



This chart covers the most common approaches in Automated Content Analysis. For specific use cases, you might also consider additional approaches: If you are interested in estimating how similar documents are, you can have a look at (a) cosine similarity, (b) Levenshtein distance, (c) Word Mover Distance. If you are interested in estimating how similar the meaning of words is (e.g., that the difference between queen and king is the same as between woman and man), you should look at word embeddings (e.g., word2vec). These approaches may be used stand-alone or in combination with the techniques in this chart.