

Depresszió keretezése online fórumokon

A társadalmi szenvedés-kutatás automatizált
szöveganalitikai eszközökkel

Sik Domonkos, Máté Fanni, Németh Renáta, Katona Eszter

Társadalmi szenvedés és depresszió

- Depresszió hegemón diskurzusa ¹ :
 - Mentális zavar/ betegség
 - Globálisan vezető ok a munkaképesség elvesztésére
 - Hatékony pszichoterápiás és gyógyszeres terápiákkal kezelhető
- Depresszió fenomenológiája:
 - Világ mint lehetőség helyett akadályok tere
 - Időtudat beszűkülése (jövő és múlt helyett jelen totalitása)
 - Interszubjektivitás, interaffektivitás, ágencia feloldódása
- Depresszió szociológiája: az interakciók és integrációs torzulásaiból fakadó társadalmi szenvedés

¹ <https://www.who.int/news-room/fact-sheets/detail/depression>

Depresszió diszkurzív keretezése

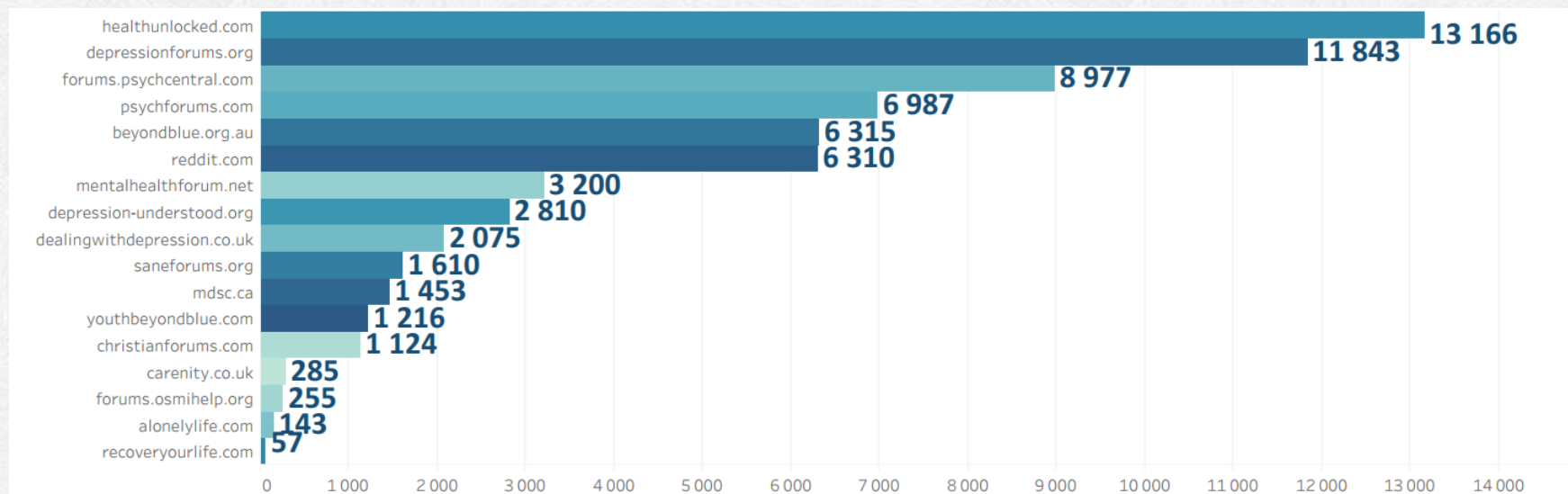
- Depresszió értelmezése a szenvedés konstitutív eleme
- Különböző keretezések eltérő attribúciós és kezelési pályákat nyitnak meg/ fednek el
- Ez jelöli ki a kutatás tétjét:
 - Az érintettek laikus diskurzusaiban milyen keretezések figyelhetők meg?
 - Ezek milyen deskriptív, expresszív és performatív sajátosságokkal jellemezhetők?
 - Dinamikájuk milyen lehetőségteret jelöl ki az érintettek önértelmezése számára?
- A kutatás terepe: depresszióval foglalkozó angol nyelvű online fórumok

Bio-medikális, pszichológiai és társadalmi keretezés

	Bio-medikális	pszichológiai	társadalmi
ontológiai modell	biológiai	kognitív-affektív tudat, viselkedés, személyiség	társadalmilag konstruált jelentéshorizont
attribúció (kialakulásának oka)	genetika, biokémiai, neurológiai	maladaptív kondicionális, megakadt/traumatizált személyiségfejlődés	társadalmi zavarok interiorizálása
megoldási lehetőség	testi zavarok kezelése (gyógyszer, beavatkozás)	torzult tanulási/adaptációs/személyiségfejlődési pálya korrekciója (terápia)	torzult személyközi és strukturális tér korrekciója

A kutatásról

- 17 honlap, depresszióval foglalkozó angol nyelvű fórum
- 79 889 poszt, 2016 februártól 2019 februárig
- Adatgyűjtés: SentiOne
- Nyilvános posztok



Annotálás

- A betanított (gyakran internetes felületeken, pl. Amazon Mechanical Turk toborzott) személyek felcímkézik az adatokat, amiből a gép tanul
- Az adatok annotálása széleskörben használt, megbízható eszköz piaci környezetben:
 - Macska van-e a képen?
 - Jó vagy rossz véleményt mondtak-e egy termékről?
- Ellenben társadalomtudományokban még kevésbé használt, hermeneutikailag bonyolult esetekben:
 - Gyűlöletbeszéd-e? Ki ellen irányul?
 - Hogyan értelmeződik a depresszió?

Annotálás a kutatásunkban

- 10 hallgató
- 4500 poszt
- Öt csoport: bio-medikális, pszichológiai, szociológiai, nem besorolható és irreleváns
- Két címke lehetősége – a posztok 34 százalékában volt második
- Oktatás – pilot kör – frissített guideline
- Első kör → két-két annotátor kódol
- Ha nem egyezik – szakértői kódolás (12,3%)

Az annotálás értékelése

- Cohen-kappa
 - Előnye: a besorolás véletlenből eredő egyezését is figyelembe veszi, több kategória esetén is használható
 - Két címke → módosított mérőszám
(Flor et al, 2016)
 - Mérsékelt egyezés (0,4-0,6)
(Landis-Koch, 1977)

$$k = \frac{p_o - p_e}{1 - p_e}$$

p_o=megfigyelt egyezés
p_e=véletlen egyezés

	Konzervatív	Liberális
Megfigyelt egyezés	58.3%	69.7%
Véletlenszerű egyezés	27.9%	28.3%
Cohen-kappa	0.42	0.58

Az eredmények értékelése

- A mérőszám értéke több dologtól függ:
 - Annotátorok – de: az „expert” kódoló sem feltétlenül jobb
 - Guideline – de: a guideline végletekig történő részletezése sem jó megoldás
 - Téma összetettsége – nem feltétlenül létezik egy és csak egy jó besorolás

Konklúzió

- Egyre elterjedtebb módszer, a társadalomtudományokban is
- Az annotálás minőségének vizsgálata is egyre fontosabb
- Általános problémát jelent
- Nem elkerülhető a folyamat és az emberi tényező figyelembe vétele:
 - Milyen módon értelmezünk egy szöveget/képet?
 - Az annotátorok milyen jellemzői (kulturális, demográfiai) befolyásolják a besorolást/értelmezést?

Terveink

- Az annotátorok közötti egyet nem értés, eltérő értelmezés okainak további vizsgálata
- Olyan módszertan kidolgozása, amellyel az annotálás értékelésekor figyelembe vehetők az annotátorok jellemzői is

Irodalom

- Flor, M., Yoon, S. Y., Hao, J., Liu, L., & von Davier, A. (2016, June). Automated classification of collaborative problem solving interactions in simulated science tasks. In Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications (pp. 31-41).
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33 1, 159-74 .