

A Twitter átfogó vizsgálata

A cikk szerzői a legátfogóbb tudományos tanulmányt készítették el a Twitterről. Kutatásuk során több mint 1100 tudományos tanulmány kivonatát vizsgálták meg 2007 és 2011 között, amelyek a Twitterre fókuszáltak. A tudományos tanulmányok kivonatainak forrása a *Scopus*, a *Web of Science*, valamint a *Google Scholar* volt.

A szerzők bemutatták a mikroblogolás kialakulását, amit a könnyű hozzáférhetőség és az olcsó mobil technológia elterjedése nagymértékben elősegített. A mikroblogolás a világhálón a 2000-es évek elején jött létre, amely lehetővé tette, hogy a használók azonnali válaszadásra legyenek képesek, kommentálásokat fogalmazzanak meg és esetlegesen felerősítsék az aktuális események hatását. Tudományos kutatásuk csak ezt követően indult el. A szerzők megítélése alapján a Twitter 2008-ra a legelterjedtebb felületté vált és azóta is vezető mikroblog. A felhasználók a gyorsabb kommunikációt 140 karakterlimiten belül valósíthatják meg. Megítélésük szerint a Twitter mára médiummá vált, mert az aktuális híreken, a politikai eseményeken át a katasztrófákról egyaránt információt biztosít használói számára, viszont ez a nyitottság véleményük szerint visszaélésekre és a lehetőségek rosszra történő felhasználására is módot adhat, ami etikai kérdéseket vethet fel.

A szerzők figyelembe vették a téma irodalmi áttekintését, amely magába foglalja a mikroblog és a Twitter szavak előfordulását a tudományos adatbázisok cikkeinek címében, kivonataiban és kulcsszavaiban, a mikroblogra megalkotott definíciókat, valamint a már korábban kialakított osztályozásokat, amelyek esetében az első csoport a felhasználó tartománya, míg a második csoport az üzenet tartománya volt. Megállapították, hogy a korai kutatások egyrészt a Twitterekben megjelenő üzenetek számát, a követőket és követések számát, a posztolások számát és posztok helyét, másrészt a Twitteren megjelenő témaköröket vizsgálták: mint például a társadalmi egészség, politika és kor-

mányzat, üzlet, oktatás és tanulás, újságírás, szemtanúi beszámolók.

A szerzők elkészítették saját osztályozásukat is. Analizálták a kivonatokat és kialakítottak egy keresési technikát a Twitterrel kapcsolatos tanulmányok beazonosítására és utána egy megalapozott kutatói megközelítést használtak az osztályozás és a prezentálás kivitelezéséhez.

A szerzők a témával kapcsolatos irodalom áttekintésében a következő négy szempontot vették alapul:

- az üzenet: a szöveg, amit a felhasználó beír vagy elküld, és a kapcsolódó metaadat, amikor elküldésre kerül az üzenet;
- a felhasználó/használó: a mikroblogolási rendszerben megjelenő digitális személyiség kialakítása, ami tartalmazhat információt arról, hogy a felhasználó kit követ;
- a technológia: az üzenet küldéséhez szükséges hardverek és szoftverek;
- a koncepció: felöleli a bevezetők áttekintését, értekezésdarabokat egészen beszámolóig.

Foglalkoztak továbbá a tartományok (domén) és az adatok áttekintésével is, azok méretével, mélységével és a tanulmányok hosszúságával. Tanulmányozták a tanulmányok osztályozáshoz kapcsolódó módszereket is. A kutatást minőségileg osztályozták a nyitott kódú elemzés használatával, alapul véve a tanulmány címét és kivonatát a Miller-féle technikával. Nyílt kódú elemzést választottak, mivel ez hozzájárult a koncepció körvonalazásához, ahol a sorról sorra történő kódolási folyamat megjelöli az adatban fellelhető változásokat.

Megállapították, hogy 1161 értekezésből 575 fókuszált a Twitterre és a mikroblogolásra, 550 tartalmazott említést a témával kapcsolatban. A maradék 27 tanulmányban utalás van a Twitter kifejezésre, de nem állnak összefüggésben a mikroblogolással, más témákkal viszont igen, mint például az állatok hangjai.

A tanulmányok teljes listáját a CentAUR-ban sorolták fel, külön kategóriákban az alábbiak szerint:

- azok a tanulmányok, amelyek a Twitterre fókuszálnak;
- amelyek említik a Twittert, de nem fókuszálnak rá;
- amelyek használják a Twitter szót, de nem állnak kapcsolatban a mikroblogolással.

Az elemzés módszerei

A kivonatokból 33 kutatási módszer volt eredetileg azonosítva, mint amelyet felhasználtak a publikált kutatásban. Az információk forrásaként szolgáló módszertanulmányok jelezték, hogy szükséges lenne a használt módszereket valahogy osztályozni. A kivonatok felülvizsgálata során észrevették, hogy néhány szerző sokkal több részletet adott a módszeréről, mint mások; és hogy van olyan kivonat ahol csak referálást láttak, míg máshol a kutatók végrehajtották a tartalmi és érzelmi területek elemzését is.

A szerzők *négy módszert* alakítottak ki, mely az összes megközelítési formát felölelte:

1. *Elemző*. Itt a kutatók néhány analízistípust hajtottak végre, például tartalom, adat, szemantikus, közösségi hálózati analízisek; mennyiségi és minőségi megközelítésekkel.
2. *Tervezés és fejlesztés*. Itt rendszereket alakítottak ki és ajánlottak fel, ami lehetett felfedező, amely tartalmazott kísérletet, bemutatót, modellt vagy szimulációt, teljes tervezést és kivitelezést.
3. *Vizsgálat*. Itt a szerző által elkezdett beszámoló és felmérés típusa jelenik meg, felölelve az alábbi megközelítéseket: például életrajz, esettanulmány, esszé, etnográfia, értékelésinterjú, vizsgálat és tanulmány.
4. *Tudásfelfedezés*, ahol a létező technikákat alkalmazták a mesterséges intelligenciakutatásokból, a matematika és a statisztika területéről adatgyűjtés céljából, az adat és szövegbányászat, valamint a természetes nyelv (szóhasználat) feldolgozása érdekében. Továbbá felölelték az új algoritmusok fejlődésének megközelítését.

Az 575 tanulmányból álló csoport módszertani megközelítéséről a szerzők táblázatot készítettek. Számos kivonatban, két vagy három metodikai megközelítést használtak. E módszertani megközelítések egyike sem terjed ki azonban mind a négy megközelítésre. A módszerek kombinációjával relatíve kevés esetben találkoztak, a leginkább említésre méltók a következők voltak:

- (1) A tudásfelfedezés módszereit használták:
 - 24 tanulmány esetében analitikai módszer,
 - 28 tanulmány esetében tervezés és fejlesztés,
 - 7 tanulmány esetében mindkettő: analitikai, tervezési és fejlesztési
- (2) Vizsgálati módszereket használtak
 - 15 tanulmány esetében analitikus,
 - 11 tanulmány esetében tervezés és fejlesztés.

A leggyakrabban kombinált módszertani megközelítés a tudásfelfedezés volt, ami magában foglalja a mesterséges intelligencia létező technikáit, a matematikát és a statisztikát.

Szempontok

A tanulmányokat kettőnél több aspektusból nem vizsgálták. Tisztán látszik, hogy a leginkább tanulmányozott terület az *Üzenet* volt, amely az üzenetek Twitteren történő cserélődését vizsgálta. A második leginkább kutatott terület a *Használó* volt, amely vonatkozott a használó profiljára és az azt követőkre. Mind a 146 tanulmány együttesen vette figyelembe az *Üzenetet* és a *Használót* (80 elsődlegesen az *Üzenetre* vonatkozik, 66 elsődlegesen a *Használóra* vonatkozik). Relatíve kevés dolgozat tanulmányozza a technológiát és későbbi továbbfejlesztésüket, amely talán a Twitter védett és limitált természetéből adódik. Az eredményeik összhangban vannak *Cheong* és *Lee* munkájával, amely megállapította, hogy a legtöbb Twitter alapú kutatás az üzenetre koncentrál. A cikk írói úgy tapasztalták, hogy a szerzők egyaránt foglalkoznak az üzenettel, és a felhasználóval: kombinálva azt, hogy mit mondanak az emberek és, kik ezek az emberek. A kutatás módszere ebben a tanulmányban a különböző aspektusok vizsgálatára koncentrált.

A szerzők megállapításai alapján a technikai tanulmányok legtöbbje a tervezés és a fejlődés módszertanából közelített. Megállapították, hogy a tanulmányok legtöbbje a kísérleti módszeren alapult, ezek beszámolókat tartalmaznak a rendszerről. Az üzenetközpontú tanulmányok legtöbbje pedig a tervezés és a fejlesztés megközelítést alkalmazták.

Az adatok

A legtöbb Twitter-fókuszú kivonat (több mint 80%) nem adott semmilyen mennyiségi információt az adatokról, amelyek használva voltak a tanulmányban és arról sem, hogy az hogyan lett összegyűjtve.

A tartomány

A korábbi kutatások több mint 280 kategóriát alakítottak ki a domének osztályozására. A szerzők újratezték a tartományok kategóriáinak kialakítását, és ennek alapján 13 kategóriát hoztak létre:

1. Üzlet: amely az összes kereskedelmi témát lefedi, beleértve a PR-t és marketinget.
2. Osztályozás: azon tanulmányokat öleli fel, amelyek beazonosítanak bármely mintát és csoportot, beleértve az intelligenciát is.
3. Kommunikáció: a kommunikációnak széles az áttekintése, amely felöleli az egyének közötti kommunikációt, amely befolyásolhat másokat is, a médiát is beleértve.
4. Oktatás: oktatási szöveggörnyezetben történő használata, kezdve a formális egyetemi környezettől, az általános társadalmi tudatosságig, tanulásig.
5. Vészhelyzet: amely tartalmazza a váratlan körülményeket, ami magában foglalja a katasztrófákat és a földrengéseket, valamint az árvizeket.
6. Földrajz: felöleli a helyet, a nevesített országokat, kultúrát és politikai szempontokat a használó lakhelye mentén.
7. Egészség: minden egészségügyi és orvosi ügyet.
8. Könyvtárak: beleértve az archívumokat és a tárhelyeket (raktárakat) is.
9. Nyelvészet: beleértve a mondattant, jelentés-ant és szentimentalizmust, kulturális protokollt, és a használatát a többnyelvű kommunikációkban.
10. Kutatás: tartalmazza az ajánlókat, a felismerési trendeket valamint a kézi (egyéni), vagy automatizált kutatást is.
11. Biztonság: beleértve a SPAM-eket, az automata tweetterek használatát (bot), valamint a meghatalmazásokat a megbízhatóság és a személyazonosság szempontjából.
12. Technikai: felölelt területébe beletartozik a vizualizáció használata, a hálózatok, és a T. specifikus hashtag-ekik (#).
13. Egyéb: az összes többi, ami a fenti kategóriákba nem fért bele.

Ezt követően a Twitter-fókuszú tanulmányokat újra kiosztották ezekbe a tárgykörökbe. Az 575 tanulmány tárgykörében megjelentek a hely kultúrájáról, politikájáról, sőt még a felhasználó fizikai elhelyezkedéséről szóló tanulmányok is. 91-ben a *Földrajz* volt a domináns tárgykör. A tanulmányokból 11 kapcsolódott más tárgykörökhöz is, ebből 4 *Vészhelyzet*ről szóló tanulmány volt, melyekben bizo-

nyos helyeken történt balesetekről adtak információt. Az *Egyéb* kategória az alábbi területekből tevődött össze: kisállatok és ruhák, celebek, jogi szemléletek tweettelése, ahol a kivonat túl általános volt és egyik specifikus tárgykörbe sem voltak beoszthatók. Ezek a tárgykörök párhuzamban állnak más kutatók által meghatározott tárgykörökkel.

Tárgykör, módszer, szemléletek

Külön táblázatban foglalták össze az egyes tárgykörhöz százalékosan kapcsolódó Twitter-fókuszú tanulmányokat, amelyek mindegyik fajta módszert tartalmaznak, és minden nézőpontra kiterjednek. Figyelembe vették, hogy néhány tanulmányban egynél több módszert használtak, és hogy a módszerek száma egy tárgykörben több lehet, mint 100%. A Könyvtár tárgykörében 14-ből 12 tanulmány Vizsgálati módszertant alkalmaz, míg a többi módszert alig használja. Az Egészség tárgykörben 31-ből csak 7 tanulmány adoptálta a Tervezés és Fejlődés módszerét. Hasonlóan kevés volt az Üzlet tárgykörének használata. Az Egészséget és a Nyelvészetet tárgyaló értekezésekben a módszer nagy részben az Elemző módszerre épült, kb. 31-ből 14 esetben és 44-ből 21-nél, ami mintegy 27%. Ez azt tükrözi, hogy mindkét tárgykörnél a kutatók gyakran elvállalták az adatok és a tartalom mennyiségi és minőségi elemzését is. A Könyvtár tárgykörét nézve, egy kiugrást találtak, ahol 14-ből 10 esetben a Konceptióra koncentrálnak a tanulmány. Megállapították, hogy a Vészhelyzet tárgykör az Üzenetre koncentrálnak, 23-ből 21 esetben, a 61%-os átlaghoz viszonyítva, ami talán azt tükrözi, hogy vészhelyzeti esetben a Twitter képes arra, hogy gyors információt nyújtson, míg a hagyományos hírszolgáltatások nem elég gyorsak, vagy egyáltalán nem is elérhetők.

A szerzők felhasználták a Pearson-féle korrelációt, mely két változó kapcsolatát méri. A korreláció értéke mindig +1 és -1 között mozog, ahol a pozitív érték azt mutatja, hogy az egyik változó együtt növekszik, míg a negatív érték azt jelzi, hogy ha az egyik érték növekszik, a másik csökken. Megállapították egyrészt, hogy láthatóan erős kapcsolat van a tudás-felfedezés módszere és az üzenet tanulmányozása között, aminek feltehetően az az oka, hogy a tudás-felfedezéses módszere alkalmas arra, hogy nagy mennyiségű adathalmazt kezeljen és hogy az üzenetek egy hatalmas mennyiségű információ forrásai, másrészt, hogy szintén erős a kapcsolat a felhasználó és a Tervezés és Fejlődés módszere között. Ábra mutatja be a kapcsolódó adatokat különböző formában úgy,

hogy feltérképezi a tanulmányok számát minden tárgykörben, ami használja a Tudás Kutatás módszert ellenben, azok tanulmányok számával, ahol az Üzenetre fókuszálnak elsősorban; a második sorozatban egy hasonló összehasonlítás van: megvizsgálják azon tanulmányok számát, melyben a Tervezés és Fejlődés módszerét alkalmazzák, összehasonlítva azokkal az értekezésekkel, ahol a felhasználón van a hangsúly.

A tárgykörök jellemvonásai

A Twitter fókuszú tanulmányok tárgyköreinek használták elemzésénél a szerzők a TAPoP szövegelemző portált és a Voyant eszköztárat használták. Megállapították, hogy minden csoportban a leggyakrabban használt szó a twitter volt. Táblázat mutatja a 10 leggyakrabban használt szót. Az összekombinált csoportokat, melyek további csupa-szítást igényelnének és a kiválasztás után létrejött listát külön táblázat mutatja. Megállapították, hogy a „hálózat” szó a lista tetején van, mivel gyakran szerepel a kivonatokban a „közösségi” szó az alábbi kifejezésekben: közösségi hálózat, közösségi Érdekesség, hogy az „új” szót két tárgykörben használták: Könyvtár, Technológia, ahol az új megközelítésekkel volt összekapcsolható. Ez eltér

az „új” szó Kommunikációban használt jelentésétől, mivel itt aktuális eseményekre utalt. A Keresés és Biztonság tárgykörök átfedik a Web és Tartalom kategóriákat ugyanúgy, mint az általános szavak, amelyek kifejezik, hogy a kutatók csak részben érdeklődnek ilyen anyagok iránt.

Végső következtetések

A tanulmány a legátfogóbb tudományos munka ebben a témakörben, ami több mint 1000 kivonat áttekintését tartalmazza. A Twitter fókuszú tanulmányok 3 dimenzióra bonthatók:

1. Aspektus, 2. Módszer és 3. Tartomány és módszertani szempontból négy csoportra osztályozhatók:

1. Elemzés, 2. Tervezés és Fejlesztés,
3. Vizsgálat, 4. Tudásfelfedezés.

WILLIAMS, Shirley A. – TERRAS, Melissa M. – WARWICK, Claire: What do people study when they study Twitter? Classifying Twitter related academic papers. = Journal of Documentation, 2013. 3. sz. p. 384–410./

(Sörény Edina)