

Kaj družabna omrežja res vedo o tebi?

Marko Tkalčič

Univerza na Primorskem

Fakulteta za matematiko, naravoslovje in informacijske tehnologije

Kazalo

- Uvod, digitalne sledi
- Napovedovanje
 - Preference
 - Osebnost
 - Spolna usmerjenost
- Uporaba
 - Usmerjeno oglaševanje
 - Politično oglaševanje
 - Manipuliranje čustev
- Zaključek

Kazalo

- Uvod, digitalne sledi
- Napovedovanje
 - Preference
 - Osebnost
 - Spolna usmerjenost
- Uporaba
 - Usmerjeno oglaševanje
 - Politično oglaševanje
 - Manipuliranje čustev
- Zaključek

Uvod

- Družabna omrežja (Facebook, Instagram, TikTok)
 - Spletne trgovine (Amazon, mimovrste ...)
 - Pretočne storitve (Netflix, Spotify ...)
 - Druge digitalne platforme (zemljevidi ...)
- Kaj družabna omrežja res vedo o tebi?
 - Ne vem
 - Danes: javno dostopne raziskave

Digitalne sledi

- **Digitalne sledi** (digitalni odtisi)
 - Zapisi našega obnašanja na spletu
 - Nakupi
 - Ogledi
 - Všečki
 - Objave
 - Prijatelji
 - Elektronska pošta
 - Sporočila
 - Iskanja
 - Komentarji
- Digitalne sledi podjetja združijo v **profil uporabnika**
 - Odvisen od namena (Netflix vs TikTok)

Zakaj podjetja rabijo profile

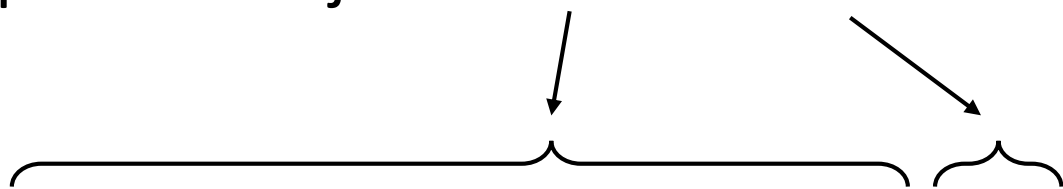
- Najpogostejši razlog:
 - Povečevanje dobičkov
 - Več prodanih izdelkov
 - Več prikazanih oglasov
 - Več naročnikov
- Stranski učinki
 - Napoved neželjenih/nepotrebnih lastnosti
 - Napačne napovedi

Kazalo

- Uvod, digitalne sledi
- **Napovedovanje**
 - Preference
 - Osebnost
 - Spolna usmerjenost
- Uporaba
 - Usmerjeno oglaševanje
 - Politično oglaševanje
 - Manipuliranje čustev
- Zaključek

Napovedovanje (Strojno učenje)

- Napovedovanje iz **značilk v razrede**



Uporabnik	Dolžina besedila	Beseda "petek"	Beseda "naloga"	Beseda "bedno"	...	Všečkal
1	125	1	0	0		1
1	45	1	0	1		0
1	230	0	1	0		1
2	16	0	0	1		0
2	35	1	0	1		1
3	78	0	1	0		1
3	64	0	0	0		0
4	55	0	0	0		0
5	112	0	0	1		0

Napovedovanje (Strojno učenje)

- Napovedovanje iz **značilk v razrede**

Uporabnik	Dolžina besedila	Beseda "petek"	Beseda "naloga"	Beseda "bedno"	...	Všečkal
1	125	1	0	0		1
1	45	1	0	1		0
1	230	0	1	0		1
2	16	0	0	1		0
2	35	1	0	1		1
3	78	0	1	0		1
3	64	0	0	0		0
4	55	0	0	0		0
5	112	0	0	1		0

	Napovedano všečkanje	Napovedano nevšečkanje
Uporabnik JE všečkal	112	44
Uporabnik NI všečkal	78	388

Točnost (Accuracy) = $(112+388)/(112+44+78+388) = 0,80$
Natančnost (Precision) = $112/(112+78) = 0,59$
Priklic (Recall) = $112/(112+44) = 0,72$

Napovedovanje preferenc

- Klasični priporočilni sistemi, npr. Netflix
 - Kateri film bi uporabniku bil všeč glede na preteklo obnašanje?

	Godfather	Inception	Hangover	Sophie's Choice
Peter	4	5	5	?
Paul	?	3	1	4
Mary	2	?	?	3
Joan	?	2	?	4



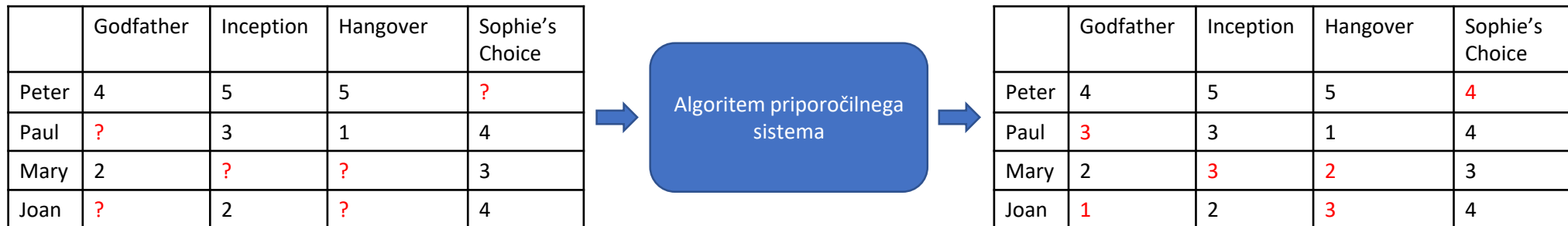
Algoritem priporočilnega sistema



	Godfather	Inception	Hangover	Sophie's Choice
Peter	4	5	5	4
Paul	3	3	1	4
Mary	2	3	2	3
Joan	1	2	3	4

Napovedovanje preferenc

- Klasični priporočilni sistemi, npr. Netflix
 - Kateri film bi uporabniku bil všeč glede na preteklo obnašanje?



- Glasbena priporočila (Spotify...)
- Priporočeni artikli (spletne trgovine)
- Vzajemna priporočila (LinkedIn, Tinder)

Napovedovanje osebnosti

- Ljudje smo različni

Napovedovanje osebnosti

- Ljudje smo različni
- Osebnostni model velikih pet
 - Odprtost
 - Vestnost
 - Ekstravertnost
 - Sprejemljivost
 - Nevroticizem

Napovedovanje osebnosti

- Ljudje smo različni
- Osebnostni model velikih pet
 - Odprtost
 - Vestnost
 - Ekstravertnost
 - Sprejemljivost
 - Nevroticizem
- Dolgi vprašalniki
 - Več deset vprašanj
 - Zamudni

Zakaj so osebnostne lastnosti pomembne?

- Osebnostne lastnosti napovedujejo (so korelirane z):
 - Preference za glasbo

Zakaj so osebnostne lastnosti pomembne?

- Osebnostne lastnosti
- Preference za glasbo

Blues
Jazz
Classical
Folk

Rock
Alternative
Heavy metal

Country
Religious
Pop

Rap/hip hop
Soul/funk
electronic

ujeje...ane

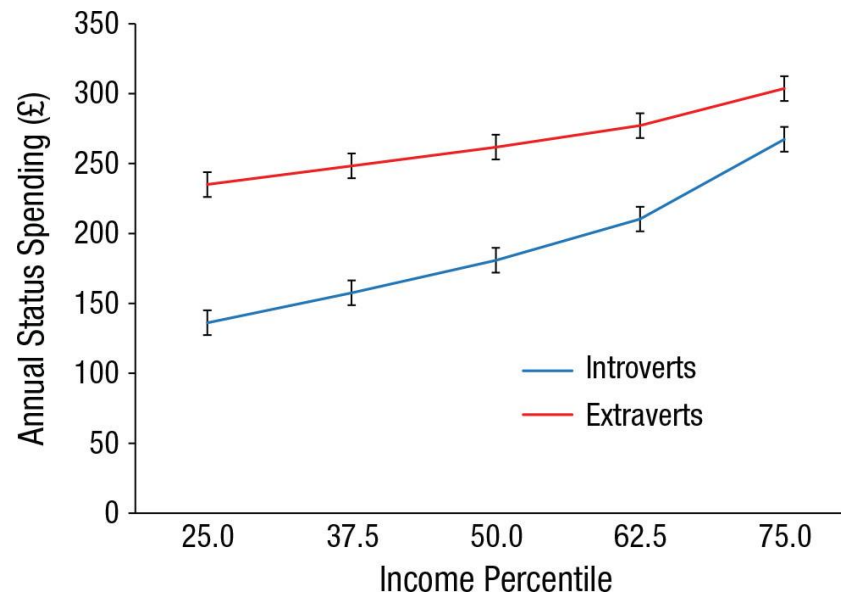
External Correlates of the Music-Preference Dimension

Criterion measure	M (SD)	Reflective and Complex		Intense and Rebellious		Upbeat and Conventional		Energetic and Rhythmic	
		S2	S3	S2	S3	S2	S3	S2	S3
Personality									
Big Five									
Extraversion	3.42 (0.85)	.01	-.02	.00	.08*	.24*	.15*	.22*	.19*
Agreeableness	3.80 (0.62)	.01	.03	-.04	.01	.23*	.24*	.08*	.09*
Conscientiousness	3.57 (0.64)	-.02	-.06	-.04	-.03	.15*	.18*	.00	-.03
Emotional Stability	3.11 (0.81)	.00*	.04	-.01	-.01	-.07	-.04	.01	-.01
Openness	3.75 (0.61)	.44*	.41*	.18*	.15*	-.14*	-.08*	.03	.04

Vir: Rentfrow, P. J., and Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6), 1236–1256.

Zakaj so osebnostne lastnosti pomembne?

- Osebnostne lastnosti napovedujejo (so korelirane z):
 - Preference za glasbo
 - Nakupovalne navade (statusni simboli)



Vir: Landis, Blaine, and Joe J. Gladstone. "Personality, Income, and Compensatory Consumption: Low-Income Extraverts Spend More on Status." *Psychological Science* 28, no. 10 (October 1, 2017): 1518–20. <https://doi.org/10.1177/0956797617714811>.

Napovedovanje osebnosti

- Digitalne sledi
 - Družabnih omrežjih
 - Nakupovalne navade
 - Nastavitve zasebnosti

Napovedovanje osebnosti – družabna omrežja

- Facebook všečki
- SVD algoritem za ustvarjanje značilk
- Logistična regresija za napovedovanje osebnosti

Napovedovanje osebnosti – družabna omrežja

- Twitter/Instagram
 - Jezikovne značilke
 - Linguistic Inquiry and Word Count (LIWC)
 - Affective Norms for English Words (ANEW)
 - Slikovne značilke
 - Svetlost
 - Nasičenost
 - Odtenki
 - Uporabniške značilke
 - Število sledilcev

Napovedovanje osebnosti – družabna omrežja

- Twitter/Instagram
 - Jezikovne značilke
 - Linguistic Inquiry and Word Count (LIWC)
 - Affective Norms for English Words (ANEW)
 - Slikovne značilke
 - Svetlost
 - Nasičenost
 - Odtenki
 - Uporabniške značilke
 - Število sledilcev

	O	C	E	A	N	AVG
B	0.75	0.75	1.04	0.67	0.98	1.14
T _m	0.74	0.78	1.16	0.71	1.07	0.89
I _i	0.80	0.70	0.98	0.74	0.90	0.83
T _{lm}	0.75	0.73	0.92	0.71	0.80	0.78
T _m I _i	0.62	0.66	0.92	0.69	0.92	0.77
I _l	0.62	0.66	0.92	0.69	0.92	0.76
T _l	0.73	0.66	0.96	0.63	0.75	0.75
T _m I _l	0.65	0.68	0.86	0.60	0.79	0.72
T _l I _l	0.61	0.68	0.86	0.63	0.76	0.71
T _m I _{li}	0.51	0.68	0.86	0.55	0.88	0.70
T _l I _{li}	0.64	0.65	0.87	0.55	0.73	0.69
T _{lm} I _{li}	0.53	0.67	0.71	0.56	0.83	0.66

Table 1: RMSE with features from (T)witter and (I)nstagram; and (B)aseline score: average value for a dimension. Subscripts indicate sets of features used: (l)inguistic, i(m)age), (m)eta.

Napovedovanje osebnosti – nakupovalne navade

ID	SKUPINA	PODSKUPINA	PODPODSKUPINA	TRGOVEC	DATUM	ZNESEK	ODPRTOST	VESTNOST	EXTRAVERTNOST	SPREJEMLJIVOST	NEVROTICIZEM
132	Blago in storitve	Olačila	Olačila drugo	H&M	01/10/2013	115,33	3,8	5,3	8,7	5,5	4,5
544	Blago in storitve	Hrana in pijača	Supermarket	Mercator	01/10/2013	36,00	4,2	4,2	6,1	4,2	7,3
543	Zdravje	Zdravje	Lekarna	Ljubljanske lekarnе	01/10/2013	126,02	5,6	6,1	5,5	3,6	2,5
398	Blago in storitve	Finančne storitve	Zavarovanje	Adriatic	01/10/2013	35,32	1,2	3,7	3,3	1,9	1,9
...

Vir: Gladstone, Joe J., Sandra C. Matz, and Alain Lemaire. "Can Psychological Traits Be Inferred From Spending? Evidence From Transaction Data." *Psychological Science* 30, no. 7 (July 2019): 1087–96. <https://doi.org/10.1177/0956797619849435>.

Napovedovanje osebnosti – nastavitve zasebnosti

		O	C	E	A	N
4	Current city	-.24*	-.20[^]	-.08	-.08	.01
5	Hometown	-.25*	-.18[^]	-.08	-.13	-.05
6	Places lived	-.12	-.12	-.08	-.20*	-.01
7	Mobile phone	-.22*	-.12	-.01	-.05	.10
8	Website	-.22*	.01	.16	.02	-.16
9	Email	-.16	.09	-.23*	.13	-.13
10	Address	-.24*	-.02	.14	-.04	-.15
11	Birth date	-.18[^]	-.18[^]	-.22*	-.12	.17[^]
32	Restaurant	.03	-.06	.22*	-.06	.09
33	Games	.10	.01	.18[^]	.02	-.13
34	Activities	.05	.03	.21*	.06	-.08
35	Interests	.09	-.04	.17[^]	-.06	-.05
37	Foods	.01	-.18	.24*	.01	-.11
38	Clothing	-.05	-.06	.19[^]	.01	-.09
40	Other	-.05	-.19[^]	.08	-.09	.02

Note. [^] $p < 0.1$, * $p < 0.05$

Table 2: Correlation Matrix of the profile items disclosure against the personality traits: (O)penness, (C)onscientiousness, (E)xtroversion, (A)greeableness, (N)euroticism. Only items that show significant levels of $p < 0.1$ are reported.

Napovedovanje osebnosti – nastavitve zasebnosti

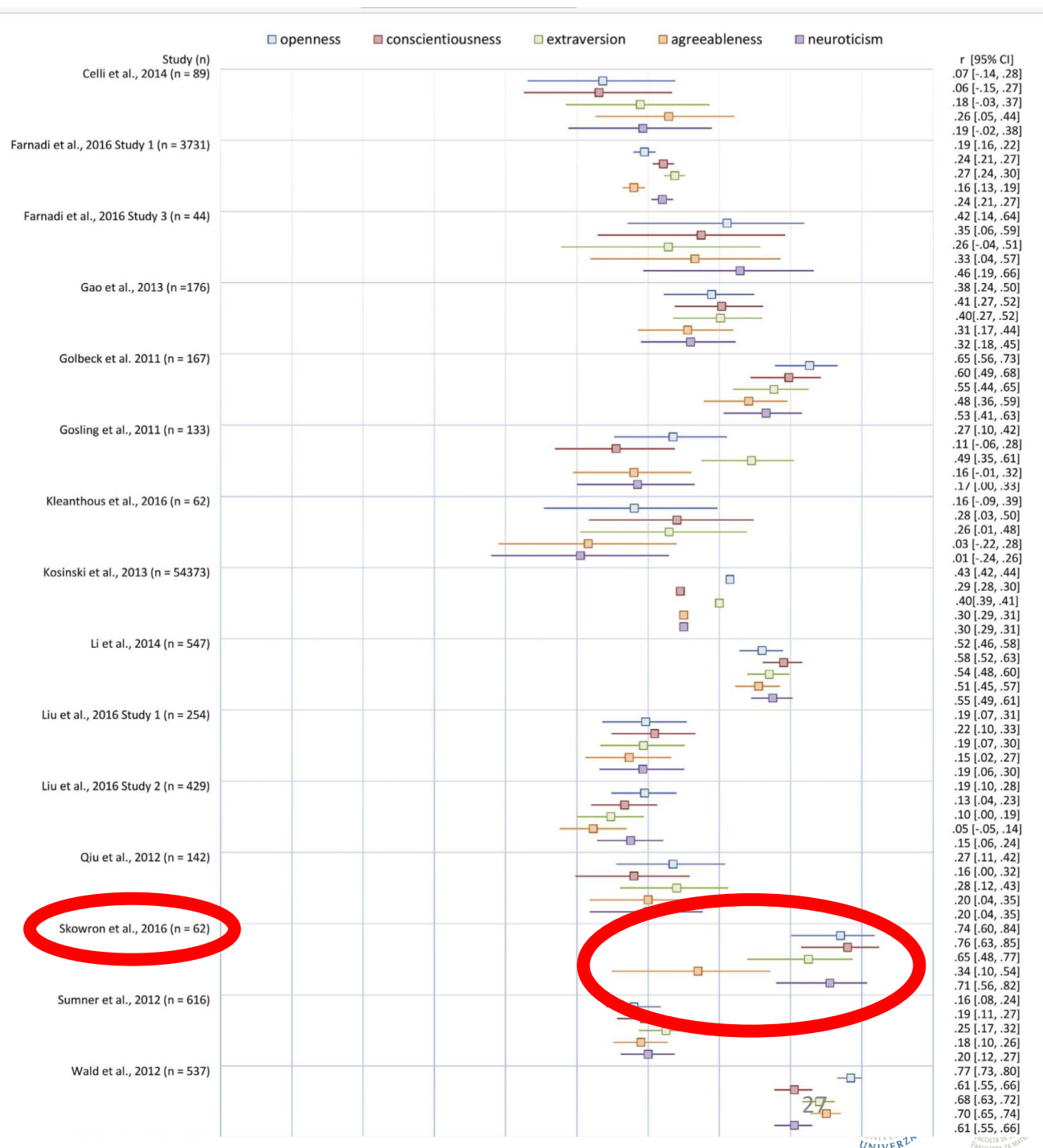
		O	C	E	A	N
4	Current city	-.24*	-.20^	-.08	-.08	.01
5	Hometown	-.25*	.18^	-.08	-.13	-.05
6	Places lived	-.12	-.12	-.08	-.20*	-.01
7	Mobile phone	-.22*	-.12	-.01	-.05	.10
8	Website	-.22*	.01	.16	.02	-.16
9	Email	-.16	.09	-.23*	.13	-.13
10	Address	-.24*	-.02	.14	-.04	-.15
11	Birth date	-.18^	-.18^	-.22*	-.12	.17^
32	Restaurant	.03	-.06	.22*	-.06	.09
33	Games	.10	.01	.18^	.02	-.13
34	Activities	.05	.03	.21*	.06	-.08
35	Interests	.09	-.04	.17^	-.06	-.05
37	Foods	.01	-.18	.24*	.01	-.11
38	Clothing	-.05	-.06	.19^	.01	-.09
40	Other	-.05	-.19^	.08	-.09	.02

Note. ^ $p < 0.1$, * $p < 0.05$

Table 2: Correlation Matrix of the profile items disclosure against the personality traits: (O)penness, (C)onscientiousness, (E)xtraversion, (A)greeableness, (N)euroticism. Only items that show significant levels of $p < 0.1$ are reported.

Napovedovanje osebnosti - primerjava

Vir: Azucar, Danny, Davide Marengo, and Michele Settanni. 'Predicting the Big 5 Personality Traits from Digital Footprints on Social Media: A Meta-Analysis'. *Personality and Individual Differences* 124 (April 2018): 150–59. <https://doi.org/10.1016/j.paid.2017.12.018>.



Napovedovanje osebnosti

- Primer: <https://applymagicsauce.com/demo> (Cambridge University)

Napovedovanje spolne usmerjenosti

- 35.000 slik
- Natančnost:
 - 0,81 moški (0,61)
 - 0,71 ženske (0,51)

Vir: Wang, Yilun, and Michal Kosinski. "Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation From Facial Images," n.d., 12.

Napovedovanje depresije

- Zapisi na Facebooku
- 683 pacientov
 - 114 z depresijo

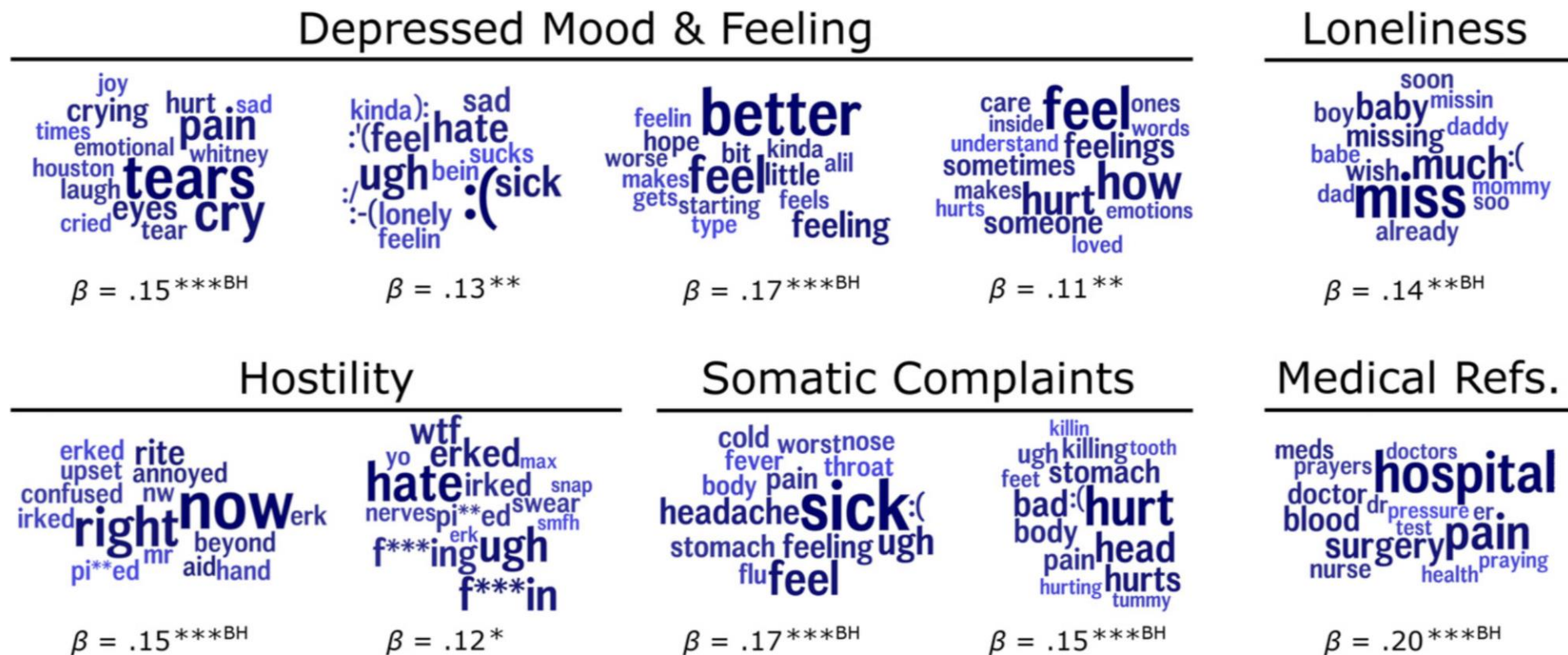


Fig. 4. Ten language topics most positively associated with a future depression diagnosis controlling for demographics (* $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$; ^{BH} $P < 0.05$ after Benjamini–Hochberg correction for multiple comparisons). Font size reflects relative prevalence of words within topics. Color shading is to aid readability and carries no meaning.

Kazalo

- Uvod, digitalne sledi
- Napovedovanje
 - Preference
 - Osebnost
 - Spolna usmerjenost
- **Uporaba**
 - Usmerjeno oglaševanje
 - Politično oglaševanje
 - Manipuliranje čustev
- Zaključek

Uporaba osebnosti

- Napovedovanje preferenc (spomni se nekaj prosojnic nazaj)
- Napovedovanje potrošnje (spomni se nekaj prosojnic nazaj)
- Usmerjeno oglaševanje (targeted advertising)
- Politično oglaševanje (Cambridge Analitica)

Usmerjeno oglaševanje

Vir: Matz, S. C., M. Kosinski, G. Nave, and D. J. Stillwell. "Psychological Targeting as an Effective Approach to Digital Mass Persuasion." *Proceedings of the National Academy of Sciences* 114, no. 48 (November 28, 2017): 12714–19. <https://doi.org/10.1073/pnas.1710966114>.

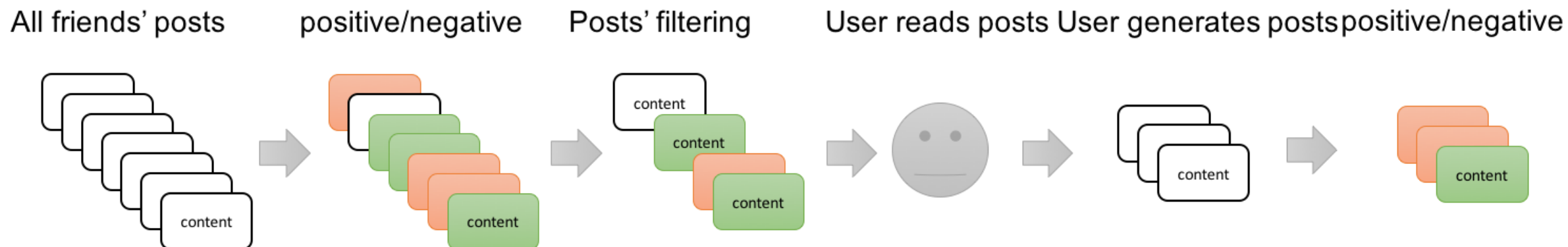
Politično oglaševanje (primer Cambridge Analytica)

- 2010-2016: osebni podatki desetine milijonov Facebook uporabnikov
 - Brez dovoljenj uporabnikov
- Napoved modela velikih pet
- Namen: politično oglaševanje
 - 2016: Donald Trump
 - 2016: Ted Cruz
 - 2016: Brexit referendum
- Npr., prepričevanje volilca, da se kandidat zavzema za službe:
 - Vestna oseba: možnost napredovanja in odgovornost, ki jo daje služba
 - Odprta oseba: možnost osebne rasti
 - Nevrotična oseba: socialna varnost, ki jo služba prinaša

Vir: <https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie>

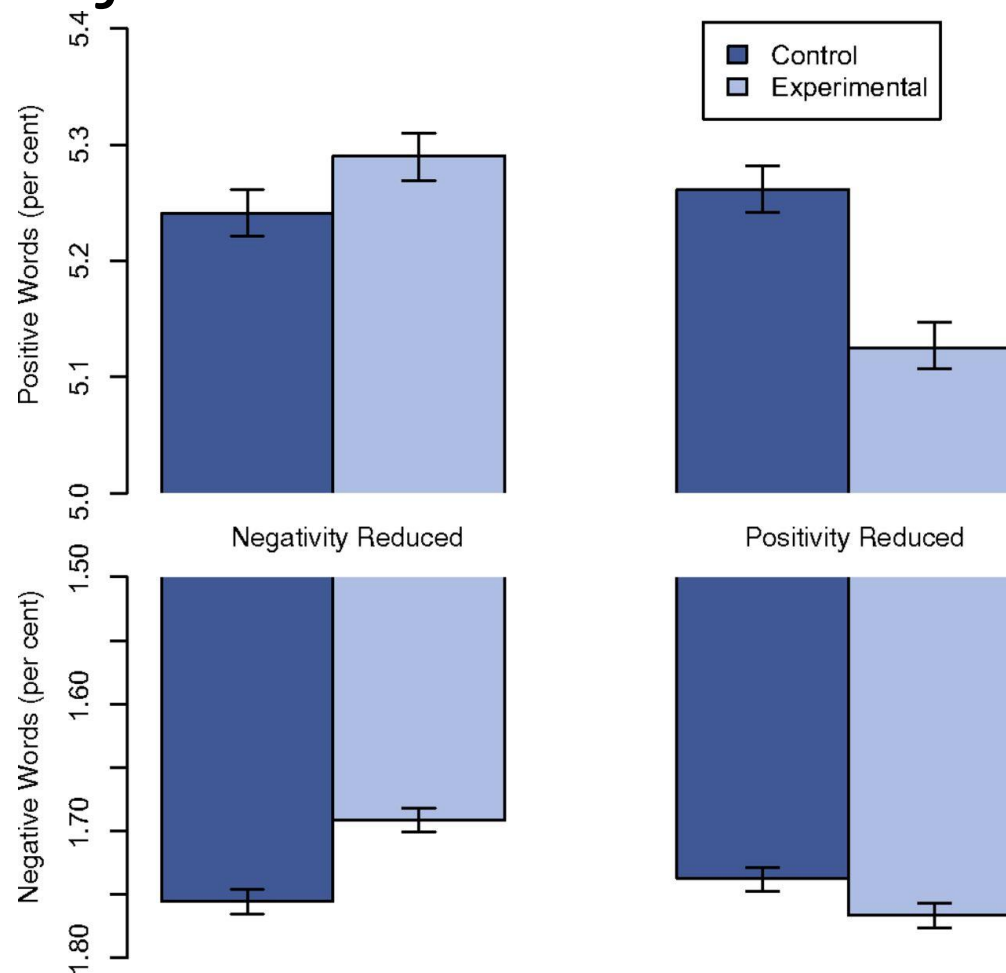
Manipuliranje čustev

- Čustvena nalezljivost
- 689.000 uporabnikov Facebooka
 - Zmanjšano število vsebin s pozitivnimi besedami
 - Zmanjšano število vsebin z negativnimi besedami



Vir: Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 111(29), 8788–8790. <https://doi.org/10.1073/pnas.1320040111>

Manipuliranje čustev



Vir: Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 111(29), 8788–8790. <https://doi.org/10.1073/pnas.1320040111>

Kazalo

- Uvod, digitalne sledi
- Napovedovanje
 - Preference
 - Osebnost
 - Spolna usmerjenost
- Uporaba
 - Usmerjeno oglaševanje
 - Politično oglaševanje
 - Manipuliranje čustev
- **Zaključek**

Zaključek

- Digitalne sledi vsebujejo mnogo informacij
- Področje je še vedno malo regulirano
- Sami poizkusimo omejiti sledi, ki jih puščamo
- Bodimo aktivni v debatah o zasebnosti, varovanju podatkov

Hvala za pozornost

Vprašanja?