

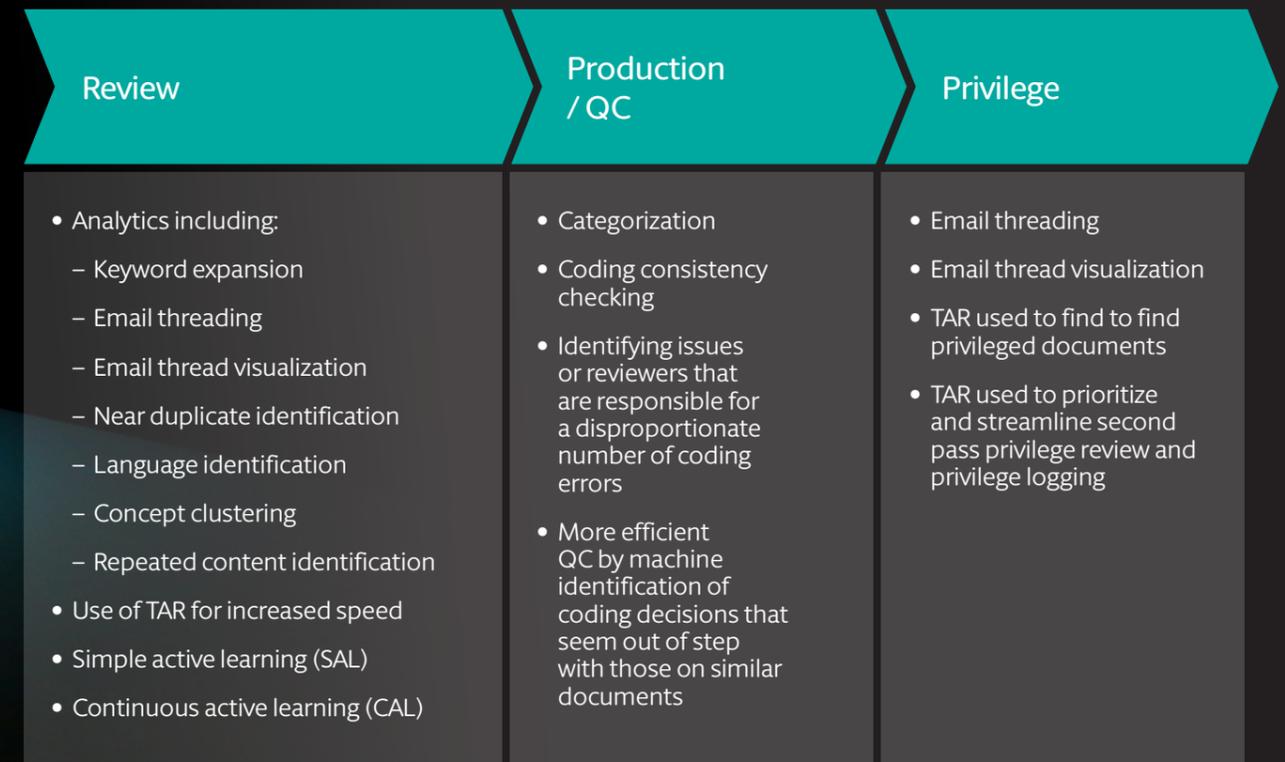
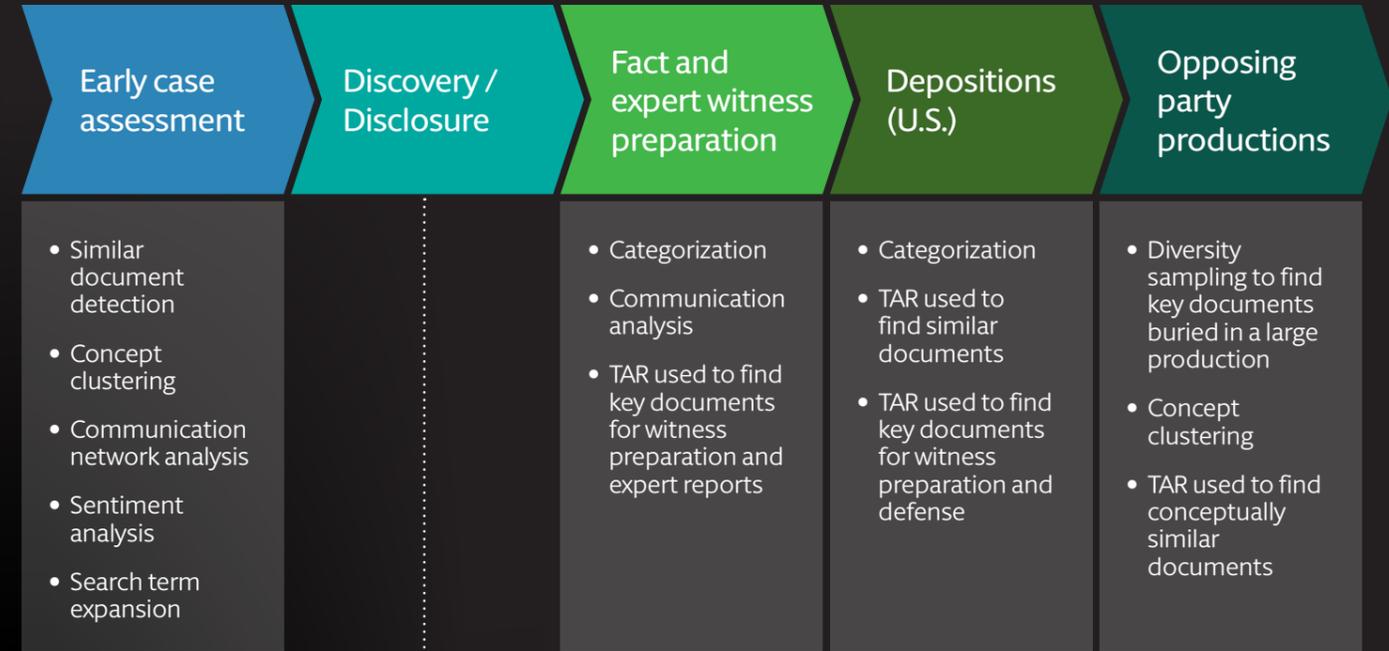
Hogan
Lovells

Litigation technology

Using technology in all stages of litigation

Technology assisted review (TAR) and analytics can be used across multiple stages of litigation. It allows us to increase efficiency and accuracy, helping to make significant reductions to the number of documents that need to be reviewed. It also helps us develop case strategy by ensuring we're able to find all the important data at the start of a case.

Litigation



Early case assessment

At the start of a matter, we can use our litigation technology to aid assessments of the case. The tools can help you get to the heart of the issues and build a robust, fully-informed case strategy.

- **Similar document detection** is used to identify key or 'hot' documents at the beginning of a matter based on known key documents. Sample documents are introduced and used to find conceptually similar documents. Sample documents could include a previously identified document, a portion of a pleading, or even a paragraph drafted by the case team.
- **Concept clustering** identifies relevant concepts for lawyer review. Once identified, the entire concept cluster can be batched for document review.
- **Communication network analysis** can be utilized to see who is talking to whom. This can also include analysis about who is talking about a particular subject. The software identifies all email addresses that a particular user uses and associates them with the individual.
- **Sentiment analysis** identifies and categorizes opinions expressed in a piece of text, especially to determine whether the writer's attitude toward a topic, product, etc., is positive, negative, or neutral.

Have court decisions supported the use of TAR in litigation?

In 2012, *Da Silva Moore v. Publicis Groupe* was the first published opinion in U.S. federal court recognizing TAR as an "acceptable way to search for relevant ESI in appropriate cases," but advised counsel to disclose the use of TAR to opposing counsel and seek agreement. In *Rio Tinto PLC v. Vale S.A.*, the same federal magistrate judge wrote that TAR can no longer be considered an "unproven technology," and, citing other courts following his lead, that "the case law has developed to the point that it is now black letter law that where the producing party wants to utilize TAR for document review, courts will permit it."

In 2015, in *Irish Bank Resolution Corp. v. Quinn*, the Irish Court of Appeal upheld the Irish High Court's order granting a responding party's motion to use TAR over the objection of the party requesting the production of documents.

In the UK, the English High Court approved the use of TAR in *Pyrrho Investments Ltd. v. MWB Property Ltd.*, when the parties jointly sought its use. The English High Court also approved the use of TAR over the objection of the requesting party in *David Brown v. BCA Trading*.

Do we need to disclose our use of analytics or TAR to opposing counsel or to the court?

Some court opinions in the U.S. suggest that parties need to disclose their use of TAR or predictive coding, which is more precisely called supervised machine learning. Many courts adopt the Sedona Conference's Cooperation Proclamation, "to promote cooperation by all parties to the discovery process to achieve the goal of a just, speedy, and inexpensive determination of every action." The intent is "to promote open and forthright information sharing, dialogue (internal and external), training, and the development of practical tools to facilitate cooperative, collaborative, transparent discovery."

The use of machine learning or analytics (e.g., email threading and near duplicate detection) assists with setting up a more efficient review. Their use is typically not disclosed. For example, email threading assembles all parts of an email string, so one reviewer can review the entire email thread. Another example, near duplicate detection will group all documents that are almost the same so that one reviewer can review similar documents allowing for a more consistent and efficient review.

Discovery/disclosure

Litigation technology comes into its own at the discovery or disclosure stages of litigation. It can be used to drive efficiencies and to spot patterns that would otherwise remain hidden.

Review

Analytics (data visualization and unsupervised machine learning) should be used to set up an efficient and consistent review.

- **Keyword expansion** is a form of a conceptual search where the analytics engine finds related words that assist with the discovery, selection, and refinement of keywords.
- **Email threading** is a process to identify and order emails belonging to the same conversation that starts with an original email (the beginning of the conversation) and includes all of the subsequent replies and forwards pertaining to that original email. Use of email threading ensures the consistency of the review so that one reviewer can review an entire email conversation at the same time, avoiding conflicting codes when multiple reviewers review individual portions of an email. The use of email threading typically reduces the dataset for review by 25%.
- **Email thread visualization** is an interactive visualization depicting the flow and branching of an email thread.
- **Near duplicate identification** identifies textually similar documents into related groups that are ranked with similarity scores. Near duplicates can be reviewed by one reviewer, which adds to consistency of the review and allows for the reviewer to understand the level of importance of the differences contained in each document.
- **Language identification** determines a document's primary and secondary languages and calculates the language breakdown.
- **Concept clustering** can be used to find terms to exclude, like junk email. The tool groups the concepts together; a human makes the decision to cull the data. It can also be used to quickly explore a large dataset or to find documents similar to one that is known to be relevant.
- **Repeated content identification** finds the repeated text, such as confidentiality footers or disclaimers, and suppresses it from the analytics index to increase the effectiveness of some analytics tools.

TAR (Supervised machine learning) is used to increase the speed of the review.

• **Simple active learning** (SAL or TAR 1.0) is machine learning that relies upon a control set, a seed set of chosen documents, and training rounds to build a classification model that will find other relevant documents in the dataset. This method is often used in litigation due to the metrics that can be shared with opposing counsel and the court, including recall and precision, which allows some room for negotiation. This older methodology is the most widely reported in court opinions and has widely accepted protocols developed around its use.

• **Continuous active learning** (CAL or TAR 2.0) is machine learning that makes use of all the available lawyer decisions, continuously re-ranking the documents in the background so that lawyers see relevant documents more quickly. This is sometimes called a prioritized review since the machine is controlling the order in which lawyers review documents by its constant re-ranking, often allowing review to defensibly stop once all of the remaining documents are substantially non-relevant.

Production/QC

TAR can be used for quality control by checking the consistency of coding by reviewers throughout the review process, including before production.

Using categorization, the analytics engine identifies and groups similar documents together based on a set of example documents manually identified by the user.

Privilege

Email threading is a structured analytics process that identifies and orders emails belonging to the same conversation. The result groups all the replies and forwards related to the original email. This technique allows for consistent coding of privileged portions of email strings since one reviewer will review the entire string at the same time.

Email thread visualization is an interactive diagram of an email thread, showing where the email trail branched off, where attachments were added, and so on. This technique allows the reviewer to see where the privileged communication begins and whether it is maintained through the email thread.

TAR can be used to find privileged communications based on examples of privileged documents and email properties, including the number of recipients and email domains indicating third-party recipients, outside counsel, or in-house counsel who perform a legal function rather than provide business advice. It is best to analyze known privileged documents first to help develop an appropriate TAR plan to find additional privileged documents.

Fact and expert witness preparation

TAR can be used for fact and expert witness preparation by using known documents to find conceptually similar documents.

The documents will contain a rank that indicates how conceptually similar the document is. This technique can be used on our client's productions or opposing party productions.

Communication analysis allows lawyers to visualize the flow of information among people, including by topic.

TAR can also be used to identify documents relevant to each issue and each witness, developing separate relevance ranking scores for every document in the collection. While lawyers are coding documents for each witness and issue, the TAR rankings for every other witness and issue are being automatically updated in the background.

Using categorization, the analytics engine can identify and group similar documents together based on a set of example documents manually identified by the user.

Opposing party productions

Diversity sampling through TAR can be used after other methods of search and review have completed.

The machine explicitly looks for unreviewed documents that are different from the ones that have already been seen. So if opposing counsel is trying to hide a small number of documents in a big production, this technique can quickly identify them.

TAR can also be used on opposing party productions by using known documents from our client's production or even artificial documents to find conceptually similar documents within the production immediately after it is received. The documents will contain a rank that indicates how conceptually similar the document is.

Concept clustering can be used to gain general knowledge of an opposing party production and to identify small groups of interesting documents within a much larger set.

Alicante
Amsterdam
Baltimore
Beijing
Birmingham
Boston
Brussels
Budapest
Colorado Springs
Denver
Dubai
Dusseldorf
Frankfurt
Hamburg
Hanoi
Ho Chi Minh City
Hong Kong
Houston
Jakarta
Johannesburg
London
Los Angeles
Louisville
Luxembourg
Madrid
Mexico City
Miami
Milan
Minneapolis
Monterrey
Moscow
Munich
New York
Northern Virginia
Paris
Perth
Philadelphia
Rio de Janeiro
Rome
San Francisco
São Paulo
Shanghai
Shanghai FTZ
Silicon Valley
Singapore
Sydney
Tokyo
Ulaanbaatar
Warsaw
Washington, D.C.
Zagreb

Our offices

Associated offices

www.hoganlovells.com

"Hogan Lovells" or the "firm" is an international legal practice that includes Hogan Lovells International LLP, Hogan Lovells US LLP and their affiliated businesses.

The word "partner" is used to describe a partner or member of Hogan Lovells International LLP, Hogan Lovells US LLP or any of their affiliated entities or any employee or consultant with equivalent standing. Certain individuals, who are designated as partners, but who are not members of Hogan Lovells International LLP, do not hold qualifications equivalent to members.

For more information about Hogan Lovells, the partners and their qualifications, see www.hoganlovells.com.

Where case studies are included, results achieved do not guarantee similar outcomes for other clients. Attorney advertising. Images of people may feature current or former lawyers and employees at Hogan Lovells or models not connected with the firm.

© Hogan Lovells 2019. All rights reserved.