PROBLEM: Nugget-based evaluations require exhaustive matching of nuggets against system responses. Nuggets have various representations in text. Example: Given nuggets \( \{N\} \) and returned (pooled) sentences \( \{D\} \), assessors need to match each nugget with every sentence.

\[
\begin{array}{c|c|c}
|N| & \{D\} & N \times D \\
\hline
TS13 & 119.7 & 1,283,298 \\
TS14 & 92.9 & 1,407,448 \\
\end{array}
\]

\*as confirmed by NIST assessors for the Temporal Summarization track 2013 (TS13).

SOLUTION: Use Continuous Active Learning for Technology-Assisted Review of candidate matches. MOST LIKELY CANDIDATE matching strategy

- Train a classifier for each nugget in a topic
- Present the highest scoring nugget-sentence pair for review
- Review L match candidates; re-train respective nugget classifiers after each review decision

COVER MAXIMUM CANDIDATES matching strategy

- Score all sentences w.r.t each classifier
- Present the sentence that likely matches multiple nuggets for review
- Review M match candidates for the sentence; re-train respective nugget classifiers on review

EVALUATION VIA SIMULATION: Compute effort over simulated assessment interfaces for strategies.

- \( \text{effort} = \lambda_{\text{read}} + \lambda_{\text{match}} \)
- \( \text{effort} = \lambda_{\text{read}} + \lambda_{\text{match}} + (M - 1) \lambda_{\text{match}}' \)

ACKNOWLEDGMENT: Support provided in part by a CIKM’16/SIGIR Travel Grant.