



A felügyelt gépi tanulás kihívásai a szociológiai alkalmazásokban

NÉMETH RENÁTA¹

ABSZTRAKT

Az ipari/üzleti alkalmazásokban már sokszorosan bizonyított felügyelt gépi tanulás szociológiai alkalmazásai sajátos kérdéseket vetnek fel. A sajátosság oka, hogy ezekben az alkalmazásokban komplex fogalmak megtanulása az algoritmus feladata (pl. hogy gyűlöletbeszédet tartalmaz-e egy tweet). A felügyelt tanulás lényege, hogy előre bekódolt (gyűlöletbeszéd/nem gyűlöletbeszéd) szövegek címkézését tanulja meg az algoritmus, jellegzetes szövegmintázatokat keresve. A felmerülő kérdések: hogyan jön létre a címkézés? Hogyan lehet betanított kódolókkal elvégeztetni egy olyan hermeneutikai kihívást, mint a gyűlöletbeszéd felismerése? Segítenek-e ezen a rutinszerűen alkalmazott, részletezett annotálási irányelvek? A cikk arra is kitér, hogyan végzik crowdsourcing platformokon a kódolást a nagy cégek, illetve ismertetem az MI-torzítást is, aminek itt az a lényege, hogy a kódolók maguk viszik be a diszkriminációt az adatokba. E kérdéseket kutatási tapasztalatainkkal illusztrálom.

KULCSSZAVAK: felügyelt gépi tanulás, annotálás, crowdsourcing, MI-torzítás

ABSTRACT

The challenges of supervised machine learning in sociological applications

The sociological applications of supervised machine learning, already well proven in industrial/business applications, raise specific questions. The reason for this specificity is that in these applications, the algorithm is tasked with learning complex concepts (e.g. whether a tweet contains hate speech). Supervised learning consists of learning to classify previously annotated (hate speech/non-hate speech) texts by the algorithm, looking for characteristic text patterns. The questions that arise are: how to prepare annotation? How can a hermeneutic challenge such as hate speech recognition be performed by annotators? Are routinely applied, detailed annotation guidelines helpful? The article also discusses how large companies perform coding on crowdsourcing platforms, and describes AI bias, which in this case means that annotators themselves introduce bias into the data. I illustrate these issues with our own research experiences.

KEYWORDS: supervised machine learning, annotation, crowdsourcing, AI bias

¹ Eötvös Loránd Tudományegyetem, Társadalomtudományi Kar, Research Center for Computational Social Science, email:nemeth.renata@tatk.elte.hu. A szerzőt a K-134428 azonosítójú NKFIH pályázat támogatta kutatási munkájában.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Bevezetés

A weboldalakon, közösségi média felületeken, szöveggé alakított videókon, digitalizált könyvtárakban felhalmozódott szövegek a társadalmi valóság sosem látott szélességű elérését teszik lehetővé. Individuális és kollektív cselekedeteinkről, szerepeinkről és csoportközi viszonyainkról, mindezek időbeli dinamikájáról és strukturális heterogenitásáról naponta többmillió terrabytenyi digitális szöveg keletkezik, s a meglévő adatvagyon néhány évente megsokszorozódik.

A digitális szövegeknek ez az univerzuma párhuzamosan olyan szöveganalitikai technológiákat hívott életre, melyek ennek az elérésnek ma már a szociológia számára is releváns mélységét nyújtják. Savage és Burrows 2007-ben az évtized egyik legtöbbet idézett szociológiai tanulmányában az empirikus szociológia közelgő válságáról írt. Azt jóslták, hogy válság következik be, ha a korábban saját módszertani szakértelméről ismert szociológia nem tud megfelelni a big data által támasztott kihívásoknak, és így elveszíti vezető szerepét. Ez nem következett be. Nyolc évvel később a British Sociological Association által kiadott „Sociological Futures” című könyvsorozat első tagja (Ryan – McKie 2015) már a címében is utalt a válság végére, és fontos lehetőségeket látott a big data kutatásban, valamint az szövegbányászatban is. Ugyanakkor a korábban a szociológia által uralt empirikus szakértelmet tekintve egyértelmű elmozdulás figyelhető meg az akadémiai szférából az ipar felé, hiszen a terület hatalmas üzleti lehetőségeket generál és az ipar finanszírozni is képes a szükséges fejlesztéseket.

Ez a cikk a szövegbányászat egy alterületét mutatja be. A szövegbányászat izgalmas, és a szociológia számára is nagyon perspektivikus terület az informatika, mesterséges intelligencia-kutatás és nyelvészet határán. Utóbbi években tapasztalt fel-futása mögött egyszerre áll a digitális szöveges források elérhetősége és a hatalmas adatbázisok feldolgozásához szükséges technológia fejlődése. Ambiciózus projektek használják üzleti alkalmazásokban, marketingben, kormányzati és védelmi területen egyaránt. Az elmúlt néhány évben a szövegbányászat a társadalomtudományokban is kezdett teret nyerni, az antropológiától a közgazdaságtanon át a szociológiáig². Természetesen, míg a szövegbányászat egy viszonylag új interdiszciplináris terület, maga a szövegelemzés több évtizedes tradícióra tekint vissza a társadalomkutatásban: kifinomult kvalitatív módszerek fejlődtek ki. A kvantitatív társadalomkutatási szövegelemzés inkább csak bizonyos kifejezések vagy kvalitatíve azonosított kódok megjelenését számszerűsítette a szövegekben, ehhez képest óriási előrelépést jelent a szövegbányászat eszköztára, ami olyan feladatok elvégzését teszi lehetővé, mint tematikus struktúrák vagy látens szemantikai relációk automatizált azonosítása (Németh – Katona – Kmetty 2020).

² Társadalomkutatónak szóló kiváló összefoglaló a módszerről: Ignatow – Mihalcea 2017.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Zárójelben megjegyezném, hogy a szövegbányászattal kapcsolatban sem az angol, sem a magyar terminológia nem kiforrott még, a természetesnyelv-feldolgozás (NLP, natural language processing), számítógépes nyelvészet, automatizált szövegelemzés diffúz körvonalakkal bír, rokon, de nem szinonim elnevezések (Németh – Katona – Kmetty 2020).

Ebben a tanulmányban azokat a lehetőségeket tekintem át, amelyeket a felügyelt gépi tanulás nyújthat a szociológia számára. Ezeknek az ipari/üzleti alkalmazásokban már sokszorosan bizonyított algoritmusoknak a szociológiai alkalmazásai sajátos kérdéseket vetnek fel. A sajátosság oka, hogy ezekben az alkalmazásokban komplex fogalmak megtanulása az algoritmus feladata. A felügyelt tanulás lényege, hogy előre bekódolt szövegek címkézését tanulja meg az algoritmus, jellegzetes szövegmintázatokat keresve. A felmerülő kérdések: hogyan jön létre a címkézés? Hogyan lehet betanított kódolókkal elvégeztetni egy olyan hermeneutikai kihívást, mint a gyűlöletbeszéd felismerése? Segítenek-e ezen a rutinszerűen alkalmazott, részletezett annotálási irányelvek? Jobban végzi-e a kutató a besorolást, mint az egyszerű kódoló, vagyis magasabbrendű-e az egyik interpretáció, mint a másik? A cikk arra is kitér, hogyan végzik crowdsourcing platformokon a kódolást a nagy cégek, hogy működik iparszerűen ez a humán/gép együttműködés, és milyen kérdések merülnek fel a crowdsourcing interpretáció kapcsán. Végül röviden kitérek az AI-torzításra, aminek itt az a lényege, hogy a kódolók maguk viszik be a diszkriminációt az adatokba.

Felügyelt vs. felügyelet nélküli gépi tanulás jellemzői

Az alábbiakban röviden ismertetem a felügyelt és felügyelet nélküli tanulás logikáját, majd a felügyelt tanulás gyakorlati kivitelezését. Nem térek ki a módszer szöveganalitikai jellegzetességeire, arra, hogy mi a „szöveg”, mint input adat sajátossága a numerikus inputtal szemben, vagy hogy milyen módon keresnek a modellek mintázatokat a szövegekben. Utóbbi kérdések iránt érdeklődő olvasóknak a Németh – Koltai (2021) szerzőpároshoz, illetve Németh, Katona és Kmetty (2020) szerzők szociológusokhoz szóló szövegbányászati bevezetőjét ajánlom.

A felügyelt és felügyelet nélküli tanulás mögött álló egzakt statisztikaelméletet Vapnik (2000) alaposan tárgyalja. Intuitíve, a felügyelt és felügyelet nélküli tanulás közötti különbség azon alapszik, hogy már létező elmélet/meglevő háttérismeret empirikus megnyilvánulásait keressük (felügyelt tanulás, ahol a „felügyelet” maga a háttérelmélet), vagy induktív módon egy még nem vizsgált téma feltárása a cél (felügyelet nélküli tanulás). A dichotómia hasonlít a klasszikus szociológia konfirmatív/exploratív módszerek közötti különbségtételéhez.

A nem felügyelt módszerek némelyikét a klasszikus társadalomkutatásban is gyakran alkalmazzák, ilyen például a szövegekre alkalmazott klaszterelemzés. Ez olyan vektortérben definiálható, ahol a tengelyek a szavakat jelölik, a dokumentu-



TEMATIKUS TANULMÁNYOK – Digitális szociológia

mok vektortérbeli helyét pedig az egyes szavak dokumentumbeli gyakorisága határozza meg. De nem felügyelt módszer a legújabb szövegbányászati megoldások alapját adó szóbeágyazás³ is. Ez a módszer az egyes szavakhoz rendel jelentést, még pedig ún. disztribúciós szemantikai megközelítésben, amikor is a jelentést kizárólag a használati környezet határozza meg. A jelentésről tehát kizárólag címkézetlen adatokból tanul az algoritmus, nincs szükség kívülről bevitt szakértői tudásra. A disztribúciós szemantika használható a legkülönbözőbb területek hatalmas skáláján, így például a gépi fordításnál is.

Ezzel szemben a felügyelt tanulás lényege, hogy (elméletünk/háttérismeretünk-re támaszkodva) előre bekódoljuk a szövegeket, majd ezeket a címkéket az algoritmus megpróbálja megtanulni. A kutató elméleti megfontolásai befolyásolják az elemzést, hiszen a címkék, mint kategóriák meghatározása megelőzi az elemzést. Egy társadalomkutatói példa Poletti és társai (2017) munkája, akik olasz nyelvű Twitter-üzeneteken igyekeztek automatikus gyűlöletbeszéd-felismerőt létrehozni. A gyűlöletbeszéd általuk alkalmazott definíció szerint valamely kisebbségi csoport ellen irányul és tartalmaz egyfajta illokúciós erőt, amely alkalmas a célcsoporttal szembeni erőszak terjesztésére, népszerűsítésére, alátámasztására vagy erre történő uszításra. Ezért a fogalom pontos megragadása érdekében a kódoknak tartalmaznia kellett a célcsoportot (vallási, etnikai kisebbségek vagy migránsok), és a kódolók a tweetet olyan jegyekkel is felruházták, mint hogy sztereotipizál-e, tartalmaz-e agressziót, támadó-e, vagy tartalmaz-e az agressziót kendőző iróniát.

A megtanulás itt azt jelenti, hogy (viszonylag alacsony hiba mellett) az algoritmus maga is képes lesz kódokat (gyűlöletbeszéd/nem gyűlöletbeszéd) rendelni még címkézetlen szövegekhez. A tanulási folyamat pedig a kutatók által címkézett két szöveghalmaz eltérését leginkább megfogó jellegzetes szövegmintázatok keresésén alapul, ahol az előre címkézett szövegek halmazát tanuló-halmaznak nevezzük. Nagyon leegyszerűsített példán: használhatunk logisztikus regressziót, ahol a függő változó a bináris címke (gyűlöletbeszéd/nem gyűlöletbeszéd), numerikus magyarázó változóink pedig azt jelölik, hogy az adott nyelv szavai hányszor fordulnak elő a szövegben; a cél a legjobban illeszkedő modellhez tartozó együttthatók megtalálása. A gyakorlatban persze ennél jóval komplexebb modelleket alkalmaznak, komplexebb (például stiláris jegyeket reprezentáló vagy a szöveg mélyebb szemantikai viszonyait megragadó) magyarázó változókkal, a nagy mennyiségű (több ezer) magyarázó változó miatt dimenziócsökkentő megoldásokkal és a túlillesztést elkerülni hivatott megoldásokkal, de az algoritmus létrejöttének logikája ott is hasonló.

A felügyelt osztályozás gyakran használt típusa a szentiment – vagy érzelem-elemzés. Egy szöveg szentimentje a szerző attitűdje egy tárgyhoz (pozitív, negatív vagy semleges), míg az érzelmek a boldogságtól a haragig terjedő érzések.⁴ Ez az

³ Angolul word embedding, a módszerről lásd: Németh – Katona – Kmetty 2020.

⁴ Összefoglaló az alkalmazott módszerekről: Nahili – Rezeg – Kazar 2020.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

alkalmazás mind üzleti, mind tudományos alkalmazásokban nagyon elterjedt. Szöveges adatok szentiment- vagy érzelem szerinti címkézésével vizsgálhatják például marketingesek, hogy hogyan reagálnak a felhasználók reklámokra, szolgáltatásokra vagy termékekre, digitális bölcsészek, hogy hogyan változnak a regényben megjelenített érzelmek, vagy szociológusok, hogy hogyan terjednek érzelmek és vélemények társadalmi hálózatokban.

A felügyelt módszerek abban támogatják tehát a kutatókat, hogy nagyobb szövegkorpuszt rövidebb idő alatt dolgozhassanak fel, mint amire az emberi kapacitás képes lenne. További (inkább tudományos elvárásokhoz igazodó) cél, hogy megértsük az automatikus címkézés mögött álló szabályokat, egyszerű példán: hogy lássuk, mely szóhasználat valószínűsíthető leginkább a gyűlöletbeszéd jelenlétét. Ez utóbbi cél tulajdonképpen azoknak a tartalmi-szociológiai interpretációra/magyarázatra lehetőségét adó jegyeknek a feltárását jelenti, amelyek az egyébként „fekete dobozként” működő kódoló algoritmus mögött állnak⁵.

A felügyelt tanulásnak több új továbbfejlesztése létezik, ahol az eredeti logika kissé módosul. Így pl. az aktív tanulás esetén nem egy fix tanuló-halmazunk van, hanem az algoritmus maga kéri menet közben konkrét, még címkézetlen, de a tanulásban fontosnak tűnő szövegek címkézését, vagy a transzfer tanulás, amikor a címkézett adatoktól eltérő besorolási problémát kell a tanuló algoritmusnak megoldania (Eisenstein 2019).

A felügyelet nélküli módszerek a felügyelt technikákkal együtt is használatosak. Például a felügyelet nélküli módszerek hatékony alkalmazása, ahogy fentebb említettük, gyakran dimenziócsökkentést igényel, s ezek a dimenziócsökkentési eljárások gyakran felügyelet nélküliek. Így egy felügyelt tanulási regressziós modellben magyarázó változóként használhatjuk a klaszterelemzés által adott besorolást. De ugyanígy gyakori megoldás az, amikor felügyelt modellben a szöveget alkotó szavak helyett a szóbeágyazási modell által adott szójelentéseket használják magyarázó változóként.

A felügyelt gépi tanulás inputja: humán annotálás

Két karakteresen különböző megoldás létezik a tanuló-halmaz létrehozására. Az egyikre példa a fent már idézett, Poletti és társai (2017) által publikált kutatás: ők kódolókat (szövegbányászati terminussal: annotátorokat) tanítottak be részletes irányelvek alkalmazásával arra, mikor minősítsenek egy szöveget agresszívnek vagy támadónak, mikor sztereotipizál és mikor irányul egy kisebbségi csoport ellen a szö-

⁵ Erről a magyarázati igényről és általában a magyarázatra/predikcióra való törekvés kettőségéről lásd: Németh 2021: 110–121.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

veg. Itt tehát humán annotátorok vannak, akik olvasnak és interpretálnak, adott annotálási irányelveket követve.

A másik lehetőségre Jelveh, Kogut és Naidu (2014) írása példa, akik amerikai közgazdászok ideológiai pozíciójának (jobboldali/baloldali) gépi tanulását végezték el a szerzők tudományos írásai alapján. A tanuló-halmaz azokból a közgazdászokból állt, akiknek pozíciója megállapítható volt politikai kampány-támogatásokat, illetve petíció aláírásokat tartalmazó nyilvántartások alapján. Itt tehát nincsenek humán annotátorok, nem olvasunk és nem interpretálunk, mert kész címkéink vannak.

A gyakorlatban leggyakrabban egy kutató (vagy egy kutatócsoport) kézzel végzi a kódolást, a szöveg annotálását, akárcsak a „klasszikus” kvalitatív szövegelemzés esetében. Az annotálás minőségének jelentősége kiemelkedő: a felügyelt tanuló algoritmus jó minőségben annotált tanuló-halmazból tud hatékonyan tanulni. Az annotálás persze időigényes és gyakran nem is egyszerű feladat. Az elsődleges cél a replikálhatóság, ami azt jelenti, hogy egy másik annotátor nagyon hasonló annotációkat készítené. Hovy és Lavid (2010) a következő strukturált eljárást javasolja az annotációk előállítására:

1. Határozzuk meg, hogy milyen kategóriákba sorolva kell annotálni. Ez általában valamilyen elmélet alapján történik (lásd: milyen jegyekkel definiálható a gyűlöletbeszéd). Itt megfelelő egyensúlyt kell találni a részletezettség/pontoság és a skálázhatóság/idővonzat között.

2. Választhatunk olyan platformot, amely támogatja az annotálás adminisztrálását. Több általános célú annotációs eszköz érhető el.

3. Az annotálási feladatra vonatkozó utasítások formalizálása annotálási irányelvek formájában. Amennyiben az utasítások nem explicitek, a kapott annotációk szubjektív benyomásokon alapulnak majd, ami a replikálhatóságot veszélyezteti.

4. Az adatok egy kis részhalmazának kísérleti annotálása (pilot), több annotátorral. A pilot előzetes benyomást ad mind a megismételhetőségről, mind az annotálási irányelvek alkalmazhatóságáról. A megismételhetőséget az annotátorok döntései közötti egyezés (inter-annotator agreement) mutatóival jellemezhetjük. A konkrét annotálási eltérések vizsgálata segíthet az utasítások pontosításában, és az annotálási feladat módosításához is vezethetnek.

5. Fő annotálás. Érdemes legalább az adatbázis egy részét párhuzamosan annotálni, azaz két vagy több annotátorral egymástól függetlenül besorolítani, hogy az annotátorok közötti egyezés kiszámítható legyen. Sok projektben a szövegek több címkét is kapnak, amelyek aztán összesítve egy „konszenzusos” címkévé állnak össze.⁶

⁶ Példaként lásd: Danescu – Niculescu – Mizil et al. 2013.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

6. A felügyelt tanulás középpontjában az annotátorok közötti egyetértés áll: ha a kódok nem megbízhatóak, a tanuló algoritmus nem tud hatékonyan tanulni belőlük, és besorolásai sem lesznek megbízhatóak. Ezért elengedhetetlen az annotálás értékeléseként az annotátorok közötti egyezés mutatójának kiszámítása. Ha ez a mutató alacsony, az az annotátorok megbízhatóságát vagy magát a teljes annotációs rendszert (a besorolás értelmességét) kérdőjelezi meg. A mutató definíciójára több matematikai megoldás létezik a konkrét feladat függvényében. Az annotátorok besorolásainak egyszerű százalékos egyezése, annak ellenére, hogy nagyon széles körben használják, nem veszi figyelembe a véletlenszerűen előforduló egyezést. Ugyanis, ha két kommentátor véletlenszerűen választ két címke között, akkor a köztük lévő egyetértés várhatóan 50%-os lesz. Ezért egy jó mutató a nyers egyezési százalékot a véletlen egyezés arányához viszonyítja. Egy ilyen széles körben használt mutató a Cohen-féle kappa is.

Érdeemes megjegyeznünk, hogy ha az általunk vizsgált eset nem túl specifikus, nem feltétlenül kell saját annotálást végrehajtanunk, használhatunk mások által annotált adatbázisokat is. Számptalan annotált, nyílt elérésű korpusz található az interneten, melyek elsősorban nyelvészeti feladatokra alkalmazhatók, de találhatunk Twitter bot-detektálásra szolgáló címkézett adatbázisokat is, sőt társadalomkutatási célúakat is, mint a Manifesto projekt, amely ötven ország politikai pártjainak választási programjainak annotált adatbázisát nyújtja 1945-től napjainkig.

Egy saját kutatási példa

Sik Domonkossal és Máté Fannival végzett kutatásunkban (Németh – Sik – Máté 2020) mi is a fentihez hasonló eljárást dolgoztunk ki. Kutatásunk célja az volt, hogy különböző felügyelt tanulói algoritmusok alkalmazásával automatikusan osztályozzuk ismert nemzetközi online depresszió-fórumok bejegyzéseit aszerint, hogy abban a depresszió milyen (bio-medikális, pszichológiai vagy társadalmi) keretezését adja a felhasználó.

Az annotátorokat a téma és a módszer iránt érdeklődő társadalomtudományi szakos hallgatók közül választottuk ki. A teljes adatbázisból, ami 70 000 posztból állt, egyszerű véletlen mintavétellel 4500-at választottunk ki felcímkézendő tanulólthalmazként. Több tréninget tartottunk az annotátoroknak, és sok valós példát felvonultató, részletes annotálási irányelveket készítettünk, amelyet a pilot szakasz után is, és a fő annotálási szakaszban is folyamatosan frissítettünk. Öt címkét használtunk, mivel a három keretezés-típus mellé a „besorolhatatlan” (depresszióról van szó, de a keretezés nem azonosítható) és az „irreleváns” (nem depresszióról van szó) is hozzáadódott. Az annotálási feladat a legkevésbé sem volt triviális, ezért (1) az annotátorok szükség esetén két címkét rendelhettek a szövegekhez, egy elsődleges és egy (opcionális) másodlagos címkét. A bejegyzések 34%-a kapott második címkét



TEMATIKUS TANULMÁNYOK – Digitális szociológia

legalább az egyik annotátortól. Továbbá (2) minden szöveghez két független annotátorunk volt, akik együttesen legfeljebb négy címkét adtak. A végső, konszenzusos címke „többségi szavazáson” alapult, tehát a négy címke közül a leggyakoribbat választottuk. A (nagyon kevés, 12,3%-os) kétértelmű esetek feloldására egy harmadik annotátort (egy vezető kutatót) kértünk fel. Másodlagos konszenzusos címkét is kiosztottunk, ha a „szavazásnak” volt egy értelmű második győztese.

Az annotátorok közötti egyezés mutatójának meghatározására a Cohen-féle kappát használtuk, amely azt mutatja meg, hogy az annotátorok mennyivel jobban teljesítenek a véletlenszerűen besoroló annotátorokhoz képest. Tökéletesen egyező besorolásoknál az értéke 1, ezzel szemben, ha az annotátorok véletlenszerűen választják ki a címkéket, akkor a kappá egyenlő 0-val.

Annotációink másik sajátossága az volt, hogy az annotátorok másodlagos címkéket rendelhettek a bejegyzésekhez. Ha az egyezést az elsődleges címkék egyezéseként határozzuk meg, egyszerűen elvetve a választható másodlagos címkéket, akkor egy túlságosan konzervatív mérőszámot kapunk. Ezért a kappá „liberális” változatát használtuk, az egyezést úgy definiálva, hogy az egyik elsődleges címke megegyezik a másik annotátor által adott elsődleges vagy másodlagos címkével. A liberális kappá előnye az eredeti, konzervatív változatával szemben az, hogy figyelembe veszi a szövegek másodlagos jelentését is. Míg a liberális kappá túl optimista képet mutathat, addig konzervatív megfelelője túl szigorú értékeléshez vezet. Az „igazság” valahol a kettő között van, ezért mindkettőt bemutattuk. Konzervatív módon mérve az annotátoraink közötti egyezés 58,3%, a liberális mérőszám pedig 69,7% volt, ami elfogadható mértékű egyezést mutat.

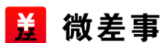
Crowdsourcing annotálás

A tanulás sikerét közvetlenül befolyásolja a tanuló-halmaz mérete – hasonlóan ahhoz, ahogyan survey-ek esetén a mintanagyság a megbízhatóság megfontosabb faktora. Ezért nem ritkák a kifejezetten nagy (több tíz- vagy százezres) elemszámú annotálandó adatbázisok. Ezekben az esetekben a crowdsourcing platformokon bérelhető bedolgozó annotátorok jelentenek megoldást, például az Amazon által működtetett Mechanical Turk, a Figure Eight,⁷ Lighttag vagy a kínai Weichaishi (1. ábra).

⁷ Korábbi elnevezése CrowdFlower.



TEMATIKUS TANULMÁNYOK – Digitális szociológia



THE PIONEER OF CROWDSOURCING IN CHINA

中文 | English

Home About Us

I'm A User

About Us

Home > About Us

About Us

Contact Us

Launched in Jan 2013, WeiChaiShi (WCS, 微差事 - Chinese word for 'Micro-Task') is the biggest crowd-tasking B2C platform in china. WCS is a mobile APP that instantly connects businesses to an on-demand mobile workforce who are incentivized to collect, capture, and report real-time data for brand clients. As of today, 3 million smart-phone registered users have completed more than 7.5 million simple tasks, acting as a go-to team to effectively connect consumers, stores, and brands.

The type of tasks includes commercial inspection, data collection, research & census and experiential marketing, etc. Through this new business model, survey companies significantly improve sampling coverage and efficiency, brands are able to create followers while lowering operations and marketing cost. In 2014, WCS won the award of 'Best business Model in China' judged jointly by '21st Century Business Review' and '21st Century Business Herald'.

WCS APP, It's EASY, It's FUN, and It PAYS!

1. ábra. A WeiChaiShi kínai crowdsourcing cég beköszönő weboldala. A bedolgozók mobiltelefonján keresztül adatgyűjtést, survey-eket és kísérleteket is kínálnak

Forrás: <http://www.weichaiishi.com/>

A crowdsourcing platformon annotáltatni kívánó kutatók egyszerűen felhívást tesznek közzé a platform bedolgozói között, melyben az annotáció darabjára kívülről az annotátorok elvárt képességeit (minimális képzettségét, anyanyelvét, korábbi munkáikkal kapcsolatos elégedettségi arányt stb.) adják meg (2. ábra). A viszonylag képzetlen „tömegmunkások” alkalmazása látszólag ellentétben áll a replikálhatóságra vonatkozó elvárásokkal, azonban számos, a platformot használó kutatás számol be megbízható annotációkról egyszerűbb feladatok kapcsán (Snow et al. 2008).

A crowdsourcing egy tágabb terület, az ember-alapú számítás (human-based computing, lásd még: elosztott gondolkodás, distributed thinking) részét képezi. Ez olyan számítástechnikai megoldásokat foglal magába, amelyekben egy számítógép úgy látja el a funkcióját, hogy bizonyos lépéseket embereknek szervez ki, egy szimbiózisszerű interakcióban (Mühlhoff 2019). Megfordulnak a szerepek: a gép kér fel embereket egy probléma megoldására, majd integrálja a megoldásokat. Az olyan számításigényes feladatok esetében, mint a képfelismerés, a humán annotáció fontos részét adja a mélytanuló algoritmusok tanításának.

A crowdsourcinggal kapcsolatban komoly etikai kérdések merülnek fel. Fort és Cohen (2011) már tíz éve arra figyelmeztetett, hogy az Amazon Mechanical Turk többszázézes bedolgozói tömeget működtet online, szerte a világban, jó részük él Indiában és Törökországban. A jellemző órabérek 2 dollár alatt vannak, és az általános elképzeléssel szemben a felhasználók jellemzően nem kismamák vagy diákok, aki hobbiként dolgoznak itt, hanem olyan munkavállalók, akik a megélhetéshez szükséges forrásként tekintenek munkájukra. Az alacsony bérrel kívül további ko-



TEMATIKUS TANULMÁNYOK – Digitális szociológia

The screenshot shows the Amazon Mechanical Turk worker interface. At the top, there's a navigation bar with 'amazonmturk Worker', 'HITS', 'Dashboard', and 'Qualifications'. A search bar for 'All HITS' is present. Below this, a table lists 'HIT Groups (1-20 of 2106)'. The table has columns for Requester, Title, HITS, Reward, Created, and Actions. A tooltip is visible over the second row, indicating a language requirement.

Requester	Title	HITS	Reward	Created	Actions
Amazon Requester Inc. - C	[French language proficiency requir...	61,046	\$0.01	17h ago	Preview Accept & Work
Amazon Requester Inc. - C	[日本語能力が必 (French language proficiency required) Questionnaire sur la relativité des produits aux intérêts (répondre par oui ou non)		\$0.01	7h ago	Preview Accept & Work
Amazon Requester Inc. - C	Product to Interest Audit (single yes/...	28,379	\$0.01	1h ago	Preview Accept & Work
Amazon Requester Inc. - C	[dominio del idioma español requeri...	27,670	\$0.01	21h ago	Preview Accept & Work
Amazon Requester Inc. - C	[Proficiência no idioma português br...	19,719	\$0.01	20h ago	Preview Accept & Work
Crowdsurf Support	Transcribe up to 35 Seconds of Med...	17,485	\$0.05	3m ago	Preview Qualify
TC Research	Find the Email for These Mental He...	13,896	\$0.12	5d ago	Preview Accept & Work
UnSpun Opinions	Opinion Survey	12,180	\$0.50	1m ago	Preview Accept & Work
KronoPin	Find the Website Address for a Con...	11,846	\$0.03	2/23/2018	Preview Qualify
Assistive Technology Rese	1 minute survey: Smart speakers at ...	10,577	\$0.15	3d ago	Preview Accept & Work
Armin Hamzic	Tell us if a picture shows a specific f...	10,557	\$0.01	1d ago	Preview Qualify
nttkKAN	Image Annotation (WARNING: This ...	8,217	\$0.05	10d ago	Preview Accept & Work

2. ábra. Az Amazon Mechanical Turk egy oldala a felkínált munkákkal, a munkaadók (requesters) által fix darabbért kínált un. HIT-ekre (Human Intelligence Tasks), elemi munkaegységekre osztott feladataival

Forrás: blog.mturk.com, <https://blog.mturk.com/quick-update-another-improvement-to-the-mturk-worker-experience-9cfd0b1963e7>

moly problémát jelent, hogy a munkavállalók (azaz „turkers”) híján vannak az olyan alapvető munkahelyi jogoknak, mint a kollektív alku, a szakszervezet alakításának lehetősége, a munkáltatói jogsérelmek orvoslásának lehetősége, ezért rendkívül kiszolgáltatottak.

Egy friss munka (Gray – Suri 2019) gazdaságtantropológiai megközelítésben vizsgálja ezeknek a láthatatlanul dolgozó „szellemmunkásoknak” a körülményeit. A szerzők szerint az átláthatóság hiánya azt a benyomást kelti a közvéleményben, hogy a mesterséges intelligencia egyedül működteti a kortárs kényelmi szolgáltatásokat, miközben számos ponton elengedhetetlen az emberi beavatkozás: a szellemmunkások az annotáláson túl számos támogató munkát (személyazonosítás, feliratozás



TEMATIKUS TANULMÁNYOK – Digitális szociológia

stb.) végeznek olyan óriások számára, mint az Amazon, a Google, az Uber és a Microsoft. A munkával járó anonimitás és rugalmasság kétségtelenül sokuknak előnyös, pl. az egyébként gyakran diszkriminált munkavállalói csoportoknak (nők, fogyatékkal élők). Ugyanakkor Fort és Cohen (2011) korábbi megfigyelései még mindig érvényesek, azzal a különbséggel, hogy ma már dolgozók millióira vonatkoznak: a szellemmunkások rendkívül alulfizetettek, kiszolgáltatottak, munkavállalói jogaik erősen sérülnek. A szerzők javaslatainak (munkavállalói juttatások bevezetése, szakszervezet alakítása, a munkavállaló és munkáltató közötti emberi kommunikációt lehetővé tevő új központok létrehozása) tépje nagy: a méltányos emberi munka jövőjét biztosítanák.

A humán annotálás kihívásai szociológiai alkalmazásokban

Az ipari/üzleti alkalmazásokban már sokszorosan bizonyított felügyelt gépi tanulás szociológiai alkalmazásai sajátos kérdéseket vetnek fel. A sajátosság oka, hogy ezekben az alkalmazásokban komplex fogalmak megtanulása az algoritmus feladata (lásd: gyűlöletbeszédet tartalmaz-e egy tweet), szemben az olyan könnyebben annotálható feladattal, hogy negatív vagy pozitív-e egy szolgáltatással kapcsolatos megjegyzés.

Kutatásunk (Németh – Sik – Máté – 2020) már idézett esete jól példázza ezt: a depresszió keretezésének eldöntése nem bizonyult egyszerű feladatnak. A pilot során, az annotálási irányelvekben felsorolt elveken alapuló értelmezést több egyéni és csoportos fordulóban gyakorolták az annotátorok. Mégis, a pilot szakasz többszöri meghosszabbítása után is elégtelen maradt az annotátorok közötti egyetértés (az elsődleges címkék százalékos egyezése 60% alatt volt). Ezen a ponton vált időszerűvé saját módszertani háttérfeltevéseink felülvizsgálata. Rá kellett jönnünk, hogy az általános kategóriák (keretezés típusok) és a konkrét hozzászólások ilyen jellegű társítása nem egyértelmű feladat. Gadamer hermeneutikai elméletét (2004) követve innét kezdve a jelentések kialakulását nyelvileg közvetített, interszubjektív, iteratív folyamatnak tekintettük, amelyben a jelentések ténylegesen egy folyamat során konstruálódnak. Módszertanilag ezt az interszubjektív folyamat az annotálási irányelvek folyamatos, iteratív frissítésében jelent meg. Ez a folyamat lényegesen különbözik azoktól a hermeneutikailag egyszerűbb üzleti alkalmazásoktól, amelyek explicit és egyértelmű kategóriák (lásd pozitív/negatív/neutrális szentiment) előre meghatározott készletét alkalmazzák. A mi eljárásunk inkább egyfajta kvalitatív kódolásként határozható meg, mivel a kategóriáinkat egy előzetes absztrakt elméletből származtattuk, és induktív módon alakítottuk ki őket a kutatás során.

Az interszubjektivitás elismerésének másik megnyilvánulása a kettős annotációra való áttérés volt. Ahelyett, hogy azt feltételeztük volna, hogy minden egyes hozzászólás egy vagy két kategóriába tartozik, amelyeket egy megfelelően képzett



TEMATIKUS TANULMÁNYOK – Digitális szociológia

annotátor azonosítani tud, úgy közelítettük meg a fórumbejegyzéseket, mint amelyek többféleképpen értelmezhetők. Annak érdekében, hogy minimalizáljuk az értelmezés esetlegességét, két független annotátorral kódoltattunk minden posztot. A végső, konszenzusos címke a két annotátor kódjának egyesítésén alapult a fent leírt módon.

Tapasztalataink szépen visszatükrözik más kutatók benyomásait. Aroyo és Welty (2015) egyenesen a humán annotálás hét mítoszaként utal azokra a naiv/pozitivisták elképzelésekre, melyek egyértelműen kódolható szövegeket tételeznek fel. Szerintük a humán annotálás egy elavult szemantikai eszményen alapul, ami az egyetlen helyes igazság meglétét tételezi fel, ebből további mítoszokat levezetve, mint például az annotátorok közötti eltérés hátrányos voltát, azt a reményt, hogy a megfelelően részletezett annotálási irányelvek megoldják a problémát, annak hitét, hogy a szakértők besorolása mindig helyesebb, mint a laikusoké, vagy annak elvárását, hogy egyetlen kategóriába besorolható legyen a szöveg. Új szemantikai elméletet javasolnak, a „crowd truth”-ra alapozva, melynek lényege, hogy az emberi értelmezés szubjektív, és hogy az annotátorok különböző interpretációi jó reprezentációját adják ennek a szubjektívitásnak, s az ésszerű interpretációk tartományának.

A Mesterséges Intelligencia torzítás és az annotálás

A Mesterséges Intelligencia torzítás (MI-torzítás, angolul AI-bias, az Artificial Intelligence rövidítéseként) lényege, hogy a nyelvtechnológia outputja maga is társadalmi torzításokat (pl. kisebbségekkel, idősekkel vagy nőkkel szembeni hátrányos megkülönböztetést) mutat, amellyel mintegy felerősíti azok társadalmi hatását (Ntoutsis et al. 2020). Az egyik legismertebb példa az Amazon kísérleti rekrutációs algoritmusának esete (Dastin 2018). Az algoritmus létrehozásának célja az volt, hogy az 1-től 5-ig automatikusan osztályozza önéletrajzok alapján a jelentkezőket. 2015-re azonban kiderült, az algoritmus gender-alapon torzított a szoftverfejlesztő és más technikai jellegű pozíciók esetén. Ennek oka az volt, hogy a modellt a céghez a megelőző 10 évben benyújtott jelentkezések alapján tanították, de a technológiai iparág férfi-domináns, így a jelentkezések is férfiakról jöttek elsősorban. Emiatt az MI rosszabbul minősítette azokat a jelentkezőket, amelyekben a „női” szó szerepelt, legyen az akár csak a „női sakkcsapat kapitány” kifejezés. A fejlesztők megkísérelték oly módon javítani, hogy az ilyen fajta nemi információtól való függést kiiktatták a tanulóból, de ez nyilván nem feltétlenül jelenti a probléma megoldását, hiszen egy ilyen rendszer más, nem vizsgált dimenziók mentén is tartalmazhat torzítást. A fejlesztők végül felhagytak ennek az algoritmusnak a használatával.

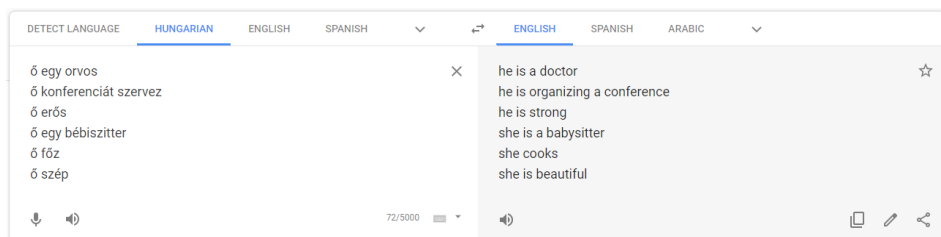
Az MI-torzítás általánosan megfogalmazva egyfajta anomália a gépi tanuló algoritmus outputjában; vagy az algoritmus fejlesztési folyamatában alkalmazott felte-



TEMATIKUS TANULMÁNYOK – Digitális szociológia

vésekre vagy a tanuló-halmazban rejlő inherens társadalmi torzításokra vezethető vissza. Az annotálás kapcsán a továbbiakban az utóbbira koncentrálnunk.

A tanuló halmaz, ember alkotta élő szöveg lévén mindazokat a viszonyokat tükrözi vissza, amik a társadalomban is megtalálhatók. Az Amazon példáján: a tanuló-halmaz kiegyensúlyozatlan volt nemi összetételét tekintve. Hasonló ismert példa a Google Translate esete is. Farkas Annával (Farkas – Németh 2022) végzett kutatásunkban azt vizsgáltuk, hogy ha magyarról angolra fordítunk foglalkozásneveket tartalmazó mondatokat („Ő egy orvos”), akkor a fordító hímnemű vagy nőnemű névmást használ, s hogy ez a döntés korrelál-e a foglalkozások tényleges nemi megoszlásával illetve a magyarok survey-jel mért attitűdjével (azaz azzal, hogy inkább férfias vagy nőies foglalkozásnak vélnék valamit a magyarok). Eredményünk szerint a fordító erősen torzít a nőkkel szemben, és működése közelebb áll az attitűdökhöz, mint a tényleges foglalkozásszerkezethez (3. ábra).



3. ábra. A Google Translate nemi torzítása
Forrás: Farkas – Németh 2022: 2

A fordító torzításának oka itt is az, hogy kétnyelvű szöveghalmazokon tanul, s ha a társadalmi attitűd vagy a tényleges foglalkozásszerkezet inkább férfiasnak mutat egy foglalkozást, akkor az a szövegekben is megmutatkozik majd, így a fordító is visszatükrözi ezt az egyenlőtlenséget. (A Google mai működésében már nem detekálható ez az anomália, a visszajelzések nyomán úgy módosították az algoritmust, hogy mindkét névmást megadja a fordítás.)

Az annotáción alapuló torzítás a tanuló-halmaz szintjén vihet anomáliát a rendszerbe, kétfajta úton. Vagy az annotátorok, maguk is társadalmi normákat képviselve, tükrözik vissza a címkézésben a társadalmi viszonyokat, vagy azzal okoznak torzítást, hogy nem reprezentálják kellőképpen a társadalom egészét. Mindkét probléma egyszerre volt jelen abban a gyűlöletbeszéd-kutatásban (Sap et al. 2019), ahol a torzítást az okozta, hogy a jobbára fehér bőrű annotátorok a címkézendő tweetek szerzői által használt afroamerikai dialektust hajlamosabbak voltak offenzívként megjelölni. E dialektus ugyanis több káromkodást tartalmaz, és még ha az nem is irányul más felhasználók ellen, akkor is sértőnek, offenzívnek érezték a más nyelvi normákat képviselő kódolók.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Eltávolítható-e a torzítás az algoritmusokból? Ahogy láttuk, az MI algoritmusok épp annyira lehetnek kiegyensúlyozottak bármilyen szempontból, mint amennyire a tanuló-halmaz az. „Reprezentatív” annotátor-csapat és „reprezentatív” tanuló minta kialakítására érdemes törekedni ugyan, de tökéletes reprezentativitás még elvileg is nehezen elképzelhető. Felmerül annak kérdése is, hogy valóban a nagy átlagot kell-e az algoritmusnak hoznia, nem lehet-e egy-egy kisebbségi csoport véleménye relevánsabb – gondoljunk itt akár egy tényellenőrzési feladatra, ahol a „tények” detektálása nem feltétlenül reprezentatív szavazással dönthető el.

Megkísérrelhetjük továbbá nem az input, hanem az output oldalon is a javítást: a torzítás eltávolítását egy-egy dimenzió mentén (ahogy az Amazon fejlesztői tették), de számtalan látens dimenzió létezhet, ezért tökéletesen torzítatlan MI algoritmus nehezen elképzelhető – ugyanakkor törekedni kell tesztelésükre és javításukra.

Összegzés

A hatalmas mennyiségű digitális szöveges adat elérhetősége és új elemzési potenciálja széles perspektívát jelent a szociológia számára. A számítógépes társadalomtudomány, jelen példákön a szövegbányászat várhatóan akkor lesz beépíthető a mindennapi kutatásba, ha az interdiszciplináris együttműködések széles körben elterjednek, ha a szükséges tudás és kompetencia beépül az egyetemi képzésbe. A fejlett programozási ismereteket nem igénylő, alacsony küszöbű szövegelemző platformok ugyanakkor már most megjelentek (mint például a Google Trends, Google Ngram Viewer, vagy az európai fejlesztésű Sketch Engine), támogatva az átalakulást – ezek további fejlődése valószínűsíthető a közeljövőben.

Láttuk, hogy a szövegeken alapuló gépi tanulás is új lehetőségeket kínál a szociológia számára. A gépi tanulási megoldások egyre nagyobb része humán annotáción alapul, és az ilyen ember-gép együttműködésen alapuló számítások még erősebb felfutása várható nem csak az iparban, de a tudományban is. Az előre annotált és nyilvánosan elérhető adatbázisok már most támogatják a felügyelt tanulás saját alkalmazását. Ugyanakkor, mivel a szociológiai fogalmak általában komplexebbek, e humán annotáció során nem kerülhető meg az „igazság” egyértelműségének elvetése és az emberi értelmezés szubjektivitásának elfogadása.

A szociológus ugyanakkor nem csak felhasználóként, hanem a kritikai nézőpont képviselőjeként, a felügyelt gépi tanulás társadalmi hatásaira, etikai problémáira rámutató aktorként is jelen kell, hogy legyen. A crowdsourcing annotálás komoly munkaerőpiaci problémákat generál, az MI-torzítás pedig társadalmi hátrányokat erősíthet fel. Ezen túl a cikk keretein túlmutató, de zárásként mindenképp megemlítenő kérdés az internetes privacy és szabadság problémája. Kérdés, milyen hatása lesz ezekre a nagytömegű szövegek gyors feldolgozására képes technológia, nem válik-e a szövegbányászati technológia a cenzúra és megfigyelés mindenható eszközévé?



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Irodalom

- Aroyo, L. – Welty, C. (2015): Truth is a lie: Crowd truth and the seven myths of human annotation. *AI Magazine* 36(1): 15-24. <https://doi.org/10.1609/aimag.v36i1.2564>
- Dastin, J. (2018): Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> (utolsó elérés: 2021. április 7.)
- Eisenstein, J. (2019): *Introduction to Natural Language Processing*, MIT Press, Cambridge, MA
- Farkas A. – Németh R. (2022): How to measure gender bias in machine translation: Real-world oriented machine translators, multiple reference points. *Social Sciences & Humanities Open*, 5(1): 100239.
- Fort, K. – Adda, G. – Cohen, K. B. (2011). Amazon mechanical turk: Gold mine or coal mine? *Computational Linguistics* 37(2):413–420. doi: 10.1162/COLI_a_00057
- Gadamer, H. G. (2004): *Truth and Method*. 2. kiadás. Crossroad, New York
- Gray, M. L. – Suri, S. (2019): *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Houghton Mifflin Harcourt, Boston
- Hovy, E. – Lavid J. (2010). Towards a 'science' of corpus annotation: a new methodological challenge for corpus linguistics. *International Journal of Translation*, 22(1): 13–36.
- Ignatow, G. – Mihalcea, R. F. (2017): *An Introduction to Text Mining: Research Design, Data Collection, and Analysis* (1st edition.). SAGE Publications, Inc., Los Angeles
- Jelveh, Z. – Kogut, B. – Naidu, S. (2014): Detecting Latent Ideology in Expert Text: Evidence From Academic Papers in Economics. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1804–1809). Doha, Qatar: Association for Computational Linguistics. <http://www.aclweb.org/anthology/D14-1191>. (utolsó elérés: 2021. 11. 20.)
- Nahili, W. – Rezeg, K. – Kazar, O (2020): Big Data Analytics using Supervised Learning: A Comprehensive Review of Recent Techniques. *International Journal for Research in Applied Science and Engineering Technology*, 8(1): 305-312.
- Mühlhoff, R. (2019): Human-aided artificial intelligence: Or, how to run large computations in human brains? Toward a media sociology of machine learning. *New Media & Society*, 22(10): 1868-1884. <https://doi.org/10.1177/1461444819885334>
- Németh R. – Koltai J. (2021): Sociological knowledge discovery through text analytics. In: Rudas T. – Péli G. (szerk.): *Pathways between Social Science and Computational Social Science – Theories, Methods and Interpretations*. Springer, New York



TEMATIKUS TANULMÁNYOK – Digitális szociológia

- Németh R. – Katona E. – Kmetty Z. (2020): Az automatizált szöveganalitika perspektívája a társadalomtudományokban. *Szociológiai Szemle*, 30(1):44-62.
- Németh R. (2021): *Az okság alternatív fogalmi és módszertani megközelítései a szociológiában*. Savaria University Press, Szombathely
- Németh R. – Sik D. – Máté F. (2020): Machine Learning of Concepts Hard Even for Humans: The Case of Online Depression Forums. *International Journal of Qualitative Methods* 19(1): 1-8. <https://doi.org/10.1177/1609406920949338>
- Ntoutsis, E. et al. (2020): Bias in data-driven artificial intelligence systems – An introductory survey. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery* 10(6): 1-14. <https://doi.org/10.1002/widm.1356>
- Poletto, F. – Stranisci, M. – Sanguinetti, M. – Patti, V. – Bosco, C. (2017): Hate Speech Annotation: Analysis of an Italian Twitter Corpus In: Basili, R. – Nissim, M. – Satta, G. (eds.): *Proceedings of the Fourth Italian Conference on Computational Linguistics CLiC-it 2017: 11-12 December 2017, Rome*. Accademia University Press, Torino <https://doi.org/10.4000/books.aaccademia.2448>.
- Ryan, L. – McKie, L. (eds.) (2015). *An end to the crisis of empirical sociology? Trends and challenges in social research*. Routledge, London
- Sap, M. – Card, D. – Gabriel, S. – Choi, Y. – Smith, N. A. (2019): The Risk of Racial Bias in Hate Speech Detection. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Florence, Italy.
- Savage, M. – Burrows, R. (2007): The Coming Crisis of Empirical Sociology. *Sociology: A Journal of the British Sociological Association*, 41(5):885–899. <https://doi.org/10.1177/0038038507080443>
- Snow, R. – O'Connor B. – Jurafsky D. – Ng A. Y. (2008): Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In: *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Honolulu, Hawaii
- Vapnik, V. N. (2000): *The Nature of Statistical Learning Theory*. 2. kiadás, Springer-Verlag, New York, Inc.