

Dataset	# of Strings	String Len (avg/max)	# of Synonym Pairs	# of Applicable Synonym Pairs Per String (avg/max)
USPS	1,000,000	6.75 / 15	300	2.19 / 5
CONF	10,000	5.84 / 14	1000	1.43 /4
SPROT	1,000,000	10.32 / 20	10,000	17.96 /68