

new sensible word count

20000

15000

10000

5000

0

0

10000

20000

30000

40000

50000

60000

number of types in the generated data

ar  
bg  
ca  
cs  
da  
de  
el  
en  
es  
et  
fa  
fi  
fr  
he  
hi  
hr  
hu  
id  
it  
ko  
lt  
lv  
ms  
nl  
no  
pl  
pt  
ro  
ru  
sk  
sl  
sr  
sv  
th  
tr  
uk  
vi  
zh

