After th se of ChatG 2022, t
nu f papers publish l every day ab
ge Lang odels (LLMs)
sec e tha 0-fold. umber o
of par er hese LLM np
340 millions mplementation
T to billions of eters models like
LL A lar of the pa
ara he word-em eddin
g lay s which are us repr
ite cabu y of characters, ts of cha
cte words. Ap on easing hoo
del complexit a fixed vocabu also re
espon for brittle mo which canno
with out-o ocab ry inp s and ca
ot gene ew languages