

# A magyarországi antiszemitizmus vizsgálata NLP topikmodell segítségével

## **Barna Ildikó**

Habilitált egyetemi docens, tanszékvezető, ELTE TáTK Társadalomkutatások Módszertana Tanszék

## **Knap Árpád**

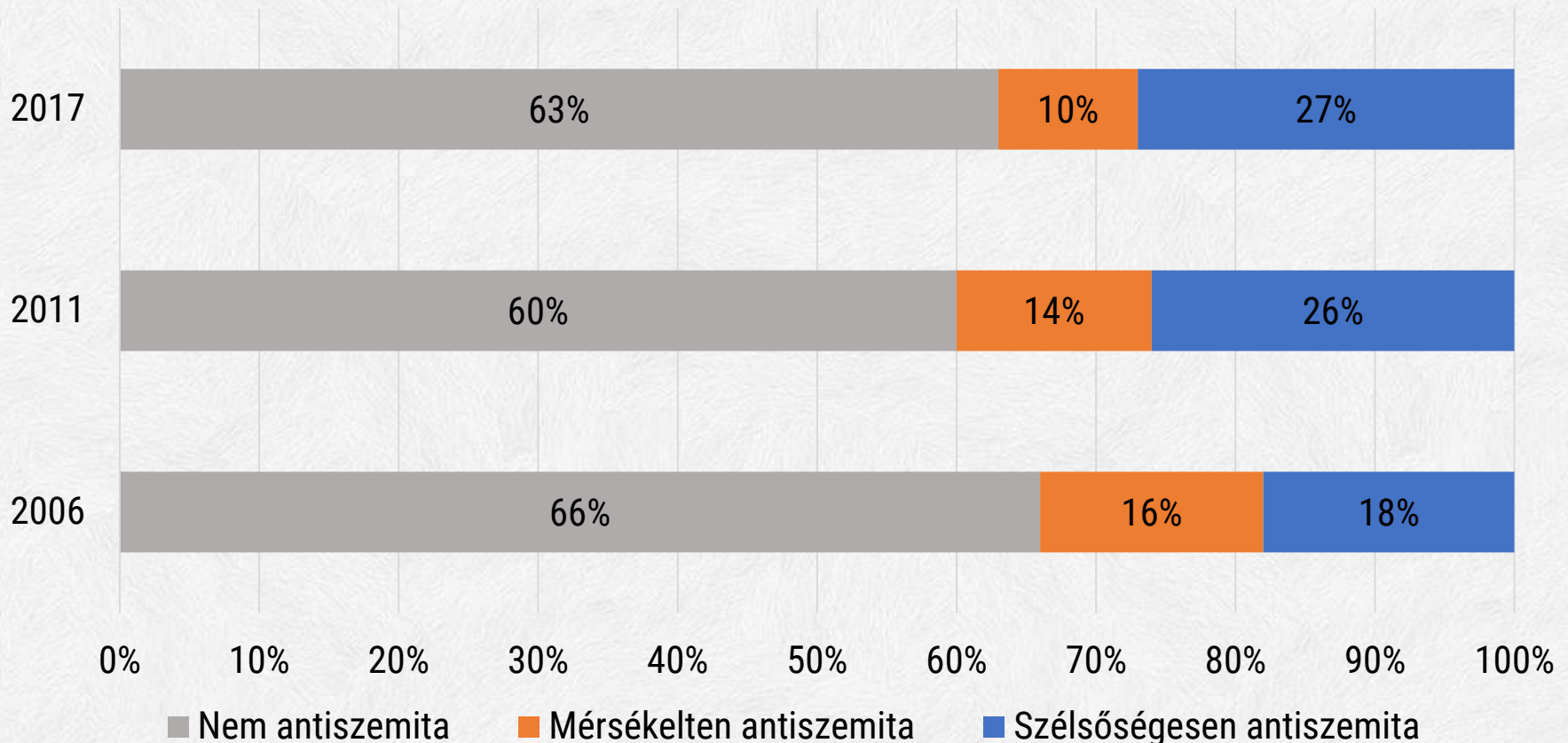
PhD hallgató, ELTE TáTK Interdiszciplináris társadalomkutatások doktori program

A társadalom kutatásának új útjai workshop, 2019. október 11.

# A kutatás áttekintése

- A projekt célja: az online gyűlöletbeszéd kutatása
- Az első lépés: antiszemitizmus kutatása olyan szöveges korpuszokon, amelyekben biztosan van antiszemita tartalom
- A korpuszok: cikkek és kommentek a szélsőjobboldali Kuruc.info portálról
- Fő kutatási kérdések
  - Mi a látens tematikus struktúrája a két korpusznak?
  - Eltér-e a két korpusz tematikus struktúrája?
  - A topikokban megjelennek-e az antiszemitizmus különböző típusai?

# Antiszemizmus Magyarországon

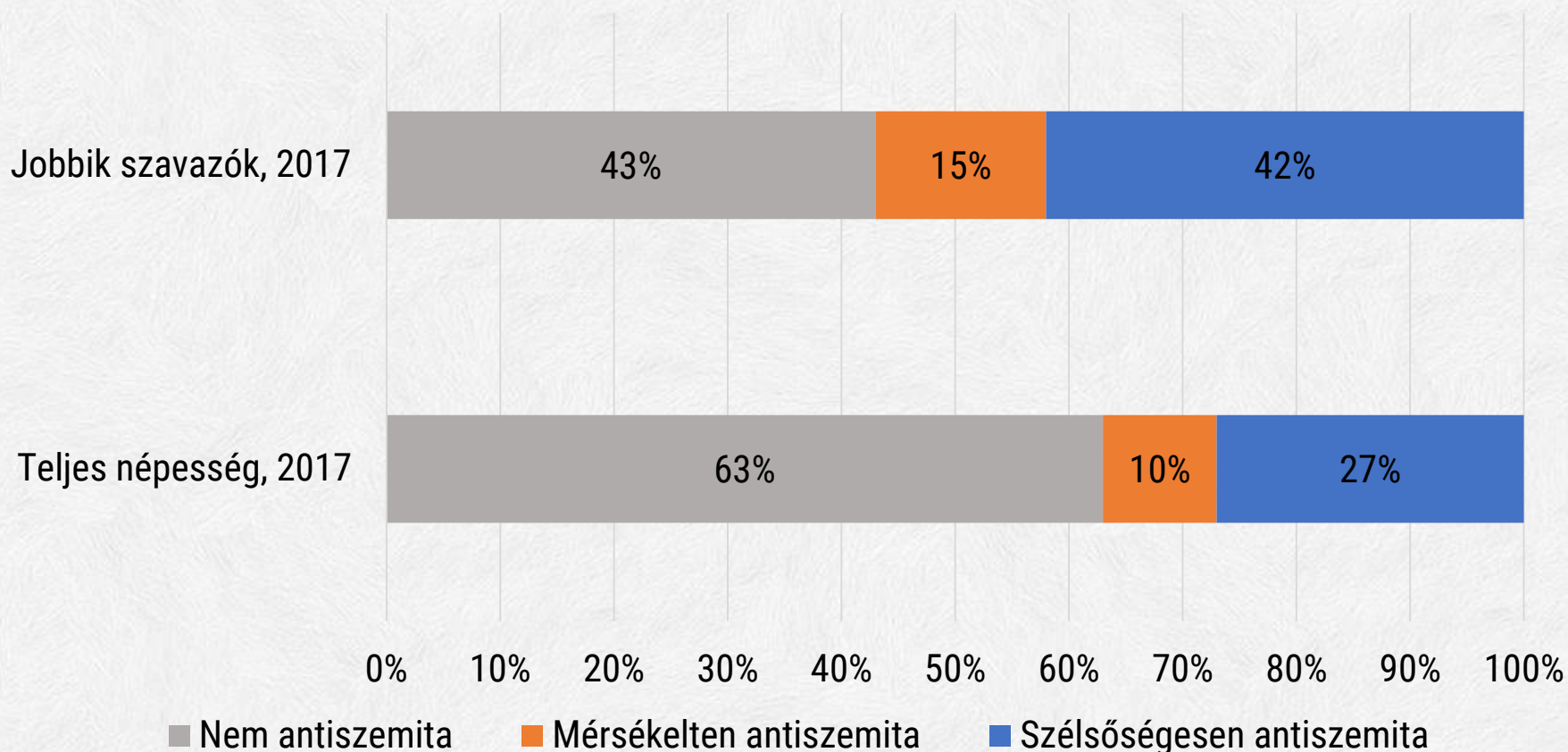


Forrás: TEV, Medián

# A Jobbik és a Mi Hazánk Mozgalom

- 2003: megalapítják a Jobbikot
- 2009-től jelentős politikai erő
- 2010: 17; 2014: 20 százalékos eredmény az országgyűlési választásokon
- 2014: stratégiaváltás, centrista, néppárti image
  - Lemondások, kizárások
  - Novák Előd és más radikális politikusok távozása
  - 2018: Mi Hazánk Mozgalom megalakulása

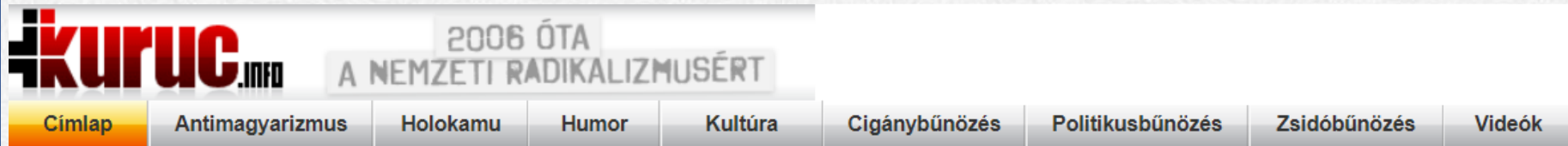
# Pártpreferencia és antiszemitizmus



Forrás: TEV, Medián

# A Kuruc.info hírportál

- Szélsőjobboldali, nacionalista hírportál, 2006 óta működik
- 2009: az 51. legtöbbet látogatott magyar weboldal (Alexa)
- 2018: a tizedik legnagyobb magyar hírportál (Medián)
  - A teljes népesség 10,
  - A Jobbik szavazók 23 százaléka látogatja legalább alkalmanként
- Nyíltan antiszemita és rasszista tartalmak



# Kuruc.info, Jobbik, Mi Hazánk

- A Jobbik online kommunikációjának fő platformja éveken keresztül a Kuruc.info volt
- 2019. április: sajtóesemény az egykori MIÉP irodában
  - Bemutatnak három újságíró-t és Novák Elődöt
  - Nyílt kiállítás a Mi Hazánk Mozgalom mellett

# A korpusz tulajdonságai

- Adatok letöltése: SentiOne
- Egyetlen keresőszó: „zsidó”
- Cikkek és kommentek a Kuruc.info-ról, 2016 márciusa és 2019 márciusa között
- 2 289 cikk és 51 060 komment elemzése



# Előfeldolgozás

1. URL-ek eltávolítása
2. Lemmatizálás
3. Tulajdonnév-felismerés (Named Entity Recognition)
4. Szignifikáns bi- és trigramok felismerése
5. Stopszavak definiálása
6. Cserelista definiálása
7. Gyakoriság alapú szűrés

# LDA (Látens Dirichlet Allokáció)

- Topikmodellezési eljárás nagy méretű szöveges korpuszokra
- Cél: egy korpusz látens, rejtett tematikus struktúrájának feltárása topikok létrehozásával
- Konceptió: egy adott topik jellemezhető a hozzá tartozó szavakkal
- Topik-koherencia mutató elemzése

# Tematikus struktúra: vallási antiszemitizmus (antijudaizmus)

- Mindkét korpusz estében
  - Alacsony prevalencia
  - Hiányzik a Krisztus-gyilkosság klasszikus vádja
- Cikkek: a zsidók meg akarják „mérgezni” a keresztényeket
- Kommentek: a keresztények a zsidók felett állnak

# Tematikus struktúra: faji alapú antiszemitizmus

- Cikkek: kiemelkedő gyakoriság
- Kommentek: szinte teljesen hiányzik
- A nagy látencia-nyomás miatt survey módszerekkel kevésbé megragadható

*~ A zsidók azért terjesztik a homoszexualitást, hogy ezzel elősegítsék a „fajok keveredését” és a terjesszék a „fehérellenes sztereotípiákat”*

*~ A zsidók „fajvédelmi akcióikkal provokálják a magyarságot”*

# Tematikus struktúra: összeesküvés alapú antiszemitizmus

- A cikkekben kevésbé jelenik meg és a faji alapú antiszemitizmushoz kapcsolódik
  - A kommentekben nagyon gyakori és önmagában is megjelenik
- ~ „A beteges, a perverz és a káros dolgok normálissá tétele a zsidó aktivizmus visszatérő módszere, legyen az politika vagy média.”
- ~ A cigányok és a bevándorlók mint a zsidók biológiai fegyverei

# Tematikus struktúra: másodlagos antiszemitizmus

- Gyakoribb a cikkek esetében
  - Mindkét korpuszban előfordul *hard-core* holokauszttagadás, de gyakoribb a *soft-core* holokauszt relativizálás, torzítás
  - A zsidók hasznot húznak a holokausztból
  - Cikkek esetében külön topik: áldozati versengés
- ~ *Visszatérő kifejezések: holoipar, holokamu, hololegendárium, holokamuszt, holokauszt dogmatika*
- ~ *„Úgynevezett holokauszt”, „állítólagos hitleri gázkamrák és a zsidók állítólagos kiirtása*

# Tematikus struktúra: új antiszemitizmus

- Natan Sharansky (2004) 3D teszt
  - Demonizáció, kettős mérce (double standard), delegitimáció
- Alacsony prevalencia mindkét korpuszban
- Gyakran nincs explicit antiszemitizmus, de rendkívül egyoldalú a nézőpont, nem állná ki a 3D tesztet
- Sok idézőjel-használat: például „béketeremtés”, „védelmi” miniszter

# Konklúzió

- Sikerült jól interpretálható topikokat azonosítani
- A cikkek és a kommentek eltérő karakterrel rendelkeznek
- Az antiszemitizmus klasszikus típusai megjelentek mindkét korpuszban
  - Ugyanakkor ezek a típusok gyakran keverednek egymással
- További tervek
  - Tematikus bővítés: az online gyűlöletbeszéd más formáinak vizsgálata, az antiszemitizmus rejtettebb formáinak vizsgálata
  - Módszertani bővítés: más NLP-s módszerek használata, például szóbeágyazási modell



# Köszönjük a figyelmet!

Megjelent cikkünk: **Barna Ildikó, Knap Árpád. "Antisemitism in Contemporary Hungary: Exploring Topics of Antisemitism in the Far-Right Media Using Natural Language Processing." In *Theo Web – Academic Journal of Religious Education* 18(2019), H.1, 75–92 (2019).**

További információk találhatóak a kutatócsoport weboldalán: <https://rc2s2.eu>