

new sensible word count

20000

15000

10000

5000

0

0

50000

100000

150000

200000

number of tokens in the generated data

