to e-discovery and digital forensics stems from Manes' academic experience. Manes has published over fifty papers on e-discovery, digital forensics and computer security, and has given hundreds of presentations to attorneys, executives, professors, law enforcement, and professional groups on topics ranging from e-discovery issues to cyber law. He has also briefed the White House, Department of the Interior, the National Security Council, and the Pentagon on computer security and forensics issues.