

# DRAFT

## PROTOCOL (EXEMPLAR) GOVERNING PRODUCTION OF RELEVANT INFORMATION USING TECHNOLOGY ASSISTED REVIEW

BOLCH JUDICIAL INSTITUTE, DUKE LAW SCHOOL  
March 2019<sup>†</sup>

The parties agree that the following protocol<sup>1</sup> shall govern the use of technology assisted review (TAR) to identify and produce electronically stored information that is potentially relevant to a claim or defense.

### GENERAL PROVISIONS

The use of TAR is intended to assist in the just, speedy, and inexpensive resolution of the discovery relating to electronically stored information in accordance with Rule 1 of the Federal Rules of Civil Procedure. The parties agree that, regardless of the review method they use to identify and review documents for responsiveness to a discovery request, the process they use should be designed to yield as much responsive material as possible while remaining reasonable and proportionate. The parties recognize that any search methodology they use, including manual human review, will miss some responsive documents.<sup>2</sup> The parties further recognize that although precision, efficiency, and effectiveness of the search methodologies are important, perfection is neither required nor possible to achieve.

---

<sup>†</sup>Copyright © 2019, Duke Law Bolch Judicial Institute, All Rights Reserved. This document does not necessarily reflect the views of Duke Law School or its faculty, or any other organization including the Judicial Conference of the United States or any other government unit.

<sup>1</sup> The protocol is an exemplar. Parties may adjust individual provisions to address the circumstances of their case, particularly the time periods within which to act. The protocol is intended to be used in conjunction with the EDRM/Duke TECHNOLOGY ASSISTED REVIEW (TAR) GUIDELINES, Bolch Judicial Institute, Duke Law School (January 2019) posted at \_\_\_\_\_ . Footnotes are not part of the protocol and are provided solely for explanatory purposes.

<sup>2</sup> Surprisingly for many, the percentage in most cases can be high, e.g., more than 20%, regardless of the search methodology, including human review. Keyword searches have been known to miss more than 70%.

The parties agree that the discovery standards in the Federal Rules of Civil Procedure, concerning cooperation and engaging in a reasonable and proportionate process, also apply to the use of TAR. The parties recognize that cooperation and transparency concerning the process for searching for potentially relevant ESI and the method for measuring the effectiveness of that process can avoid unnecessary disputes and facilitate discovery provided that non-responsive material and attorney work product or other privileged material is not required to be disclosed.<sup>3</sup> The parties remain responsible for how they conduct discovery.

## **I. TAR PROCESS**

The parties agree to use the TAR process as follows:

### **A. Description of the TAR Process**

Although the processes followed by different TAR software employ different workflows, each includes: (1) assembling the TAR team; (2) collecting and analyzing the TAR set; (3) “training” the TAR software to classify documents; (4) applying quality-control measures; and (5) completing the process and validating results. The producing party will disclose in advance any significant departures from these procedures, which are generally described in the EDRM-DUKE TAR GUIDELINES ([HTTPS://WWW.EDRM.NET/WP-CONTENT/UPLOADS/2019/02/TAR-GUIDELINES-FINAL.PDF](https://www.edrm.net/wp-content/uploads/2019/02/TAR-GUIDELINES-FINAL.PDF)), for the particular TAR software that is being used.

The producing party will use one or more attorneys who are familiar with the claims and defenses and who will consult with the TAR software provider or vendor as needed. If at any point during the TAR process, the producing or receiving party suspects the TAR process is not

---

<sup>3</sup> Although the Federal Rules of Civil Procedure may not require the disclosure of non-responsive, non-privileged documents, such disclosure of all or a random sample may streamline and improve the TAR process and will significantly strengthen the requesting party’s confidence in the review and provide greater assurance to the court.

functioning properly, the parties should meet and confer regarding the potential problems and proposed resolutions.

## **B. Identification of Documents Subject to TAR**

The requesting party must describe with reasonable particularity each item or category of items to be produced. The producing party will provide the following information, as well as any later modifications, about the TAR process.<sup>4</sup>

1. **TAR Software.** The producing party will disclose the name of the TAR software, the software provider or vendor, and the person(s) administering/managing the TAR process.
2. **Data Sources and Custodians Included in TAR Review.** The producing party will disclose sources of data that will be subject to review using TAR<sup>5</sup> and the names and titles of custodians, including the custodian's group or department if not reasonably apparent from the job title. On timely request, the producing party will provide information sufficient for the receiving party to evaluate the responsibilities of a reasonable number of the producing party's custodians if not otherwise apparent.
3. **Date-Range Restrictions.** The producing party will disclose the date range(s) as well as any specific date filters that will be applied to documents subject to the TAR process.
4. **Global De-Duplication across Custodians/Sources.** Unless provided under a general ESI discovery protocol, the producing party will disclose the de-duplication processes for documents that will be subject to the TAR process, including: (i) global de-duplications based on an industry-accepted hash algorithm, i.e., either MD5 or SHA-1 hash value, and, in the case of email and email families, the fields concatenated to form the value to which the hash algorithm will be applied; (ii) e-mail threading; and (iii) other structured analytics or other de-duplication techniques before applying the TAR process.
5. **Other Filtering.** The collection of documents subject to the TAR process will be DeNIST'ed, and any system or below 200 byte or above 30 mb files removed before applying the TAR process.
6. **Excluded Files and Documents.** The producing party will disclose the criteria for excluding documents and a general description of the files, types, or sources of documents (e.g., audio, excel spreadsheets, foreign language, text messages, chat logs, and call history) that will be excluded from the TAR review.<sup>6</sup>
7. **Culling Documents by Keyword Searches and Other Techniques.** The producing party will disclose the terms of any method it plans to use to cull documents subject to the TAR process, including the use of keyword searches and the keyword terms, and, on request,

---

<sup>4</sup> There are different opinions as to what level of disclosure, if any, is required under the Federal Rules of Civil Procedure. Nonetheless, the disclosure does not create a new burden, will significantly strengthen the requesting party's confidence in the review, and provide greater assurances to the court.

<sup>5</sup> Potentially responsive documents are not necessarily in the possession of a specific custodian. In some types of litigation, responsive documents may be stored in a specific data source rather than by a custodian, such as employee contracts or technical documents.

<sup>6</sup> Of course, responsive documents excluded from the TAR review remain subject to discovery obligations and must be reviewed using some other search and review method for responsiveness.

will provide an explanation of the reasons justifying such culling.<sup>7</sup> The parties understand that the culling methods and terms used on TAR documents should be developed in good faith in a manner that will not attempt to prevent the TAR process from identifying potentially responsive documents. If the producing party uses search terms, it will sample the results, including sampling documents that are not captured by those search terms, refine the search as needed, disclose adjustments to the terms, and, in addition, separately validate the results of the combined search terms and TAR process, e.g., by drawing an additional validation sample from the entire collection to which the search terms were applied.

If keyword searches are used to cull documents, the parties understand and agree that only the resulting documents will be subject to the TAR process. The producing party is under no obligation to review every document subject to the TAR process, which may include documents identified by the initial keyword searches but determined by the TAR process to be likely nonresponsive.

8. Confidence Level and Interval. Unless the producing party discloses otherwise, it will use a confidence level of 95% and a confidence interval of 2.5 to calculate recall.<sup>8</sup>

### **C. Description of Responsive-Review Criteria during TAR Process**

The producing party will prepare a document identifying the responsiveness criteria, including a list of the main issues (i.e., information that it considers responsive to the discovery requests and relevant to claims and defenses) that will be used as its guide in coding and reviewing documents for responsiveness during the TAR training or review process.<sup>9</sup> The producing party will provide the document and updates of modifications to the requesting party. Disclosure of the responsiveness-criteria document does not obligate the producing party to produce any other documents that also contain information about responsiveness criteria (e.g., all reviewer-training documents).

---

<sup>7</sup> The explanation is intended to preempt disputes about the breadth of the culling. Disputes about the keyword terms are addressed in section “D. Modification of TAR Process.” Many parties cull documents using keyword terms before applying TAR, but objections have been raised that no culling should be permitted under TAR without both parties’ agreement.

<sup>8</sup> These are provided merely as default measurements. The circumstances of each case may call for different measurements. Parties must carefully examine and set the measurements appropriate in the case.

<sup>9</sup> The Federal Rules of Civil Procedure may not require such a disclosure. Nonetheless, the disclosure does not create an onerous burden, will significantly strengthen the requesting party’s confidence in the review, and provide greater assurances to the court. But this provision could potentially lead to tangential discussions if the parties have not agreed on the scope of discovery. The parties may need to seek the court’s guidance to determine the proper scope of discovery prior to commencing any discovery review and production to avoid later disputes.

The producing party will also disclose whether exemplar, strategic samples, or randomly selected documents will be used to begin training the TAR software as well as the total number of documents used for training. The requesting party may suggest exemplar samples of documents to be used for training purposes.

Neither the disclosure of the responsiveness-criteria document nor the disclosure addressed immediately above waives the producing party's attorney work-product protections or any privilege that would attach to any other responsive information in the case.

#### **D. Modification of Proposed TAR Process**

Within 14 days of the producing party's disclosures of the terms of the TAR process or within a time agreed upon by the parties, the requesting party will submit any reasonable and proportionate modifications. Within 14 days after receiving any proposed modifications, the parties will meet and confer to attempt to resolve any outstanding disagreements. If the dispute is not resolved within 14 days, the parties will jointly seek resolution by the court.

#### **E. Modification of Ongoing TAR Process**

If the requesting party identifies a new claim or defense or a significant new fact or issue after the TAR process has been initiated, the producing party shall take and disclose additional steps to train or otherwise refine the TAR process, or employ an alternative search methodology, which targets the new claim or defense or significant fact, taking into account the Rule 26(b)(1) proportionality factors. Within 14 days after receiving any proposed modifications, the parties will meet and confer to attempt to resolve any outstanding disagreements. If the dispute is not resolved within 14 days, the parties will jointly seek resolution by the court.

#### **F. Validation/Quality Control**

The producing party shall validate the results of the TAR process, and also, if search terms were applied, the results of the TAR process and keyword search terms combined. TAR software generates, or allows for the generation of, metrics or effectiveness measures, which allow the producing party to evaluate the results of the TAR process. At a minimum, the producing party should evaluate the recall percentage of the TAR set, and also, if search terms were applied, the recall percentage of the combined result set. The parties may consider whether to use other effectiveness measures in addition to recall. These methods are not mutually exclusive. The producing party will also analyze responsive documents “missed” by the TAR process, and also, separately, if search terms were applied, missed by the search terms, to determine whether they contain relevant, unique content sufficiently different from produced documents as to demonstrate that the TAR process, or search terms, was deficient for categories of documents (e.g., TAR or search terms missed significant documents relating to a specific issue, or the strongest evidence for the case). The producing party will separately disclose “missed,” responsive documents.

The producing party will produce all responsive, non-privileged documents subject to any different agreement between the parties. If the validation efforts fail to demonstrate the desired level of recall, or demonstrate a deficient process, the producing party shall take and disclose additional steps to train or otherwise refine the TAR process or shall provide a sufficient justification for not doing so. If the parties cannot resolve validation disagreements within 28 days after disclosure of the TAR results and production of validation documents, the parties agree to stop discussions and jointly seek resolution by the court.

#### **G. Disclosure of the TAR Process Results**

After the producing party has validated the TAR process results, it will promptly disclose the following information:<sup>10</sup>

1. The total number of documents subject to the TAR process;
2. The final estimated richness of the TAR set both at the outset and at the completion of the TAR process and the sampling method used to estimate richness;
3. The estimated recall for the TAR set, with corresponding confidence interval;
4. The total number of documents in the predicted relevant set (not family complete);
5. The total number of documents in the predicted nonrelevant set (null set) (not family complete); and
6. The total number of documents reviewed for validation,<sup>11</sup> with a breakdown of the total number of responsive and nonresponsive documents identified.<sup>12</sup>

If search terms were used, the producing party will also disclose:

1. The total number of documents in the collection to which search terms were applied;
2. The number of documents in the collection to which the search terms were applied which were excluded from the TAR process because they did not contain a search term.

#### **H. Review and Quality Assurance of TAR Results**

After the requesting party has reviewed the documents produced as a result of the TAR process, the producing party on request will undertake an additional search limited to addressing specific deficiencies if the requesting party establishes: (1) that specific categories or types of responsive and proportional information were found to be omitted from production, e.g., emails from a particular individual during a specific time period; minutes of meetings covering particular time periods; or (2) specific, critical responsive information was omitted from production.<sup>13</sup>

---

<sup>10</sup> The Federal Rules of Civil Procedure may not require such disclosures. Nonetheless, the disclosures do not create new burdens, will significantly strengthen the requesting party's confidence in the review, and provide greater assurances to the court.

<sup>11</sup> A control set is sufficient for validation, provided other quality control is done to ensure the quality as well as quantity of "missed" responsive documents is acceptable, as described herein.

<sup>12</sup> Privileged documents included in the responsive subset are not subject to production.

<sup>13</sup> The requesting party always has the option to file a motion to compel if it finds the search or production inadequate, even after the producing party undertakes the additional limited search and production. The provision as a whole is intended to preempt unnecessary disputes about requests for a later, *limited* supplemental search and incentivize requesting parties to prepare initial, targeted discovery requests, confident that a second opportunity is available.

- The parties will meet and confer within 14 days after receiving the request for the follow-up information or search to identify the designated types of documents and information. The producing party must take reasonable steps to address and, if possible, to correct the identified deficiencies, refining its use of TAR or employing an alternative search methodology, which targets the information at issue, taking into account the proportionality factors in Rule 26(b)(1).
- The producing party will produce any responsive, non-privileged results of the follow-up search within 28 days after the parties enter into an agreement for additional steps in accordance with the meet-and-confer process above. If the parties cannot resolve the disagreement within 28 days after the follow-up disclosures, the parties agree to stop discussions and jointly seek resolution by the court.
- If a reasonable amount of time is not available to complete an additional search and review of the results as provided under this provision before a discovery-deadline set by the court expires, the parties agree to jointly request the court to extend the time.

### **I. No Waiver of Any Right to Serve Additional Discovery Requests**

Nothing in the TAR protocol prevents a party from serving additional document requests or objecting to requests, consistent with the Federal Rules of Civil Procedure, any local rules or standing order of court, and any discovery or case-management order issued by the presiding judge.

### **II. DEFINITIONS**

- **CONFIDENCE INTERVAL (MARGIN OF ERROR) AND CONFIDENCE LEVEL.** The confidence interval and confidence level characterize the certainty of the point estimate.<sup>14</sup> For example, the recall point estimate of 80% can be combined with a margin of error of 5%, allowing for a confidence interval of 75% (5% below 80%) and 85% (5% above 80%). Moreover, a confidence interval is meaningful only if accompanied by a confidence level, which is a measure of how likely the sample is to represent the true set, within the confidence interval. Continuing the example used here, a confidence level of 95% means that 95 times out of 100, the result of the sample will include the actual recall within its confidence interval. Put another way, there is a 95% chance that the actual recall is between 75% and 85%.
- **CONTROL SET.** A control set is a random sample taken from the entire TAR set that acts as a relevancy truth set against which the computer's decisions can be judged. It is used to estimate the computer's effectiveness in classifying documents during TAR. It may also be used to estimate the richness of the TAR set. Not all workflows use a control set.

---

<sup>14</sup> **POINT ESTIMATE.** A point estimate is an estimate that is a single value. Based on the recall definition example below, the point estimate for recall is the single value of 0.80 (80%), since the computer correctly identified 80,000 of the 100,000 total relevant documents. However, as provided in the confidence interval and level definitions, a point estimate alone is of limited use, and therefore should be combined with how confident we are in the point estimate.



- **ELUSION.** Elusion estimates how many relevant documents were missed and are in the predicted nonrelevant set. In the example used below in the recall definition, the computer identified 800,000 documents as potentially nonrelevant. Because there are a total of 100,000 relevant documents and 80,000 documents were identified within the 100,000 potentially relevant documents, 20,000 relevant documents were potentially missed. The elusion of the TAR predictive model is therefore  $20,000 / 800,000 = 0.025$  or 2.5%.
- **ESTIMATE OR ESTIMATION.** Knowing the exact value of an effectiveness measure (such as recall) would require knowing the true relevancy status of every document in the TAR set. In practice, therefore, one must estimate the effectiveness using sampling techniques. These estimates allow for a statistical certainty that the estimated values are close to the true value.
- **PRECISION.** Precision measures the percentage of documents that are truly relevant among all the documents the TAR software identified as potentially relevant. Using the example in the recall definition, the TAR software identified 200,000 documents as potentially relevant, of which 80,000 were identified as relevant by human-review, resulting in a precision of 40% (80,000/200,000).
- **PREDICTED NONRELEVANT SET.** The predicted nonrelevant set is a subset of documents in the TAR set. It contains those documents in the TAR set that are predicted as nonrelevant by the software and thus would be excluded from further review or production efforts workflow.<sup>15</sup>
- **PREDICTED RELEVANT SET.** The predicted relevant set is a subset of documents in a TAR review set. These are the documents that the software identifies as relevant and subject to potential production as a result of the TAR process. No matter what form of TAR used, the identification of the potential production set is derived from the software's predictions regarding what is relevant and non-relevant. In some workflows, the entire predicted relevant set is reviewed by humans during the TAR training process. For example, in common TAR 2.0 workflows, the software is only trying to return relevant documents to the humans, and the humans review all of the documents returned by the computer as predicted relevant). In other workflows, the predicted relevant set is not reviewed in its entirety during the TAR training process. For instance, in common TAR 1.0 workflows, the TAR process is designed to build a predictive model to make relevancy calls on the entire TAR Set, and after TAR is complete, the human review team makes the decision to review the entire relevant review set or to simply accept the software's relevancy decisions. In any event, documents predicted to be relevant can be subsequently reviewed and determined to be relevant or nonrelevant under both TAR 1.0 or TAR 2.0 workflows. Despite no longer being a "prediction" at that point, these documents continue to be classified as part of the "predictive relevant set" to eliminate confusion that would otherwise arise.

With this in mind, it is important to note that, like manual reviews, TAR classifications are not perfect. The "predicted relevant set" will not contain all of the relevant documents from the TAR set: its recall will not be 100%. Nor will it contain only relevant documents: its precision will not be 100%. Any documents in the predicted relevant set that are subsequently determined to be non-relevant by a human reviewer can always be excluded from production (insofar as they are not part of a family that includes relevant documents).

---

<sup>15</sup> Just as there will be nonrelevant documents in the predicted relevant set, there will be some estimated number of relevant documents in the "predicted nonrelevant set." But, for simplicity purposes, we identify this as the predicted nonrelevant set because most of these documents have been identified by the computer as nonrelevant, and thus will be excluded from further human review.

- **RECALL.** Recall measures the percentage of documents found to be relevant. Consider a workflow in which a TAR set of one million documents are collected, of which 100,000 are relevant.<sup>[2]</sup> The TAR software identifies 200,000 documents as potentially relevant and 800,000 documents as potentially nonrelevant. A human review of the 200,000 potentially relevant documents shows that 80,000 are relevant. Therefore, the effectiveness of the classification system, when measured using recall is 80%, since the TAR software identified 80,000 of the 100,000 relevant documents. The producing party may represent that their workflow achieved an 80% recall, i.e., the documents being produced represent 80% of the relevant population prior to any possible privilege review.
- **QUALITY CONTROL.** During a document review, the team may engage in quality control efforts to ensure the human reviewer and computer's relevancy decisions are as accurate as reasonably possible.
- **RICHNESS.** Richness (or prevalence) is the estimated proportion of documents in a data set that are relevant. For example, if a set of one million documents contains 100,000 relevant documents, it has 10% richness. Richness is also known as prevalence.
- **TAR SET.** This is the total set of documents that the workflow (the document review) will be conducted on.
- **TRAINING SET.** The training set is the subset of documents in the TAR set that the human reviewer(s) reviews to teach the software what is relevant. The training set will contain relevant and nonrelevant documents. The TAR software uses the training set to produce a predictive model, and the predictive model will be used to define the predicted relevant set. The number of relevant and nonrelevant documents necessary to produce a predictive model with good effectiveness will depend on the nature of the documents in the TAR set, the difficulty of the relevance definition, and the particular TAR software and method used.

---

<sup>[2]</sup> In order to estimate recall, the total number of relevant documents in the TAR set must be known. Because the only way of identifying the total number of relevant documents in a set is to review the entire TAR set, the total number of relevant documents must also be estimated.