

new sensible word count

20000

15000

10000

5000

0

0

25000

50000

75000

100000

125000

150000

175000

200000

number of tokens in the generated data

