

new sensible word count

8000

6000

4000

2000

0

0

100000

200000

300000

400000

500000

number of characters 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