

new sensible word count

20000

17500

15000

12500

10000

7500

5000

2500

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
sv
sr
th
tr
uk
vi
zh