



# Guide to Using Analytics Technology During Discovery

| Technology Type                 | Description  | Classification | Best Case Size for Implementation | Minimum # of Documents Required | Use with OCRd* | Use with EDA*/ ECA* Workflow |
|---------------------------------|--|----------------|-----------------------------------|---------------------------------|----------------|------------------------------|
| <b>Near Duplicate Detection</b> | Identifies documents that are near-duplicates of each other based on textual content and then groups those documents according to similarity.  | Structured     | All Cases                         | No                              | Yes            | Yes                          |
| <b>Email Threading</b>          | Determines the relationship between email messages by grouping related email items together and creating a visualization that helps users track the progression of an email chain.<br><br>*email threading will not be as accurate when used with OCR versus ESI however it can be done  | Structured     | All Cases                         | No                              | Yes            | Yes                          |
| <b>Keyword Expansion</b>        | Keyword expansion can be used to identify and validate words, terms, names, email domains, variations of words and/or conceptually related terms that a legal team wasn't aware of..   | Conceptual     | All Cases                         | No                              | Yes            | Yes                          |
| <b>Clustering</b>               | Identifies groups of conceptually similar documents, however, because clustering does not require user input, there is no way to identify which concepts are of particular interest to a legal team. Clustering is most useful when working with unfamiliar data sets..  | Conceptual     | All Cases                         | No                              | Yes            | Yes                          |
| <b>Categorization</b>           | Users create a set of example documents that will be used as the basis for identifying and grouping together other conceptually similar documents. As documents are reviewed, users can designate example documents and add them to various categories. These examples can then be used to apply categories to the remaining document population.. | Conceptual     | Medium – Large Cases              | 50,000+ docs                    | Yes            | Yes                          |

\*OCR - Optical Character Recognition, EDA = Early Data Assessment, ECA = Early Case Assessment



# Guide to Using Analytics Technology During Discovery

| Technology Type                   | Description  | Classification | Best Case Size for Implementation | Minimum # of Documents Required | Use with OCRd* Documents | Use with EDA*/ ECA* Workflow |
|-----------------------------------|--|----------------|-----------------------------------|---------------------------------|--------------------------|------------------------------|
| <b>Concept Searching</b>          | Identifies documents with similar conceptual content and is very different from keyword or metadata searching. A concept search reveals conceptual matches between a query and textual content versus matching a specific word or search term / set of search terms. This can help prioritize or find important content quickly.   | Conceptual     | Medium – Large Cases              | 50,000+ docs                    | Yes                      | Yes                          |
| <b>Relativity Assisted Review</b> | <p>Relativity Assisted Review uses categorization to train/teach the system on how to determine if a document is responsive or not responsive.</p> <p>Subject Matter Experts (SMEs) on the review team begin by coding a sample set of documents. Based on those coding decisions, Relativity determines how the remaining document population should be coded. Reviewers then validate Relativity's automated decisions by manually reviewing statistically-relevant subsets of documents in order to ensure coding accuracy.</p> | Predictive     | Medium – Large Cases              | 50,000+ docs                    | Yes                      | Yes                          |
| <b>Equivio Relevance</b>          | <p>Equivio Relevance is a Predictive Coding workflow that uses machine-learning technology that is trained, by way of examples, to imitate the decisions of Subject Matter Experts (SMEs).</p> <p>Relevance uses a scoring system to organize and rank documents according to relevance. This allows legal teams to quickly cull down raw data collections to more substantive datasets, focus on key documents and to prioritize data for review.</p>   | Predictive     | Medium – Large Cases              | 20,000 – 50,000+ docs           | Yes                      | Yes                          |

\*OCR - Optical Character Recognition, EDA = Early Data Assessment, ECA = Early Case Assessment

**Contact your local D4 office to schedule a technology needs assessment.**

e | [consulting@d4discovery.com](mailto:consulting@d4discovery.com)

w | [d4discovery.com/about-us/locations](http://d4discovery.com/about-us/locations)

D4 | 800.410.7066 | [www.d4discovery.com](http://www.d4discovery.com)



# Guide to Using Analytics Technology During Discovery

| Technology Type             | Description   | Classification            | Best Case Size for Implementation | Minimum # of Documents Required | Use with OCRd* Documents | Use with EDA*/ ECA* Workflow |
|-----------------------------|---|---------------------------|-----------------------------------|---------------------------------|--------------------------|------------------------------|
| <b>Brainspace Discovery</b> | Brainspace Discovery is a visual analytics solution that allows users to access clustering, concept searching and near-duplicate identification technology all within one seamless interface.<br><br>This collective and intuitive approach allows legal teams to quickly cull down raw data collections to more substantive datasets while at the same time connecting data, people and knowledge through the relationships the technology identifies. | Conceptual/<br>Predictive | Medium –<br>Large Cases           | 20,000 –<br>50,000+ docs        | Yes                      | Yes                          |

---

\*OCR - Optical Character Recognition, EDA = Early Data Assessment, ECA = Early Case Assessment

**Contact your local D4 office to schedule a technology needs assessment.**

**e** | [consulting@d4discovery.com](mailto:consulting@d4discovery.com)

**w** | [d4discovery.com/about-us/locations](http://d4discovery.com/about-us/locations)

**D4** | 800.410.7066 | [www.d4discovery.com](http://www.d4discovery.com)