

new sensible word count

8000

6000

4000

2000

0

0

50000

100000

150000

200000

number of tokens in the generated data

