

8.2.14 Equivio Relevance

Equivio Relevance is a lawyer guided software tool for scoring documents by relevance. From a sample of documents that have been classified by a lawyer as ‘relevant’ or ‘not relevant’, the software generates relevance scores for all documents in a collection. Equivio Relevance can be used at various stages in the legal document review process to save time and cost including:

- Initial case assessment – to rapidly assess key issues
- Culling of irrelevant documents that are responsive to typical keyword / phrase search techniques.
- Prioritisation of review – start the document review with the most relevant documents
- Post-review quality assurance – to assist in identifying review inconsistencies.

How does the Service Work?

1. The documents to be ranked for relevance are selected. For instance, they may be documents identified via the application of searches / filtering by say custodians, date range, file type etc. Alternatively, they may be the documents as selected by the client or documents received on disclosure.
2. The collection of documents to be reviewed are firstly de-duplicated and text extracted from electronic documents such as Word, email, PowerPoint etc. and for hard copy or scanned files text is created via optical character recognition ‘OCR’.
3. The text from the files to be reviewed is loaded into the Equivio Relevance software. There is the option to ‘seed’ the review with sample relevant documents and / or keywords. Alternatively, the software will select a representative batch of 40 documents for the lawyer to review.
4. The software is simple to use and does not require any training (see screenshot below). All that is required is for the lawyer to read the text from each document and decide whether the document is ‘relevant’ or ‘not relevant’. Note that it is also possible to review documents by ‘issue’.

The screenshot displays the Equivio Relevance software interface. At the top, there are tabs for 'Setup', 'Interactive Ranking', 'Sample Results', 'Batch Ranking', 'Final Results', and 'Utilities'. The main window shows a table of documents with columns for 'Row Id', 'Document Identifier', 'Topic 102', 'Date', and 'Size'. Row 10 is highlighted in blue. To the right of the table, a red arrow points to the list with the label 'List of Documents'. Below the table, a preview window shows the text of the selected document. A red arrow points to the text with the label 'Document Text for Review'. At the bottom right of the preview window, there is a 'Ranking Palette' with three buttons: 'Not Relevant', 'Relevant', and 'Relevant'. A red arrow points to the 'Not Relevant' button with the label 'Click 'Not Relevant' or 'Relevant''. At the bottom of the interface, there are buttons for 'Get More Documents' and 'Calculate Sample Results'.

Row Id	Document Identifier	Topic 102	Date	Size
6	2720319-mel2100.pdf	Not Relevant	3/30/2009	13
7	2056369-jhc39800.pdf	Not Relevant	3/30/2009	3
8	4823549-usc32a20.pdf	Not Relevant	3/30/2009	9
9	4516499-klt39c50.pdf	Not Relevant	3/30/2009	7
10	4407120-lav1000.pdf	None	3/30/2009	7
11	397408-buz24c00.pdf	None	3/30/2009	2
12	1075229-ewm53a00.txt	Relevant	3/30/2009	77
13	6039459-mt1300.pdf	Not Relevant	3/30/2009	2
14	806619-dip59e30.pdf	Not Relevant	3/30/2009	0
15	3836889-pcs62f00.txt	Relevant	3/30/2009	7

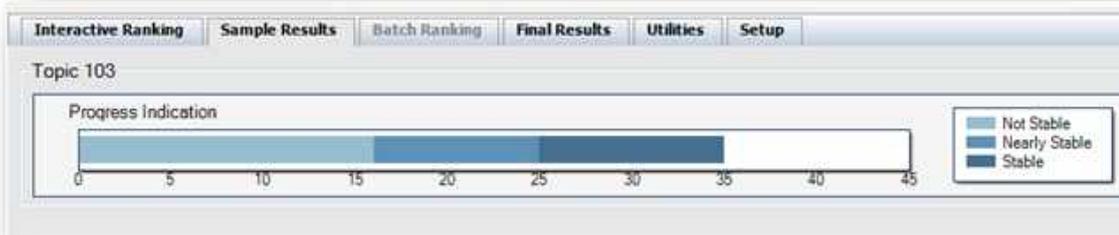
Preview Window Content:

PHENIX MIDSTREAM COMPANIES INC. I
 INTER-OFFICE CORRESPONDENCE
 120 PARK AVENUE, NEW YORK, N.Y. 10017
 TO: Distribution DATE: October 23, 1999
 FROM: Murray H. Beng
 SUBJECT: Product Liability Litigation
 As you may know, Victor Schwab, "is an attorney in Washington and a former professor of tort law who is widely regarded as one of the country's leading experts on product liability. Financial analysts and reporters frequently solicit his views on recent developments relating to product liability generally, and to our litigation in particular. On the whole, his comments have been quite supportive of our position regarding smoking and health litigation. Mr. Schwab recently gave an interview, to a Washington firm which publishes an investor information newsletter. I understand that the newsletter is distributed to most tobacco analysts. I am attaching hereto a copy of the newsletter containing a summary of that interview. As you will note, Mr. Schwab expresses a generally positive view, about our prospects for cases that may come to trial in 1999-1.

Distribution
 William I. Campbell
 Donald Reed
 Hannah Marquet
 Michael A. Niles
 John A. Murphy

Ranking Palette: [Not Relevant] [Relevant] [Relevant]

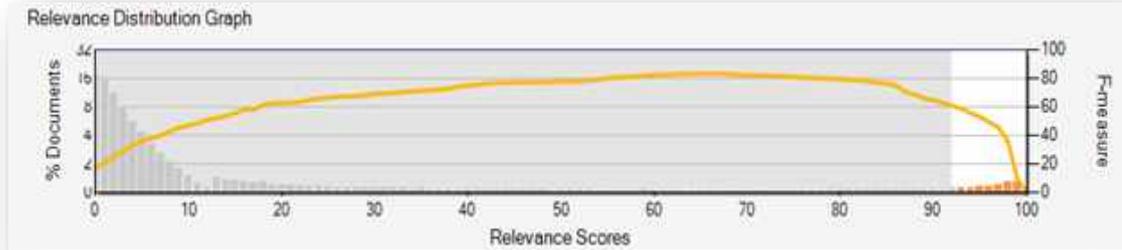
- The software presents to the lawyer batches of 40 documents at a time. Upon completion of each batch of 40 documents the software takes between 2 and 5 minutes to assess the results and create another batch of 40 documents for review. In the background, the Relevance software is effectively ‘second guessing’ the lawyer’s classification of documents. The process typically requires between 30 and 45 batches of 40 documents each (i.e. 1,200 – 1,800 documents) to become ‘stable’ (see the screenshot below showing the number of batches reviewed and progress towards reaching ‘stability’). ‘Stable’ means the software has reached a point where it is correctly assessing the relevant / not relevant classification in line with the lawyer review with a high degree of consistency.



- At completion of the review process, the software typically takes between 1 and 3 hours to calculate a ‘relevance score’ for all of the documents in the collection. The score is a number between 0 and 100.
- The Relevance software also provides useful statistics including a list of keywords / phrases used ranking in relation to the importance the software has placed upon such words / phrases in relation to determining the ‘relevant’ and ‘not relevant’ status of documents (see below for an example screenshot).

Include		Exclude	
Keyword	Weight	Keyword	Weight
brand	0.20	tobacco	-0.13
value	0.18	day	-0.09
display	0.17	received	-0.07
accounts	0.16	were	-0.07
field	0.14	ph	-0.07
stores	0.14	vantage ultra	-0.06
program	0.14	tie	-0.06
spanish speaking	0.14	protocol	-0.06

- Depending on the nature of the documents to be reviewed, it is normal to find that the relevance scores are grouped towards either end of the 0-100 scale. The final analysis includes various statistics including a graphical representation of the distribution of the relevance scores (see screenshot below) so as to assist with formulating an approach to reviewing documents thereafter.



- Finally, the documents along with their relevance score are loaded into an online review platform such as Relativity, CaseLogistix or Concordance. The legal review team can then use the relevance score to prioritise, quality check and potentially eliminate documents with very low relevance scores.

Owned/Supplied by :	Equivio (Acquired by Microsoft in January 2015)
Used by :	<p>AlixPartners, FRA, Legastat, Lighthouse.</p> <p>Equivio is often embedded into other litigation support products (particularly Relativity) to provide "Predictive Coding" functionality.</p> <p>Expect this to change as Microsoft terminate support for embedded options as contracts with suppliers expire.</p>