

Data set	Number of docs		K	\bar{D}	St. Dev.	Avg. Test perplexity per word (in nats)			
	Train	Test				LDA-50	LDA-200	R. Soft-50	Unigram
NIPS	1,690	50	13,649	98.0	245.3	3576	3391	3405	4385
20-news	11,314	7,531	2,000	51.8	70.8	1091	1058	953	1335
Reuters	794,414	10,000	10,000	94.6	69.3	1437	1142	988	2208