

Most underrepresented words in LLM texts

Rates relative to human chunk 2

GPT-4o		GPT-4o Mini		Llama 3 70B Instruct		Llama 3 8B Instruct		Llama 3 70B Chat	
Word	Rate	Word	Rate	Word	Rate	Word	Rate	Word	Rate
i.e.	0.0026	extremely	0.0034	yep	0.0075	anyways	0.0097	jeez	0
blah	0.0063	spokesman	0.0059	fucking	0.0075	analyse	0.011	donnie	0
fuck	0.0079	i.e.	0.0095	horrible	0.0077	somebody	0.012	alexithymia	0
fucking	0.0083	unhappy	0.010	ok	0.011	characterise	0.013	frampton	0
asshole	0.0093	bitch	0.012	fuckin	0.013	fuck	0.014	y'all	0
nasty	0.011	fucking	0.012	amid	0.015	obviously	0.014	analogous	0
chum	0.013	kg	0.013	fortunately	0.017	ok	0.014	carotenoid	0
shit	0.014	fuck	0.014	i.e.	0.018	fuckin	0.014	mussel	0
ok	0.014	visa	0.015	obviously	0.018	blah	0.015	que	0
and/or	0.016	ter	0.016	amongst	0.019	yep	0.016	whatsapp	0