

Bulk Tagging Best Practices

Streamline Management of Duplicate Document, Conversation Threads and Family Groups

As the number of electronic documents created for business grows, increasing the speed at which large record sets can be reviewed is critical to meeting tight production deadlines in any e-discovery document review project. One of the biggest challenges attorneys face in document review is how to efficiently and effectively handle the multitudes of duplicate documents found in virtually every electronic discovery project.

Applied Discovery uses the term “bulk tagging” to describe the set of features that allow users to apply codes to bulk sets of documents within our Online Review Application (ORA). These types of document sets include family groups (i.e., zip files and emails with attachments), email conversation threads, duplicate documents and even annotations.

Benefits of Bulk Tagging

The ability to apply review codes to bulk sets of documents can dramatically increase reviewer productivity and ensure consistent coding of documents throughout the dataset.

Consider this scenario: The CEO of a company had sent an email with an attached spreadsheet to 12 members of the executive team. Not only does that same email exist in the folders of 13 different custodians (the CEO plus the 12 other executives), that spreadsheet does as well. If the reviewer assigned to review the CEO's documents determines that the email and the attachment are privileged, it is likely that all instances of both the email and the attachment should be coded as privileged, too.

It is clear that the ability to apply bulk tags to the conversation threads (the email chain) and the attachments (the family group) ensures that all the documents are consistently coded and saves time by eliminating the need to review each of the documents individually.

Technology Used to Identify Document Sets

Understanding how duplicate documents, family groups and conversation threads are identified can provide a level of comfort for those who might be wary of using bulk tagging features.

Conversation Threads

The technology used by Applied Discovery® to determine whether a document is part of a conversation threads in Microsoft® Outlook® is straightforward. During the data processing phase, the metadata is retained for each document and that metadata is analyzed to determine whether the document is related to any other document in the data set.

Family Groups

Applied Discovery treats each member of a family group—an email and its attachments—as separate documents so each can be coded separately. However, clients have the options to apply bulk tagging to ensure consistent coding for each member of the family group.

Duplicate Documents

Applied Discovery has implemented a proprietary process to identify duplicate documents during the data processing phase using an industry standard algorithm known as the MD-5 hash. The MD-5 hash was developed by Professor Ronald L. Rivest of MIT. The MD-5 algorithm takes as input a message of arbitrary length and produces as output a 128-bit fingerprint of the input. Security experts consider it to be infeasible to produce two different documents with the same MD-5 hash.

The MD-5 hash is the most secure method of verifying data integrity and is much more reliable than checksum and other commonly used methods.

Bulk Tagging Considerations

Because every e-discovery project is unique, it is critical that the review and production requirements, the scope of the project and the review team are thoroughly analyzed to determine if and how bulk tagging can be used in a given case. The criteria for managing document sets are different for every case. At the beginning of every project it is critical to work with account managers from your e-discovery service provider to develop a strategy and communicate that to the review team to streamline the document review. At Applied Discovery, once that strategy is developed, our clients use the ORA to assign custom administrative rights to enable or disable bulk tagging features for individual users to ensure that attorneys adhere to the review rules of the project.

Bulk Tagging Best Practices

The trick to using bulk tagging during document review is for the case administrator to develop rules for bulk tagging, clearly communicate those rules to the review team, and finally to perform scheduled quality checks to ensure the rules are consistently followed.

Some issues for the case administrator to consider when developing the rules are as follows:

- Which reviewers are allowed to use bulk tagging?
- Can reviewers bulk tag documents across the entire case or only within their assigned review documents (i.e., within a particular custodian)?
- What kinds of documents should be bulk tagged (non-responsive, privileged, etc.)?
- What is the protocol for tagging family groups and conversation threads?
- Should annotations automatically be applied when bulk tagging?
- What are the production requirements (i.e., will duplicates be included or just the originals)?
- Will bulk tagging be available throughout the duration of the document review or only in specific phases (i.e., available during initial review then disabled during second-level review)?

Conclusion

Without the tools needed to manage duplicate documents, family groups and conversation threads, the review team may end up spending an extraordinary amount of time double-checking to make sure that document sets were coded the same during review. By following the best practices outlined above, bulk tagging features can help increase reviewer productivity, flexibility and ensure consistency in your document review.

The Discovery Experts: Industry Relations

LexisNexis Discovery Services has the right consulting and technology choice for every discovery need. Top law firms, corporations and government agencies rely on the LexisNexis® products and services, Applied Discovery®, Concordance™, and Hosted FYI™, to meet their discovery obligations on time, accurately and cost-effectively. Services include records management consulting, data collection, forensics, media restoration, data filtering, data processing, review and document production in the format that each matter requires. The Industry Relations team works to educate the legal community on the continually evolving case law and technology of electronic discovery.

For more information or to contact the experts, please visit [lexisnexis.com/discovery](https://www.lexisnexis.com/discovery).

