

dataset	#categories	#documents	#vocabulary	#avgLen	#cardinality
Ohsumed	23	12,929	12,711	96.67	1.66
Delicious	20	21,670	33,769	140.97	1.96