

Appendix for: Large Language Models, Social Demography, and Hegemony: Comparing Authorship in Human and Synthetic Text

AJ Alvero^{1*}, Jinsook Lee², Alejandra Regla-Vargas³,
René Kizilcec⁴, Thorsten Joachims⁵, anthony lising antonio^{6*}

^{1*}Department of Sociology, University of Florida.

^{2,4-5}Department of Information Science, Cornell University.

³Department of Sociology, University of Pennsylvania.

⁵Department of Computer Science, Cornell University.

⁶Graduate School of Education, Stanford University.

*Corresponding author(s). E-mail(s): aalvero@ufl.edu;
aantonio@stanford.edu;

Contributing authors: jl3369@cornell.edu; aregla@sas.upenn.edu;
kizilcec@cornell.edu; thorsten.joachims@cornell.edu;

Appendix

Table A1 Essay Prompts

Public	Private
Describe an example of your leadership experience in which you have positively influenced others, helped resolve disputes, or contributed to group efforts over time.	Some students have a background, identity, interest, or talent that is so meaningful they believe their application would be incomplete without it. If this sounds like you, then please share your story.
Every person has a creative side, and it can be expressed in many ways: problem solving, original and innovative thinking, and artistically, to name a few. Describe how you express your creative side.	Reflect on a time when you questioned or challenged a belief or idea. What prompted your thinking? What was the outcome?
What would you say is your greatest talent or skill? How have you developed and demonstrated that talent over time?	Reflect on something that someone has done for you that has made you happy or thankful in a surprising way. How has this gratitude affected or motivated you?
Describe the most significant challenge you have faced and the steps you have taken to overcome this challenge. How has this challenge affected your academic achievement?	The lessons we take from obstacles we encounter can be fundamental to later success. Recount a time when you faced a challenge, setback, or failure. How did it affect you, and what did you learn from the experience?
Think about an academic subject that inspires you. Describe how you have furthered this interest inside and/or outside of the classroom.	Describe a topic, idea, or concept you find so engaging that it makes you lose all track of time. Why does it captivate you? What or who do you turn to when you want to learn more?
What have you done to make your school or your community a better place?	Discuss an accomplishment, event, or realization that sparked a period of personal growth and a new understanding of yourself or others.
Beyond what has already been shared in your application, what do you believe makes you stand out as a strong candidate for admissions to the University of California?	Share an essay on any topic of your choice. It can be one you've already written, one that responds to a different prompt, or one of your own design.
Describe how you have taken advantage of a significant educational opportunity or worked to overcome an educational barrier you have faced.	

LIWC Variable	Definition	Public Human	Public GPT 3.5	Public GPT 4	Private Human	Private GPT 3.5	Private GPT 4
verb	Total number verbs	16.27 (2.55)	8.20 (1.45)	6.93 (1.48)	12.94 (2.91)	9.02 (1.94)	7.30 (1.66)
affect	Affect words	5.44 (1.12)	7.73 (1.08)	6.54 (1.23)	5.10 (1.56)	6.78 (1.81)	5.66 (1.62)
posemo	Positive emotion	4.03 (1.01)	6.60 (0.88)	5.24 (1.01)	3.39 (1.22)	4.94 (1.17)	3.99 (1.00)
focuspresent	Present focus	8.54 (2.33)	4.48 (1.24)	2.99 (1.04)	5.38 (2.12)	3.88 (1.56)	2.76 (1.14)
function	Function words	55.84 (3.28)	44.46 (2.02)	42.23 (2.15)	50.88 (6.23)	47.96 (3.14)	47.36 (2.29)
pronoun	Total pronouns	17.37 (2.44)	12.91 (1.14)	11.51 (1.30)	15.31 (3.20)	13.40 (2.17)	11.20 (1.67)
ppron	Personal pronouns	12.05 (1.83)	10.06 (0.96)	8.24 (1.08)	11.07 (2.66)	9.33 (1.77)	7.06 (1.38)
you	Second person	0.25 (0.36)	0.01 (0.11)	0.01 (0.09)	0.26 (0.49)	0.04 (0.19)	0.02 (0.08)
social	Social words	7.59 (2.10)	6.92 (1.78)	6.28 (1.72)	6.62 (3.02)	5.39 (2.27)	4.17 (1.90)
prep	Prepositions	14.95 (1.39)	14.92 (1.12)	14.50 (1.20)	14.72 (2.18)	15.57 (1.27)	16.13 (1.23)
ipron	Impersonal pronouns	5.32 (1.40)	2.85 (0.64)	3.27 (0.76)	4.24 (1.48)	4.07 (1.03)	4.14 (0.90)
drives	Drives and needs	10.72 (1.99)	14.87 (2.30)	13.44 (2.41)	8.56 (2.61)	11.31 (2.99)	8.72 (2.73)
achieve	Achievement	3.76 (1.23)	6.96 (1.36)	6.00 (1.35)	2.92 (1.55)	5.22 (2.19)	4.12 (1.88)
reward	Reward focus	2.05 (0.72)	2.31 (0.66)	1.56 (0.58)	1.44 (0.75)	1.82 (0.79)	1.24 (0.64)
i	First person singular	9.71 (1.79)	9.21 (1.16)	7.19 (1.18)	8.86 (2.56)	7.91 (2.03)	6.01 (1.41)
auxverb	Auxiliary verbs	8.39 (1.69)	3.39 (1.06)	2.55 (0.89)	5.65 (1.83)	3.88 (1.14)	3.50 (1.00)
article	Articles	6.22 (1.18)	5.63 (0.97)	6.36 (1.05)	6.90 (1.72)	7.20 (1.23)	9.07 (1.38)
bio	Biological processes	1.54 (0.83)	1.19 (0.65)	1.25 (0.72)	1.99 (1.35)	1.50 (0.84)	1.16 (0.58)
sexual	Sexuality	0.08 (0.12)	0.25 (0.17)	0.25 (0.19)	0.09 (0.17)	0.33 (0.30)	0.23 (0.21)
adj	Adjectives	5.05 (0.93)	5.60 (0.88)	4.88 (0.97)	4.98 (1.24)	4.59 (0.97)	4.37 (0.94)
percept	Perceptual processes	1.88 (0.78)	1.56 (0.62)	1.56 (0.62)	2.60 (1.31)	1.87 (0.99)	2.07 (1.25)
feel	Feeling	0.55 (0.33)	0.17 (0.17)	0.23 (0.22)	0.70 (0.53)	0.40 (0.34)	0.32 (0.27)
work	Work	6.75 (2.10)	12.16 (2.08)	10.96 (2.31)	4.20 (2.41)	6.73 (2.48)	5.76 (2.13)
affiliation	Affiliation	2.75 (1.17)	3.66 (1.39)	3.75 (1.38)	2.37 (1.54)	2.80 (1.76)	2.21 (1.40)
cogproc	Cognitive processes	12.82 (2.24)	9.94 (1.70)	10.48 (1.67)	11.25 (2.77)	11.72 (1.91)	11.13 (1.85)
cause	Cause	2.55 (0.79)	2.83 (0.83)	2.78 (0.86)	1.95 (0.91)	2.15 (0.92)	1.74 (0.77)
power	Power	3.63 (0.97)	4.13 (1.02)	3.71 (1.12)	2.73 (1.14)	3.24 (1.30)	2.39 (1.11)
insight	Insight	3.26 (0.88)	4.08 (0.94)	4.07 (0.96)	3.23 (1.22)	4.84 (1.20)	4.45 (1.14)
relativ	Relativity	13.67 (1.87)	11.39 (1.48)	12.75 (1.71)	14.79 (2.95)	13.07 (1.92)	12.62 (1.78)
space	Space	6.46 (1.09)	7.04 (1.26)	7.48 (1.30)	6.91 (1.67)	7.19 (1.51)	7.18 (1.35)
negate	Negations	1.27 (0.55)	0.31 (0.22)	0.80 (0.32)	0.80 (0.51)	0.66 (0.37)	1.14 (0.43)
differ	Differentiation	2.89 (0.88)	1.63 (0.61)	2.14 (0.70)	2.36 (0.95)	2.29 (0.96)	2.76 (0.88)
interrog	Interrogatives	1.63 (0.57)	0.81 (0.35)	0.79 (0.38)	1.40 (0.66)	0.80 (0.43)	0.87 (0.45)
see	Seeing	0.65 (0.43)	0.41 (0.33)	0.47 (0.33)	1.02 (0.77)	0.52 (0.41)	0.67 (0.46)
conj	Conjunctions	6.90 (1.10)	7.04 (0.83)	6.12 (0.92)	6.12 (1.34)	6.75 (1.11)	5.84 (0.99)
focuspast	Past focus	5.70 (1.69)	2.87 (0.93)	2.83 (0.94)	5.67 (2.12)	3.99 (1.72)	3.49 (1.29)
focusfuture	Future focus	0.87 (0.41)	0.27 (0.23)	0.31 (0.26)	0.75 (0.47)	0.62 (0.40)	0.67 (0.38)
time	Time	5.40 (1.25)	2.81 (0.88)	3.25 (1.03)	5.87 (1.71)	3.54 (1.17)	3.42 (1.04)
adverb	Common adverbs	4.16 (1.05)	2.20 (0.63)	2.64 (0.74)	3.98 (1.27)	2.86 (0.82)	2.92 (0.78)
motion	Motion	1.99 (0.61)	1.60 (0.52)	2.07 (0.62)	2.29 (0.96)	2.52 (0.83)	2.20 (0.72)
relig	Religion	0.13 (0.24)	0.04 (0.09)	0.13 (0.16)	0.15 (0.41)	0.15 (0.27)	0.23 (0.23)
negemo	Negative emotion	1.31 (0.64)	0.89 (0.52)	1.04 (0.55)	1.62 (0.99)	1.51 (1.24)	1.41 (1.11)
tentat	Tentativeness	1.89 (0.73)	0.67 (0.35)	0.08 (0.40)	1.76 (0.86)	1.17 (0.64)	1.33 (0.63)
compare	Comparatives	2.82 (0.71)	2.28 (0.56)	1.85 (0.61)	2.85 (0.93)	2.32 (0.67)	2.25 (0.69)
quant	Quantifiers	2.20 (0.63)	1.26 (0.43)	1.15 (0.50)	1.98 (0.81)	1.20 (0.53)	1.39 (0.58)
number	Numbers	1.11 (0.50)	0.31 (0.25)	0.28 (0.28)	0.99 (0.58)	0.38 (0.28)	0.41 (0.29)
certain	Certainty	1.72 (0.62)	1.05 (0.42)	0.92 (0.42)	1.56 (0.72)	1.25 (0.52)	0.99 (0.45)
leisure	Leisure	1.47 (1.01)	1.28 (0.77)	1.31 (0.73)	1.58 (1.30)	1.10 (1.09)	1.27 (1.11)
health	Health/illness	0.72 (0.54)	0.37 (0.37)	0.44 (0.41)	0.74 (0.74)	0.56 (0.55)	0.42 (0.36)
risk	Risk focus	0.51 (0.32)	0.44 (0.31)	0.47 (0.33)	0.57 (0.45)	0.78 (0.71)	0.70 (0.63)
discrep	Discrepancies	1.56 (0.62)	0.49 (0.31)	0.55 (0.35)	1.38 (0.72)	1.03 (0.49)	0.82 (0.45)
anx	Anxiety	0.37 (0.28)	0.22 (0.25)	0.35 (0.32)	0.47 (0.47)	0.40 (0.52)	0.34 (0.44)
anger	Anger	0.19 (0.20)	0.12 (0.16)	0.18 (0.20)	0.27 (0.32)	0.18 (0.23)	0.23 (0.25)
body	Body	0.35 (0.32)	0.15 (0.17)	0.18 (0.19)	0.67 (0.67)	0.41 (0.39)	0.31 (0.27)
hear	Hearing	0.47 (0.48)	0.42 (0.44)	0.37 (0.42)	0.58 (0.75)	0.30 (0.72)	0.69 (1.10)
money	Money	0.36 (0.36)	0.32 (0.34)	0.30 (0.34)	0.38 (0.52)	0.26 (0.34)	0.18 (0.25)
informal	Informal speech	0.18 (0.17)	0.09 (0.12)	0.09 (0.16)	0.18 (1.00)	0.06 (0.12)	0.07 (0.14)
nonflu	Nonfluencies	0.12 (0.13)	0.08 (0.12)	0.03 (0.08)	0.07 (0.12)	0.04 (0.09)	0.02 (0.07)
we	First person plural	0.68 (0.60)	0.48 (0.48)	0.76 (0.62)	0.65 (0.84)	0.80 (0.99)	0.64 (0.81)
friend	Friends	0.21 (0.20)	0.18 (0.18)	0.15 (0.19)	0.24 (0.29)	0.11 (0.16)	0.07 (0.13)
shehe	Third person singular	0.52 (0.62)	0.01 (0.06)	0.01 (0.07)	0.70 (1.09)	0.15 (0.41)	0.11 (0.31)
male	Male referents	0.48 (0.60)	0.07 (0.12)	0.04 (0.11)	0.70 (1.12)	0.15 (0.44)	0.12 (0.33)
they	Third person plural	0.89 (0.58)	0.36 (0.31)	0.26 (0.26)	0.60 (0.58)	0.44 (0.45)	0.28 (0.28)
home	Home	0.44 (0.36)	0.19 (0.21)	0.25 (0.29)	0.52 (0.56)	0.20 (0.29)	0.22 (0.32)
family	Family	0.68 (0.61)	0.12 (0.18)	0.13 (0.21)	0.71 (0.88)	0.12 (0.27)	0.11 (0.23)
female	Female referents	0.51 (0.64)	0.01 (0.08)	0.03 (0.11)	0.58 (1.07)	0.11 (0.42)	0.08 (0.27)
sad	Sadness	0.29 (0.25)	0.17 (0.24)	0.16 (0.21)	0.40 (0.39)	0.51 (0.73)	0.44 (0.56)
netpeak	Netpeak	0.02 (0.06)	0.01 (0.04)	0.05 (0.13)	0.04 (0.14)	0.01 (0.06)	0.03 (0.11)
assent	Assent	0.03 (0.06)	0.00 (0.02)	0.00 (0.03)	0.05 (0.98)	0.01 (0.04)	0.02 (0.06)
ingest	Ingesting	0.25 (0.31)	0.29 (0.29)	0.23 (0.36)	0.40 (0.77)	0.15 (0.31)	0.17 (0.28)
death	Death	0.05 (0.10)	0.04 (0.10)	0.02 (0.07)	0.07 (0.18)	0.02 (0.06)	0.03 (0.09)
swear	Swear words	0.01 (0.03)	0.00 (0.00)	0.00 (0.01)	0.01 (0.07)	0.00 (0.02)	0.00 (0.03)
Analytic	Analytical thinking	13.08 (6.55)	24.71 (3.23)	27.24 (3.33)	19.01 (7.57)	24.12 (6.52)	30.61 (4.48)
SemiC	Semicolons	0.10 (0.16)	0.01 (0.04)	0.18 (0.20)	0.15 (0.22)	0.08 (0.13)	0.29 (0.23)
Period	Periods	4.94 (0.94)	5.28 (0.42)	4.88 (0.46)	5.22 (1.18)	4.91 (0.58)	4.49 (0.58)
Sixltr	Words longer than 6 letters	19.78 (3.48)	39.31 (2.57)	40.06 (2.81)	20.86 (4.75)	33.18 (4.12)	34.69 (3.26)

Table A2 LIWC Features and definition mean and SD

Groups	LIWC Features	GPT4 (Public)	GPT4 (Private)	GPT3.5 (Public)	GPT3.5 (Private)
Gender	negemo	male	male	male	male
	percept	male	male	male	male
	family	male	male	male	male
	female	male	male	male	male
	focuspast	male	female	male	female
	pronoun	male	male	male	male
	quant	female	female	female	female
	male	female	female	female	female
	time	female	female	female	female
	social	male	male	male	male
	feel	male	male	male	male
	body	male	male	male	male
	i	male	male	male	male
	verb	male	female	male	female
	achieve	male	male	male	male
	ingest	male	male	female	male
	negate	male	female	male	male
	ppron	male	male	male	male
	bio	male	male	male	male
	hear	male	female	male	male
	leisure	female	female	female	female
	home	male	male	male	male
	anx	male	male	male	male
	shehe	male	male	male	male
	Period	male	male	female	male
	article	male	male	female	male
	compare	female	female	female	female
	reward	female	female	male	male
	work	male	male	male	male
First Gen	number	First Gen	First Gen	First Gen	First Gen
	discrep	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	article	First Gen	Cont. Gen	First Gen	Cont. Gen
	i	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	Analytic	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	prep	First Gen	Cont. Gen	First Gen	Cont. Gen
	social	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	sexual	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	sad	Cont. Gen	First Gen	Cont. Gen	First Gen
	pronoun	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	interrog	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	ipron	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	focusfuture	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	affect	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	you	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	ppron	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	work	First Gen	Cont. Gen	First Gen	Cont. Gen
	auxverb	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	relative	First Gen	First Gen	First Gen	First Gen
	health	First Gen	Cont. Gen	First Gen	Cont. Gen
	adj	First Gen	First Gen	Cont. Gen	First Gen
	female	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	posemo	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	verb	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	focuspresent	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	quant	First Gen	First Gen	First Gen	First Gen
	negate	Cont. Gen	First Gen	Cont. Gen	Cont. Gen
	space	First Gen	Cont. Gen	Cont. Gen	Cont. Gen
	friend	First Gen	First Gen	First Gen	First Gen
	function	First Gen	Cont. Gen	Cont. Gen	Cont. Gen
EC	leisure	First Gen	First Gen	First Gen	First Gen
	they	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	family	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	adverb	First Gen	Cont. Gen	Cont. Gen	Cont. Gen
	home	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	Sixltr	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	focuspast	Cont. Gen	Cont. Gen	Cont. Gen	Cont. Gen
	article	High EC	High EC	Low EC	High EC
	Analytic	High EC	High EC	High EC	High EC
	i	High EC	High EC	High EC	High EC
	prep	Low EC	High EC	Low EC	High EC
	pronoun	High EC	High EC	High EC	High EC
	interrog	High EC	High EC	High EC	High EC
	ipron	High EC	High EC	High EC	High EC
	focusfuture	High EC	High EC	High EC	High EC
	you	High EC	High EC	High EC	High EC
	ppron	High EC	High EC	High EC	High EC
	auxverb	High EC	High EC	High EC	High EC
	health	Low EC	High EC	Low EC	High EC
	adj	Low EC	Low EC	High EC	Low EC
	posemo	High EC	High EC	High EC	High EC
	verb	High EC	High EC	High EC	High EC
	focuspresent	High EC	High EC	High EC	High EC
	negate	High EC	Low EC	High EC	High EC
	space	High EC	High EC	High EC	High EC
	function	High EC	High EC	High EC	High EC
	leisure	Low EC	Low EC	Low EC	Low EC
	family	High EC	High EC	High EC	High EC
	achieve	Low EC	High EC	Low EC	High EC
	adverb	High EC	High EC	High EC	High EC
	Sixltr	High EC	High EC	High EC	High EC
	focuspast	High EC	High EC	High EC	High EC

Table A3 LIWC features and demographic subgroups are closer to each feature across GPT written essays