

new sensible word count

8000

6000

4000

2000

0

0

1000000

2000000

3000000

4000000

5000000

6000000

number of types in the reference data

