



conTEXT 2023  
Central European University  
2023. november 14.

# SZÓBEÁGYAZÁSOK ÉS NAGY NYELVMODELLEK TÁRSADALOMTUDOMÁNYI ALKALMAZÁSÁNAK PÉLDÁI

Rakovics Zsófia



ELTE | TáTK  
TÁRSADALOMTUDOMÁNYI KAR



Research Center  
for Computational  
Social Science



TÁRSADALOMTUDOMÁNYI  
KUTATÓKÖZPONT

# Bevezetés

## Motiváció

- Egy időbeli szóbeágyazási modell ismertetése és alkalmazásának demonstrálása
- Nagy nyelvmodellek társadalomtudományi alkalmazásának lehetőségei és korlátai

## Relevancia

- Számítógépes társadalomtudomány
  - Természetesnyelv-feldolgozás
    - A társadalom vizsgálata a nyelven (nagy mennyiségű szöveges adaton) keresztül
  - Gépi- és mélytanulási eljárások
    - Nagy mennyiségű adat automatizált feldolgozása, társadalmi összefüggések feltárása

# Egy újszerű időbeli szóbeágyazási vektortérmodell ismertetése és alkalmazásának lehetőségei



ELTE | TáTK  
TÁRSADALOMTUDOMÁNYI KAR



Research Center  
for Computational  
Social Science

HUN  
REN



TÁRSADALOMTUDOMÁNYI  
KUTATÓKÖZPONT

# Szövegtörpusz

## Körpusz

- Orbán Viktor beszédek
- 2007-2021
- 1000+ beszéd

## Előfeldolgozás

- Tokenizálás és lemmatizálás
- Szignifikáns n-grammok és névelemek felismerése
- Stopszavazás

## Az elemzett szövegtörpusz kapcsolódik

- az [ELTE TáTK Szociológia Doktori Iskolában](#) végett PhD kutatáshoz
- az [ELTE RC2S2](#) kutatócsoport [A politikai nyilvánosság rétegei Magyarországon \(2001-2020\)](#)\* című kutatásához is.

\* A Nemzeti Kutatási, Fejlesztési, és Innovációs Hivatal (NKFIH) által támogatott kutatás, azonosítója K-134428.

# Szóbeágyazási vektortérmodellek

A szóbeágyazási vektortérmodellek nyelvelméleti háttérét az eloszlási hipotézis adja (Firth 1957)

- Egy szónak nincsen inherens jelentése, annak kontextusa adja a jelentést

## Típusok

- Neurális háló alapú szóbeágyazás
- Együttes eloszláson alapuló szóbeágyazás

## Módszertan

- Az egyik legismertebb neurális háló alapú modell-implementáció a word2vec (Mikolov et al. 2013)
- Összehasonlító vizsgálatok (Wendlandt 2018), általános vélekedés szerint a neurális háló alapú szóbeágyazási modellek jobban teljesítenek
- Elemzés, mely szerint megfelelő beállítások mellett a két modell azonos (Levy et al. 2015)

# Pointwise Mutual Information

Az együttes eloszláson alapuló szóbeágyazáshoz kapcsolódó mérőszám

Pointwise Mutual Information (PMI) (Church & Hanks 1990)

$$PMI(w, c) = \log \frac{\hat{P}(w, c)}{\hat{P}(w)\hat{P}(c)}$$

ahol  $w$ : a szótár szavai,  $c$ : a kontextus szavak

$\hat{P}(w, c)$ : a  $w$  szó és a  $c$  kontextus szó együttes előfordulásának valószínűsége

$\hat{P}(w)\hat{P}(c)$ : a  $w$  szó és a  $c$  kontextus szó előfordulási valószínűségének szorzata

# Positive Pointwise Mutual Information

A PMI mátrix negatív értékeinek figyelmen kívül hagyása

- A figyelem azokra a szópárokra összpontosul, amelyek a véletlenhez képest gyakrabban járnak együtt

Positive Pointwise Mutual Information (PPMI) (Levy et al. 2015)

$$PPMI(w, c) = \max(PMI(w, c), 0)$$

ahol  $w$ : a szótár szavai,  $c$ : a kontextus szavak

$PMI(w, c)$ : a  $PMI$  mátrix értéke a  $w$  szó és a  $c$  kontextus szó párra

$\max(PMI(w, c), 0)$ : a maximuma a  $PMI$  értékének és a nullának

# Temporal Positive Pointwise Mutual Information

## Időbeli szóbeágyazási modellek

- A PPMI-hez hasonló beágyazás (Yao et al. 2018)
- Mikolov féle sekély neurális háló Kálmán-szűrős kiegészítése (Bamler és Mandt 2017)

## Temporal Positive Pointwise Mutual Information (TPPMI) (Rakovics 2022)

- Dinamikus szóbeágyazási modell, az időbeli változás (időegység lehet például egy hónap) lekövetésére képes szóbeágyazási vektortérmodell
  - Különböző hónapokból származó szövektorok egymáshoz képesti viszonyát lehet modellezni
  - A kontextus szavak megegyeznek a hónaponként számolt PPMI mátrixokban
  - Mozgó átlagoláshoz hasonló eljárással dolgozik: az aktuális hónap szövektorának kiszámolásánál a környező hónapok szövektorait is figyelembe veszi



# Elemzés és a szóvektorok értelemezése

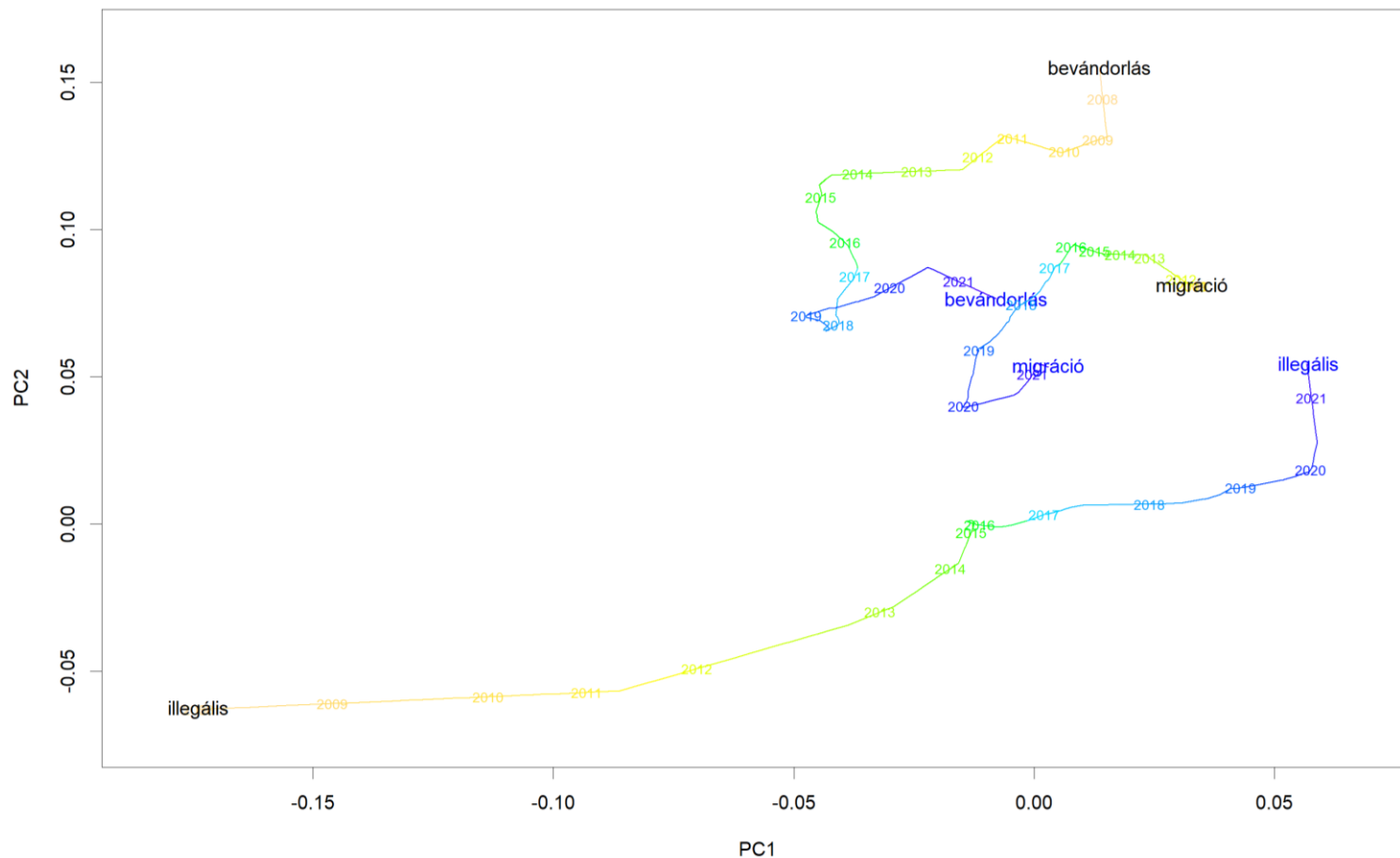
## Elemzés és értelmezés

- A Temporal Positive Pointwise Mutual Information (TPPMI) szóbeágyazás használata
- A vektorok egymáshoz képesti helyzete a vektortérben, illetve azok változása jól megragadható
  - Egy adott szó szóvektorának időbeli változása alapján szó jelentésváltozására következtethetünk
  - A szóvektorok egymáshoz képesti változása alapján a vizsgált szavak jelentéseinek egymáshoz képesti alakulására következtethetünk

## Vizualizáció

- Két dimenzióban grafikusán is jól követhetők a vektorok időbeli trajektóriái

# Eredmények (Rakovics 2022)



## Vizsgált kulcsszavak

- bevándorlás
- migráció
- illegális

Az ábrán a **szóvektorok** láthatók

- Vízszintes tengely: az első főkomponensen mért értékek
- Függőleges tengely: a második főkomponensen mért értékek

**Időbeli elmozdulás:** sárgából (2007) a kékbe (2021) hajló színskála

# Nagy nyelvmodellek társadalomtudományi alkalmazásának lehetőségei és korlátai



ELTE | TáTK  
TÁRSADALOMTUDOMÁNYI KAR



Research Center  
for Computational  
Social Science



Új Nemzeti  
Kiválóság Program



NEMZETI KUTATÁSI, FEJLESZTÉSI  
ÉS INNOVÁCIÓS HIVATAL

# Adatok és módszertani keret

## Adatbázis

- European Social Survey (ESS)
- 2020-2021 (10. hullám)
- 30+ ország, köztük Magyarország
- Kérdésblokkok (pl. politika, demokrácia)

## Kutatási kérdés

- Használható-e a GPT mint virtuális válaszadó?

## Módszertani keret

- Virtuális válaszok generálása a GPT-3.5 Turbo nyelvmodell-lel
  - A virtuális válaszadó szociodemográfiai tulajdonságainak és bal-jobb skálán mért politikai beállítódásának meghatározása a prompt-oláskor
  - Minden válaszadóhoz 10 virtuális válasz generálása, majd azok összegzése és értékelése
- Valódi és virtuális válaszok szisztematikus kiértékelése és az eredmények összehasonlítása

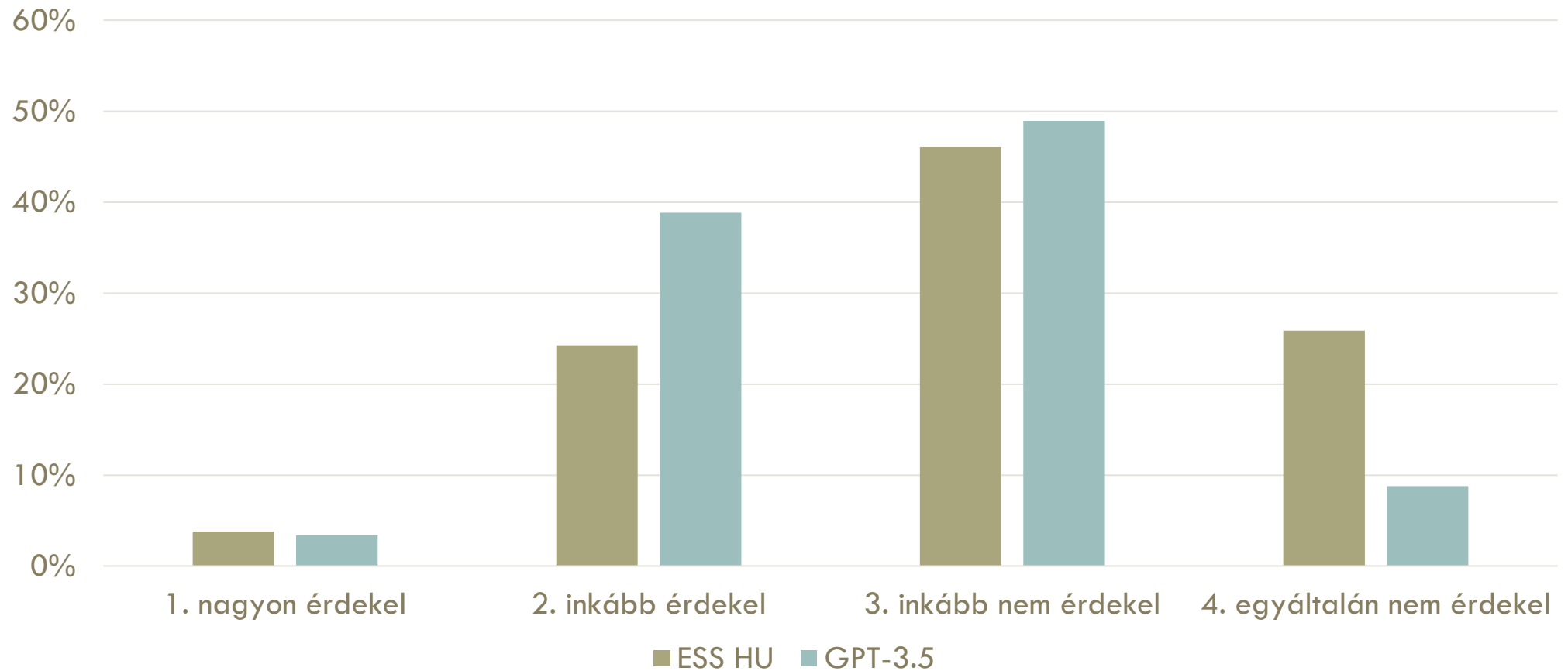
# Virtuális válaszok generálása

Prompt engineering (Yao et al. 2023)

Prompt meghatározása

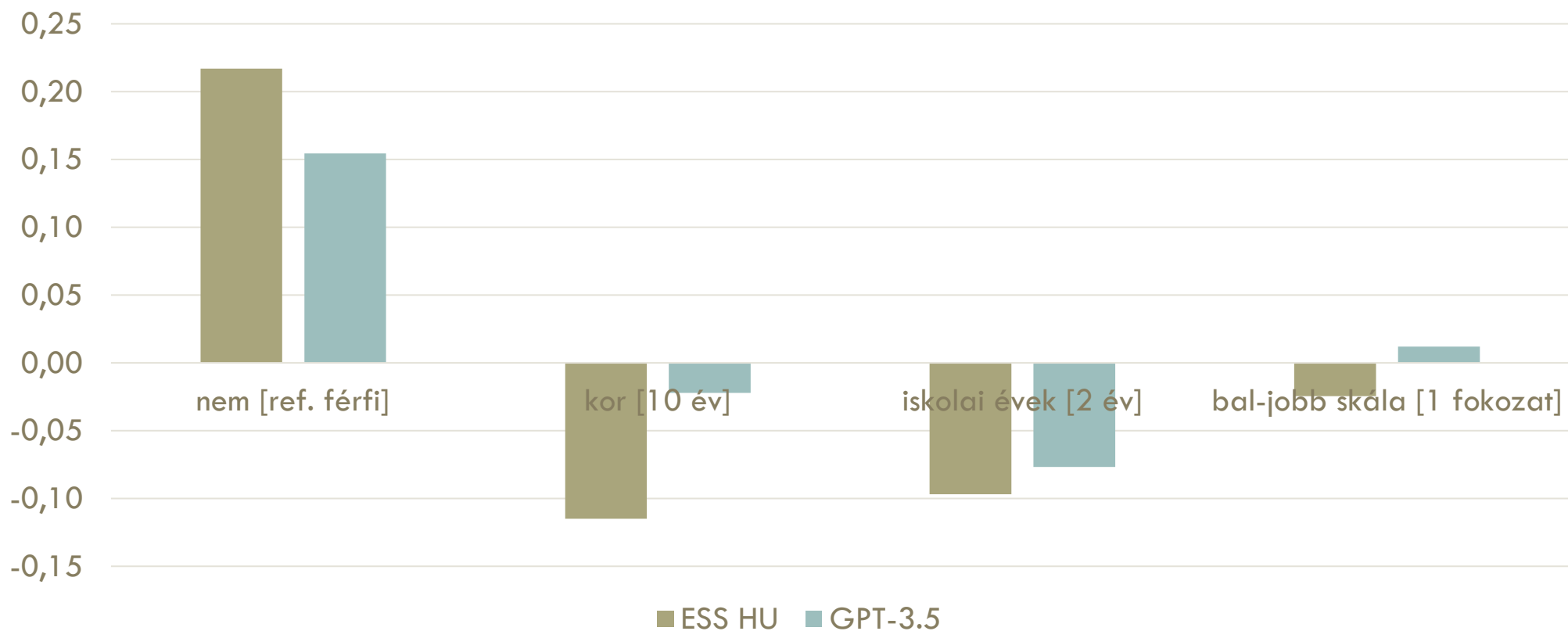
- Virtuális válaszadó attribútumai: „Pretend you are a Hungarian 42-year-old female with 16 years of education who is very right leaning politically.”
  - Magyar válaszadók használata
  - Nem, életkor és iskolai évek száma a válaszadók tényleges neme, életkora és iskolai évek száma alapján került meghatározásra
  - A bal-jobb skálán mért politikai beállítódás fokozatai a left/right, very/slightly szavakkal kerültek beállításra
- Kérdés: „How interested would you say you are in politics? Are you 1. very interested, 2. quite interested, 3. hardly interested, 4. not at all interested? Answer with the listed options!”

# Mennyire érdeklí Önt a politika?



# Regressziós együtthatók

politikai érdeklődés ~ nem + életkor + iskolai évek száma + bal-jobb skála értéke



# Összegzés

## Időbeli szóbeágyazási vektortérmodellek

- Szavak relatív jelentésváltozásának kvantitatív vizsgálatát teszik lehetővé
- Érdeemes lehet kiegészíteni ezt a megközelítést kvalitatív elemzéssel is

## Nagy nyelvmodellek társadalomtudományi alkalmazása lehetőségei és korlátai

- Lehetséges-e a kérdőíves kutatásokat nagy nyelvmodellekkel helyettesíteni?
  - Valódi és virtuális válaszadók válaszainak kiértékelése

## Számítógépes társadalomtudomány

- Módszerek lehetőségeinek és korlátainak feltérképezése
- Kritikus, reflexív és tudományos szemlélet



# Irodalomjegyzék (1)

- Bamler, R., & Mandt, S. (2017). Dynamic word embeddings. In *International conference on Machine learning* (pp. 380-389). PMLR.
- Church, K., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational linguistics*, 16(1), 22-29.
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*.
- Levy, O., Goldberg, Y., & Dagan, I. (2015). Improving distributional similarity with lessons learned from word embeddings. *Transactions of the association for computational linguistics*, 3, 211-225.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

## Irodalomjegyzék (2)

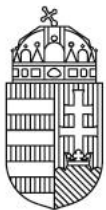
- Rakovics, Zsófia. (2022). A Temporal Positive Pointwise Mutual Information (TPPMI) időbeli szóbeágyazási modell alkalmazásában rejlő lehetőségek demonstrálása: A miniszterelnöki beszédek szavainak jelentésváltozása. In: Feledy, A.; Egle, B. (szerk.) *Van új a nap alatt: Az ELTE Angelusz Róbert Társadalomtudományi Szakkollégium konferenciájának tanulmánykötete*. Budapest, Magyarország. 67 p. pp. 31-48. 18 p.
- Wendlandt, L., Kummerfeld, J. K., & Mihalcea, R. (2018). Factors influencing the surprising instability of word embeddings. *arXiv preprint arXiv:1804.09692*.
- Yao, Z., Sun, Y., Ding, W., Rao, N., & Xiong, H. (2018). DynamicWord Embeddings for Evolving Semantic Discovery. In *Proceedings of the eleventh ACM international conference on web search and data mining* (pp. 673-681).
- Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T. L., Cao, Y., & Narasimhan, K. (2023). Tree of thoughts: Deliberate problem solving with large language models. *arXiv preprint arXiv:2305.10601*.

# KÖSZÖNÖM A FIGYELMET!

Rakovics Zsófia

ELTE TáTK Szociológia Doktori Iskola, Társadalomkutatások Módszertana Tanszék  
Research Center for Computational Social Science  
HUN-REN Társadalomtudományi Kutatóközpont

A Kulturális és Innovációs Minisztérium ÚNKP-23-3 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.



KULTURÁLIS ÉS INNOVÁCIÓS  
MINISZTERIUM



NEMZETI KUTATÁSI, FEJLESZTÉSI  
ÉS INNOVÁCIÓS HIVATAL

**UNKP** Új Nemzeti  
Kiválóság Program