

Summary statistics

Run ID	FW14Docs50
Task	Resource selection
Resources used	Used result documents only
External data used in run	<none>
Description	Run LDA topic analysis over the whole collection of sampled documents. The topic distribution of each resource is calculated as the average of topic distributions of all the sampled documents in this resource. Each query is expanded by Google and then its topic distribution is inferred using the trained topic models. Finally the resources are ranked by according to the similarities of their topic distribution vectors to that of the query. In this run, the number of topics in training LDA topic models is set to 50.
Number of topics	50

Metric	Average	Median	Min	Max
nDCG@20	0.418	0.403	0.133	0.736
nDCG@10	0.292	0.277	0.026	0.806
nP@1	0.173	0.000	0.000	1.000
nP@5	0.203	0.113	0.000	0.983

