DRAFT

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

BOLCH JUDICIAL INSTITUTE, DUKE LAW SCHOOL
May 2019[†]

The parties agree that the following protocol[1] will govern the use of technology assisted review (TAR) to identify and produce electronically stored information that is relevant to a party's claim or defense.

GENERAL PROVISIONS

TAR is intended to assist in the just, speedy, and inexpensive conduct of discovery relating to ESI in accordance with Rule 1 of the Federal Rules of Civil Procedure. The parties agree that, whatever review method they use to identify and review documents for responsiveness to a discovery request, the process should not be designed to exclude relevant information, but should yield results that are reasonable and proportionate. The parties recognize that any search methodology they use, including manual human review, will miss some responsive documents. The parties further recognize that although precision, efficiency, and effectiveness of the search methodologies are important, perfection is neither required nor possible.

The parties agree that the discovery standards in the Federal Rules of Civil Procedure, on engaging in a reasonable and proportionate process, also apply to the use of TAR. The parties

---

[1] 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 _____.

recognize that cooperation and transparency about the process used to search for relevant ESI and the method used to measure the effectiveness of that process can avoid unnecessary disputes and facilitate discovery, can minimize the disclosure of nonresponsive or work product or privileged material, and can prevent untoward consequences – such as forfeiture or waiver – if disclosure does occur. The parties remain responsible for how they conduct discovery.

I. TAR PROCESS

The parties agree to use the following TAR process:

A. Description of the TAR Process

Although different TAR software use different workflow processes, 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 the results. These are generally described in the EDRM-DUKE TAR GUIDELINES (HTTPS://WWW.EDRM.NET/WP-CONTENT/UPLOADS/2019/02/TAR-GUIDELINES-FINAL.PDF) for the TAR software that is being used. The producing party will disclose in advance any significant departures from these processes.

The producing party will use one or more attorneys who are familiar with the claims and defenses and the TAR protocol. If at any point during the TAR process, the producing or receiving party suspects the TAR process is not 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 will describe with reasonable particularity each item or category of relevant items to be produced. The producing party will provide the following information, as well as later modifications, about the TAR process.

1. TAR Software. The producing party will disclose the name of the TAR software, the software provider or vendor, and the person(s) administering or managing the TAR process.
2. Data Sources and Custodians Included in TAR Review. The producing party will disclose sources of custodial and noncustodial data that will be subject to review using TAR and the names and titles of custodians, including the custodian's group or department if that is not reasonably apparent from the job title. On timely request, the producing party will provide sufficient information for the receiving party to evaluate the roles and 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 it intends to apply to documents subject to the TAR process.
4. Global Deduplication across Custodians and Sources. Unless provided under a general ESI discovery protocol, the producing party will disclose the deduplication processes for documents that will be subject to the TAR process, including: (i) global deduplications based on an industry-accepted hash algorithm, e.g., MD5 or SHA-1 hash value, and, in the case of email and email families, the fields attached to form the value to which the hash algorithm will be applied; (ii) e-mail threading; and (iii) other structured analytics or other deduplication techniques used before applying the TAR process.
5. Other Filtering. The collection of documents subject to the TAR process will be DeNIST'ed.
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), and any system files, embedded object, or below-200 byte (too small) or above-30 mb (too large) files removed before applying the TAR process.
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. The culling methods and keywords used on TAR documents will be developed in good faith in a manner that will not attempt to prevent the TAR process from identifying 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 appropriate, and disclose adjustments to the terms.

   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 5.0 to calculate recall.

C. Description of Responsive-Review Criteria

The producing party will disclose 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 but not privileged or work-product) that will be used as its guide in coding and reviewing documents for responsiveness during the TAR training or review process. Disclosure of the responsiveness-criteria does not obligate the producing party to produce other documents that also contain information about responsiveness criteria (e.g., all reviewer-training documents).

The producing party will also disclose the process it intends to use to train the TAR software (e.g., TAR 1, TAR 2, hybrid, or other variations) as well as the total number of documents used for training (e.g., seed set for TAR 1; number of documents reviewed for TAR 2). If the TAR process uses a seed set for training, the producing party will also disclose the documents used in the seed set, other than privileged and work-product documents. The requesting party may suggest exemplars of documents to be used for training purposes.

Neither the disclosure of the responsiveness-criteria document nor their disclosure waives the producing party's attorney work-product protection or privilege that applies to the responsive information.

D. Receiving Party's Response to Terms of TAR Process

Within 14 days of the producing party's disclosures of the TAR process terms or within a time the parties agree to, the requesting party will submit any reasonable modifications, including modifications of keywords used to cull the documents. Within 14 days after receiving the proposed modifications, the parties will meet and meaningfully confer to try to resolve any outstanding disagreements. Under Section I, the court will take into account the responding

party's response to the proposed modifications if the TAR results omitted significant responsive and relevant information.

E. New Claim or Defense Requiring 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 begun that has not been adequately searched for, the producing party will take and disclose additional steps to train or otherwise refine the TAR process, or use an alternative search methodology that targets the new claim or defense or significant fact, taking into account the Rule 26(b)(1) proportionality factors. Within 14 days after receiving notice of a new clam or defense and any proposed modifications, the parties will meet and meaningfully confer to try to resolve any outstanding disagreements. Under Section I, the court will take into account the responding party's response to the proposed modifications if the TAR results omitted significant responsive and relevant information.

F.   Validation, Quality Control, and Production

The producing party will statistically validate whether the results of the TAR process demonstrated the desired level of recall and disclose the precision percentage. 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 statistically sample and analyze responsive documents that the TAR process omitted (null set) to determine whether the documents contain significant relevant information sufficiently different from the information produced as to demonstrate that the TAR process was deficient for certain categories of documents. Examples are significant information in documents relating to a specific issue or to important evidence in the case. The producing

party will separately produce the responsive, nonprivileged or non-work-product documents found in the null set.

The producing party will produce responsive, nonprivileged and non-work-product 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 will take and disclose additional steps to train or otherwise refine the TAR process or will provide a sufficient justification for not doing so.

G. Disclosure of the TAR Process Results

After the producing party has validated the TAR process results, it will promptly disclose the following information:

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 the corresponding confidence interval;
4. the total number of documents classified as relevant (not family complete);
5. the total number of documents classified as nonrelevant (null set) (not family complete); and
6. the total number of documents reviewed for validation, with a breakdown of the total number of responsive and nonresponsive documents identified.

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

H. Request for Additional Search

If the requesting party requests an additional search to provide additional information reasonably limited to addressing deficiencies, the producing party will undertake the search if the requesting party establishes that: (1) specific categories or types of responsive and proportional information were found to be omitted from production, such as emails from a particular

individual during a specific time period or minutes of meetings covering particular time periods; or (2) significant responsive information, which can be described with reasonable particularity, was found to be omitted from production.

The parties will meet and meaningfully confer within 14 days after receiving the request for the additional search or information to identify the designated types of documents and information. The producing party will take reasonable steps to address and, if possible, to correct the identified deficiencies, by refining its use of TAR or using an alternative search methodology that targets the information at issue, taking into account the proportionality factors in Rule 26(b)(1). Under Section I, the court will take into account the responding party's response to the proposed additional search if the TAR results omitted significant responsive and relevant information.

The producing party will produce any responsive, non-privileged, non-work-product results of the additional search or information within 28 days after the parties enter into an agreement for the additional search after meeting and meaningfully conferring to reach agreement. 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. Disputes about TAR Results

If the TAR results omitted significant responsive and relevant information, the requesting party may request the court to resolve the dispute. In resolving a dispute about the TAR results, the court should take into account: (1) an earlier request of the requesting party under section D to modify the terms of the proposed TAR process and the responding party's response: (2) an earlier request of the requesting party under section E to modify the ongoing TAR process and

the responding party's response; and (3) a request by the requesting party under section H for an additional search and the responding party's response.

Alternatively, the parties can agree jointly to appoint a special master to sample a reasonable number of documents randomly selected from the TAR null set and determine whether the TAR's results omitted significant responsive and relevant information. If the special master finds that the responding party reasonably conducted the TAR process and significant responsive and relevant information was not omitted, the requesting party will pay the special master's full compensation. If the special master finds that the responding party did not reasonably conduct the TAR process and significant information was omitted, the responding party will pay the special master's full compensation and will conduct an additional search. The special master may also impose additional costs on the responding party – in addition to but not to exceed the amount of the special master's compensation -- after taking into account the responding party's failure to address or adopt earlier requests by the requesting party to modify the TAR process pursuant to sections D, E, or H.

J.   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, a court's local rules or standing order, 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.[2] For example, the recall point

---

[2] 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

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.

•   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 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.[3]

•   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

---

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.

[3] 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.

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).

- 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.[2] 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.

---

[2] 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.