The Sedona Conference Glossary: eDiscovery and Digital Information Management - Proposed AI-related Terms and Definitions (2024)

The Sedona Conference Technology Resource Panel

The Sedona Conference Glossary: eDiscovery and Digital Information Management - Proposed AI-related terms and Definitions

The following are terms and definitions related to Artificial Intelligence (AI) that are either currently included or proposed for the next edition of *The Sedona Conference Glossary: eDiscovery and Digital Information Management*, currently available at https://thesedonaconference.org/publication/The_Sedona_Conference_Glossary (log-in required).

Active Learning or Active Machine Learning: Technology-assisted-review algorithm for the selection of training documents, in which the machine selects sets of additional documents that should best improve results beyond the training that has already been done. Compare to Passive Learning. See Machine Learning.

Artificial Intelligence (AI): A subfield of computer science focused on the development of intelligence in machines so that the machines can react and adapt to their environment and the unknown. AI is the capability of a device to perform functions that are normally associated with human intelligence, such as reasoning and optimization through experience. It attempts to approximate the results of human reasoning by organizing and manipulating factual and heuristic knowledge. Areas of AI activity include expert systems, natural language under-standing, speech recognition, vision, and robotics. See Ma-chine Learning.

Bayesian Search: An advanced search that utilizes the statistical approach developed by Thomas Bayes, an 18th century mathematician and clergyman. Bayes published a theorem that describes how to calculate conditional probabilities from the combinations of observed events and prior probabilities. Many information retrieval systems implicitly or explicitly use Bayes's probability rules to compute the likelihood that a document is relevant to a query. For a more thorough discussion, see The Sedona Conference, Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, 15 SEDONA CONF. J. 217 (2014), available at https://thesedonaconference.org/publication/Commentary_on_Search_and_Retrieval_Methods.

Bias: As used in Artificial Intelligence: systematic errors or prejudices in the outputs produced by AI systems. It can be categorized into two types: intentional and unintentional bias. (from the forthcoming *Artificial Intelligence in the Practice of Law)*. See also Intentional Bias and Unintentional Bias.

Binomial classifier: an algorithm that identifies data into one of two categories or models; i.e. Hot or Not Hot. See Classifier, Multinomial Classifier.

ChatGPT: is a Generative AI tool developed by OpenAI.

Classifier: An algorithm that puts data into defined categories or classes. Classifiers can be either Unsupervised or Supervised. See Supervised and Unsupervised learning.

Clustering: Unsupervised machine learning in which thematically similar files are grouped together based on the text of the individual files.

Concept Search: The method of search that uses word meanings and ideas, without the presence of a particular word or phrase, to locate electronically stored information related to a desired concept. Word meanings can be derived from any of a number of sources, including dictionaries, thesauri, taxonomies, and ontologies, or computed mathematically from the context in which the words occur.

Conceptual Analytics: Using one or more of a number of mathematical algorithms or linguistic methodologies to analyze unstructured data by themes and ideas contained within the documents, enabling the grouping or searching of documents or other unstructured data by their common themes or ideas.

Confidence Interval: The range of values that is likely to contain the true parameter for a population to the specified confidence level. For example, sampling a set of documents at a 95 percent confidence level with an interval of plus-or-minus 2 percent means that 95 per-cent of samples will produce a result within 2 percent of the actual population.

Confidence Level: The percentage of samples for which the results are expected to correctly describe a population parameter within a provided confidence interval. For example, sampling a set of documents at a 95 percent confidence level means that 95 percent of samples taken from the population would contain the correct result within a specified interval. See Margin of Error.

Contextual Search: Using one of a number of mathematical algorithms or linguistic methodologies to enlarge search results to include not only exact term matches but also matches where terms are considered in context of how and where they frequently occur in a specific document collection or more general taxonomy. For example, a search for the term “diamond” may bring back documents related to baseball but with no reference to the word diamond because the term frequently occurs within similar documents and therefore has a logical as-sociation.

Continuous Active Learning (CAL): A machine-learning algorithm that periodically analyzes users' decisions in order to rank unreviewed data, with the most likely desired data ranking first based on the users' previous decisions. See also Technology-Assisted Review.

Data Categorization: The process of classifying electronically stored information with supervised machine learning software, using categories created by either the user or automatically by the software based on the similar content of the individual files.

Decision Tree: In Artificial Intelligence, a type of supervised classifier algorithm that uses "if-then" logic to segregate data into specific hierarchical categories. For example: a real estate website could use a decision tree classifier to divide houses from apartments, then stand-alone homes versus town homes, and so on.

Deep Learning: involves a series of layered algorithms, starting with an input layer, followed by some number of hidden layers, and then an output layer. Each layer performs a different function and passes certain information on to the next layer. (from the forthcoming *Artificial Intelligence in the Practice of Law)*. See also Machine Learning.

Deepfake: A term used to define the output of a deep learning algorithm that synthesizes an image or video rather than an image or video which reflects an accurate recording of a real person or thing.

Early Case Assessment (ECA): The process of assessing the merits of a case early in the litigation lifecycle to determine its viability. The process may or may not include the collection, analysis, and review of data.

Early Data Assessment (EDA): The process of separating possibly relevant electronically stored information from nonrelevant electronically stored information using both computer techniques, such as date filtering or advanced analytics, and human-assisted logical determinations at the beginning of a case. This process may be used to reduce the volume of data collected for processing and review. See also Early Case Assessment.

Elusion: The percentage of documents of a search's null set that were missed by the search, usually determined with re-view of a random sample of the null set. The elusion rate can be multiplied by the number of documents in the null set to estimate how many documents were missed by the search.

F-Measure: Also known as the F1 Score or the F Score, a measure of a search's accuracy calculated by using precision and recall. (Precision x Recall) / (Precision + Recall).

Generative AI: a form of technology that uses algorithms to analyze the composition of a corpus to then predict or generate results based on a user prompt using the patterns identified in the original corpus. Generative AI can be used to generate text, images, 3d models or sound.

General AI: Also known as Strong AI is the branch of AI that would be able to understand, learn, adapt, and implement knowledge on a wide range of tasks, in a manner akin to a human being. (from the forthcoming *Artificial Intelligence in the Practice of Law)* See also, **Narrow AI.**

Hallucination: an error in Generative AI output based upon inaccurate analysis of the original set or model of data by identifying patters or relationships that do not exist in the data.

Intentional Bias: occurs when an AI system is deliberately designed to favor certain outcomes over others. (from the forthcoming *Artificial Intelligence in the Practice of Law)*. See Bias.

Judgmental Sampling: The human selection of a subset of documents from a larger population based on some logical criteria, such as search-term hits or the searcher's own experience and knowledge.

K-Nearest Neighbors (k-NN): an algorithm that is used to recognize patterns in a data set and predict the likelihood of associated words unprocessed data. For example, if John appears next to Doe repeatedly in a data corpus, the algorithm will associate John with Doe as closely related and can be used to classify the data in unprocessed data sets even if they don't appear together.

Large Language Model (LLM): A model built using a large set of data with the purpose of analyzing and predicting accurate results based on human language, which by its very nature is extremely complex. For example, the feature of predicting the next word in a sentence is based on a large language model.

Machine Learning: A subset of artificial intelligence enabling a system to automatically improve at a task on its own based upon experience and data, without being explicitly programmed for that task. See Artificial Intelligence, Unsupervised Learning, Supervised Learning, Reinforcement Learning and Deep Learning.

Model: As used in Artificial Intelligence discussions, a model is the data set result of a training process.

Multinomial classifier: an algorithm that identifies data into one or more categories or models; i.e. Issue 1, Issue 2 and\or Issue 3. See Classifier; Binomial Classifier.

Naïve Bayes Classifier: a text-based, probabilistic algorithm that is used to predict the likelihood that data can be classified into a specific model by using the logic that if data falls is identified as one category, it is likely to be in a second category. For example, if a string of text in a published decision reads "convicted of assault" then the case could be categorized as criminal as opposed to civil.

Narrow AI: Also known as Week AI is the branch of AI systems that are designed to perform a specific task and operate under a limited set of constraints. (from the forthcoming *Artificial Intelligence in the Practice of Law)* See also, General AI.

Natural Language Processing (NLP): the branch of AI focused on providing computers with the ability to interpret, understand, and generate language by pre-processing data into smaller units, like the root

form of a word and removing valueless words, like "the" "a" and "an" (paraphrased from the forthcoming *Artificial Intelligence in the Practice of Law)*

Natural Language Search: A manner of searching that permits the use of plain language without special connectors or precise terminology, such as "Where can I find information on William Shakespeare?" as opposed to formulating a search statement, such as "information" and "William Shakespeare." See Boolean Search.

Null Set: A set of files that are not positive results of a search.

Null Set Testing: Sampling a null set to search for false negatives of the search that created the null set.

Neural Network: In its most simple form, a collection of algorithms used to solve a problem based on how the human brain functions.

Ontology: A collection of categories and their relationships to other categories and to words. An ontology is one of the methods used to find related documents when given a specific query.

Passive Learning. A technology-assisted review workflow in which documents are randomly selected for training by hu-man review. See also Active Learning.

Pattern Recognition: Technology that searches electronically stored information for like patterns and flags and extracts the pertinent data, usually utilizing an algorithm. For instance, in looking for addresses, alpha characters followed by a comma and a space, followed by two capital alpha characters, followed by a space, followed by five or more digits, are usually the city, state, and zip code. By programming the application to look for a pattern, the information can be electronically identified, extracted, or otherwise utilized or manipulated.

Precision: When describing search results, precision is the number of true positives retrieved from a search divided by the total number of results returned. For example, in a search for documents relevant to a document request, it is the percentage of documents returned that are actually relevant to the request. See The Sedona Conference, Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, 15 SEDONA CONF. J. 217 (2014), available at https://thesedonaconference.org/publication/Commentary_on_Search_and_Retrieval_Methods.

Prevalence: The percent of a population that has a specific characteristic, such as responsiveness.

Prompt: As used in generative AI: a natural language search submitted to a large language model that can be comprised of any string of text that the model will then use to predict a response.

Proximity Search: A search syntax written to find two or more words within a specified distance from each other.

Random Sampling: The process of selecting data from a population with no bias or input from the person performing the sampling, in which each item has an equal chance of being selected as any other item. See also Sampling.

Recall: When describing search results, recall is the number of documents retrieved from a search divided by all of the responsive documents in a collection. For example, in a search for documents relevant to a document request, it is the per-centage of documents returned compared against all documents that should have been returned and exist in the data set. See The Sedona Conference, Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, 15 SEDONA CONF. J. 217 (2014), available at https://thesedonaconference.org/publication/Commentary_on_Search_and_Retrieval_Methods.

Reinforcement Learning: an area of machine learning that involves teaching a system how to behave or take action using a reward and punishment structure. (from the forthcoming *Artificial Intelligence in the Practice of Law)*. See also Machine Learning.

Sentiment Analysis: Sometimes referred to as opinion mining or emotion AI, sentiment analysis uses Natural Language Processing to determine the emotional tenor of each component (phrase, sentences, segments). Basic examples would be positive or negative sentiment. See also Natural Language Processing.

Static Search: A search that is constructed to return the same records regardless of ongoing activity in the database, such as newly added documents or updated tagging. See Dynamic Search.

Stop Words: Common words (e.g., all, the, of, but, not) that are purposefully excluded from a search index when it is created in order to make the index more efficient. Also knows as Noise Words.

Stratified Sampling: A method of data sampling where data is initially divided into subgroups (e.g., by age range or a geo-graphic criteria) or strata, and then each group is sampled in order to ensure that each subgroup is properly represented. See also Sampling.

Supervised Learning: Use of machine learning to analyze data, using training examples that have been coded by humans, such as categorization. See also Machine Learning.

Support Vector Machine (SVM): an algorithm that creates a three-dimensional classification model to categorize data to a more finite certainty than two dimensional classifiers.

Technology-Assisted Review (TAR) : A process for prioritizing or coding a collection of electronically stored information using a computerized system that harnesses human judgments of subject-matter experts on a smaller set of documents and then extrapolates those judgments to the remaining documents in the collection. Some TAR methods use algorithms that determine how similar (or dissimilar) each of the remaining documents is to those coded as relevant (or nonrelevant) by the subject-matter experts, while other TAR methods derive systematic rules that emulate the experts' decision-making processes. TAR systems generally incorporate statistical models and/or sampling

techniques to guide the process and to measure overall system effectiveness. (from the Grossman & Cormack Glossary).

Underinclusive: When referring to data sets returned by some method of query, search, filter, or cull, results that are returned incomplete or too narrow. See False Negative.

Unintentional Bias: when an AI system produces biased or unfair results due to issues in the training data or errors in the design of the algorithm. (from the forthcoming *Artificial Intelligence in the Practice of Law)*. See also Bias.

Unsupervised Learning: Use of machine learning to analyze data without training examples, such as clustering. See Machine Learning.

Validation: an objective assessment of whether the AI is working as intended (i.e., valid) and produces accurate results under substantially similar circumstances (i.e., consistency and reliability). (from the forthcoming *Artificial Intelligence in the Practice of Law)*.

Wildcard Operator: A character used in text-based searching that assumes the value of any alphanumeric character, characters, or in some cases, words. Used to expand search terms and enable the retrieval of a wider range of hits.