

## TÉZISGYŰJTEMÉNY

**Kollányi Bence**

# **A közösségi média kommunikáció automatizációja**

Bot kódok fejlesztésének vizsgálata a GitHubon  
és a nyílt-forráskódú botok használata a Twitteren

Témavezető:  
**Ságvári Bence, PhD**

Budapest, 2022

Szociológia és Kommunikációtudomány Doktori Iskola

## TÉZISGYŰJTEMÉNY

# A közösségi média kommunikáció automatizációja

Bot kódok fejlesztésének vizsgálata a GitHubon  
és a nyílt-forráskódú botok használata a Twitteren



**Kollányi Bence**

## Tartalomjegyzék

<b>1. Kutatási előzmények és a téma indoklása</b> .....	<b>4</b>
<b>2. A felhasznált módszerek</b> .....	<b>6</b>
<b>3. Kutatási kérdések</b> .....	<b>7</b>
<b>3.1. Botfejlesztőkre vonatkozó kutatási kérdések</b> .....	<b>8</b>
<b>3.2. Kutatási kérdések a Twitter botokkal kapcsolatban</b> .....	<b>9</b>
<b>4. Az értekezés tudományos eredményei</b> .....	<b>11</b>
<b>5. A témakörrel kapcsolatos publikációk jegyzéke</b> .....	<b>14</b>
<b>5.1. Folyóirat cikkek</b> .....	<b>14</b>
<b>5.2. Konferencia prezentáció és publikáció</b> .....	<b>14</b>
<b>6. Főbb hivatkozások</b> .....	<b>15</b>

## 1. Kutatási előzmények és a téma indoklása

A közösségi média platformok fejlődése lehetővé tette, hogy egyes intézményi szereplők, így például vállalkozások és kormányzati intézmények, automatizálják a közösségi média oldalakon zajló kommunikációt. A programozott közösségi média profilok (botok) az automatikus kommunikáció mellett akár online tranzakciók lebonyolítására is képesek. Az automatizált közösségi média profilok jelenléte és a botok által használt technológia fejlesztése eközben egyre erősebb hatással van az online kommunikációs szféra erőviszonyaira. A közösségi média botok fejlesztése és használata megváltoztatja a hatalmi viszonyokat, új technológia szereplők számára biztosít jelentős erőforrásokat. A bot technológia elterjedése hatással van a marketingre, a politikára, és egyes esetekben az emberek magánkommunikációját is átalakítja. A terület fontossága ellenére nagyon kevés tudással rendelkezünk a botok előállításáról és a GitHub kódmegosztó oldal szerepéről a bot kódok fejlesztésében és megosztásában.

Az automatizált közösségi média profilokat a szakirodalom gyakran állítja be a közösségi médiára leselkedő veszélyként – a botok kéretlen reklámüzeneteket küldenek a felhasználóknak, rosszindulatú kódokat terjesztenek vagy éppen hamis profilokat szerveznek hálózatba, azért hogy bizonyos politikai célokat támogassanak.

A korai publikációk közül számos tanulmány foglalkozott az automatizált Twitter-fiókok felderítésével, beazonosításával, illetve azzal hogy hogyan lehet megakadályozni, hogy ezek a botok spamet vagy a rosszindulatú webhelyekre mutató linkek terjesszenek (Chu et al., 2010; Lee et al., 2011; Song et al., 2011). A szakirodalom a kifinomultabb közösségi botok (social bots) jelenségével is foglalkozik – ezek a botok képesek az emberi kommunikációt utánozni, ezért lényegesen nehezebben detektálhatóak, mint a korai egyszerű botok (Ferrara et al., 2016). Ezeknek a közösségi botoknak a fejlesztése és használata mögött sokféle motiváció húzódhat meg. Stieglitz et al. (2017) felismerték, hogy ezek a botok egyaránt szolgálhatnak semleges, jó és rossz célokat. Az automatizált közösségi média-fiókokkal foglalkozó tágabb szakirodalom ír a marketing vagy spam botokról (Chen & Subramanian, 2018, Cresci et al., 2017), az akadémiai publikációkat népszerűsítő robotokról (Haustein et al., 2014), a média által használt hírbotokról és chatbotokról (Diakopoulos, 2019; Jones és Jones, 2019), de beszámol arról is, hogy automatizált közösségi médiaprofilokat használnak a válságkommunikációban (Brachten és mtsai, 2018; Hofeditz és mtsai, 2019), illetve a politikában (Bastos és Mercea, 2019; Caldarelli et al., 2020; Ferrara, 2017; Howard, Woolley és Calo, 2018; Wooley és Howard,

2016a, 2016b), valamint olyan társadalmi célok elérése érdekében, mint az online zaklatás visszaszorítása (Geiger, 2016).

A doktori kutatásom kezdetén írtam egy cikket a nyílt forráskódú Twitter botok fejlesztéséről és a botok kódjának GitHubon való közzétételéről (Kollanyi, 2016). A GitHub a nyílt forráskódú fejlesztők számára az egyik legfontosabb együttműködési felület, ezen a platformon osztják meg jelenleg a legtöbb számítógépes programhoz fejlesztett kódot (Gousios et al., 2014). Ezért a GitHub alkalmas arra, hogy megértsük a nyílt forráskódú Twitter botok működését, illetve tanulmányozzuk a botfejlesztés közösségi aspektusait.

A közelmúltban számos tanulmány vizsgálta GitHub-adatok segítségével a szabad forráskódú programok fejlesztői között megvalósuló együttműködést. A fokozott tudományos érdeklődést részben a relatív könnyen hozzáférhető adatok indokolják, részben az adatgyűjtés és az adatok elemzése során használt technológiák fejlődése magyarázza.

A doktori kutatásom elsődleges célja, hogy megértssem a Twitteren tevékenykedő automatizált felhasználók mögött futó kód írásának és használatának gyakorlatát. A botfejlesztés folyamatának és a botok gyakorlati használatának megértése érdekében doktori kutatásomban egyaránt vizsgálom a GitHub platformon elérhető kódokat (a botok működése) és Twitteren zajló, gyakran botokat és emberi felhasználókat egyaránt érintő kommunikációs folyamatot. A kutatási módszerem egyesíti a közösségi média platformokon zajló automatizált kommunikáció tanulmányozását a botok mögött futó forráskód vizsgálatával.

A digitális adatok összegyűjtésére építő kutatási módszertan és a klasszikus társadalomtudományi megközelítés ötvözése segít megérteni a botok írásának és használatának gyakorlatát. A doktori kutatásom részeként elsőként elemeztem a GitHubon található bot repozitóriumokról elérhető információkat (a botok forráskódjait és a botok kódjaira vonatkozó meta-adatokat). Ezt követően egy kérdőíves vizsgálat keretében megkerestem a botokat fejlesztő programozókat, hogy minél inkább betekintést nyerjek a botkód írásának folyamatába. A kérdőíves vizsgálat abban segített, hogy megismerjem a fejlesztési szakasz legfontosabb kihívásait és a botok (gyakran hosszú-távú) üzemeltetésének legfontosabb nehézségeit. A kérdőíves felmérés eredményei arra is rávilágítanak, hogy az automatizált közösségi média-fiókok létrehozása mögött mik lehetnek a legfontosabb motivációk. Arra is kíváncsi voltam, hogy a Twitter platformon a nem-automatizált (azaz emberi) felhasználók hogyan reagálnak a botok jelenlétére – ezt elsősorban a botok kommunikációjára adott reakciók (például kedvelés vagy megosztás)

elemzésével tudtam megvizsgálni. Dolgozatom utolsó része a fenti adatok alapján a nyílt forráskódú közösségi média-botok tipológiájának kialakításához járul hozzá.

## 2. A felhasznált módszerek

Kutatásomban egy olyan új módszert javasoltam, amelynek segítségével a Twitter botok GitHubon található forráskódja és az egyes botok Twitteren keresztül elérhető kommunikációját egyszerre lehet elemezni. A kutatási projektek jellemzően „fekete dobozként” tekintenek a botokra, és csupán a Twitteren hagyott nyomok elemzése révén kívánják megérteni az automatizált fiókok mögött futó algoritmusokat. Saját kutatásom ezzel szemben a botokat a GitHub platform vizsgálatával egészen forráskódjukig követi, így a bot mögött futó algoritmus (kódot) elemzését össze tudtam kötni a Twitteren tevékenykedő botokról gyűjtött adatokkal. E két adatforrás összekapcsolásával bepillantást nyerhettem a Twitter botok belső működésébe, és egy nagy, „való világot leíró” adatbázis segítségével vizsgálhattam meg, hogy a robotokba épített algoritmusok milyen kommunikációs mintázatokat hoznak létre.

A doktori kutatásomhoz kidolgozott adatgyűjtés egy négylépéses, gyakran iteratív folyamat volt, amely magában foglalja a kezdeti adatgyűjtést a GitHub platformról a GitHub API-jain keresztül, a botfejlesztők kérdőíves vizsgálatát, a Twitter-alapú adatgyűjtést és az ismét a GitHubra fókuszáló célzott adatgyűjtést. Ezek a lépések egymásra épültek, ezért csak meghatározott sorrendben lehetett őket végrehajtani.

Az adatgyűjtést azzal kezdtem, hogy megvizsgáltam, hogyan lehet hozzáférni a legnagyobb online kódtár, a GitHub adataihoz. A platform lehetővé teszi, hogy adott kulcsszavakra keressek rá egy API segítségével. A gyakorlatban ez azt jelentette, hogy írtam egy programot, ami a repozitóriumok nevében és leírásában keresett. Ennek az adatgyűjtésnek az eredménye egy olyan adatbázis lett, amelyben a bot repozitóriumok mellett valamennyi olyan fejlesztőről találhatóak adatok, akik legalább egy saját bot repozitóriummal rendelkeznek a GitHubon.

A GitHub-alapú adatgyűjtést követően felvettem a kapcsolatot a fejlesztőkkel, hogy további kérdéseket tegyek fel többek között az alábbi témákkal kapcsolatban: a botfejlesztés mögött meghúzódó motivációk, a botok programozásához szükséges készségek, kommunikáció más bot fejlesztőkkel, valamint a bot fejlesztésével és a botok üzemeltetésével (a Twitteren történő használatukkal) kapcsolatos legnagyobb kihívások.

Az adatgyűjtés első szakaszában összeállított GitHub-adatbázis lehetővé tette, hogy a repók metaadataiból kinyerhessem valamennyi fejlesztő elérhetőségi adatait, feltéve, hogy a fejlesztő nyilvánossá tette e-mail címét a platformon. A kérdőíves kutatás lebonyolítására a Survey Gizmo nevű fizetős, online kérdőíves szolgáltatást használtam, amely lehetővé tette az összes fejlesztővel felvegyem a kapcsolatot csoportos e-mailek küldésével.

A bot repozitórium egy része tartalmazott vagy egy Twitter felhasználónevet vagy egy Twitterre mutató linket a repozitórium leírásban. Sok esetben az repozitóriumba feltöltött readme.md fájl tartalmazta ezt az információt. Ez az információ azonban nem minden esetben volt elérhető, ezért úgy döntöttem, hogy beteszek egy kérdést a kérdőívbe, amelyben arra kérem a fejlesztőket, hogy adják meg a Twitteren használt botok Twitter elérhetőségét (felhasználó nevét). Így össze tudtam kapcsolni a GitHubon lévő fejlesztők adatait (beleértve a bot repozitóriumokra vonatkozó adatokat) a Twitteren működő botokkal. A botfejlesztőkkel kapcsolatban a kérdőíves vizsgálatból további adatokkal rendelkezem, amelyeket később szintén fel tudtam használni az elemzésben.

Az alábbi lista a kutatási projektem adatgyűjtési szakaszában gyűjtött négy fő adatbázist mutatja be:

**GitHub adatbázis:** metaadatok hozzávetőlegesen 19 ezer nyílt-forráskódú bot repozitóriumról;

**Kérdőíves vizsgálat eredményei:** válaszok 860 bot fejlesztőtől;

**Twitter adatbázis:** 321 Twitteren használt nyílt-forráskódú bot és a hozzájuk kapcsolódó GitHub repozitóriumok adatai, valamint 500 ezer letöltött tweet üzenet és az üzenetekre érkezett felhasználói reakciók;

**Manuális kódolt adatbázis:** felcímkézett adatbázis a fent említett 321 Twitter botról.

### 3. Kutatási kérdések

A következő részben azt az öt kutatási kérdést mutatom be, amelyekkel PhD kutatási projektemben foglalkoztam. Ezek a kutatási kérdések két fő téma köré összpontosulnak. Az első három kutatási kérdés fókuszában a nyílt forráskódú Twitter-botokat létrehozó fejlesztők állnak, illetve maga a botfejlesztés folyamata. A második kérdéscsoport a Twitter botok tevékenységét vizsgálja, beleértve, hogy a botok hogyan állítanak elő tartalmakat a Twitteren.

### 3.1. Botfejlesztőkre vonatkozó kutatási kérdések

Ezek a kérdések a bot fejlesztőkre vonatkoznak. A botfejlesztők egyaránt lehetnek professzionális fejlesztők és olyan „amatőrök”, akik képesek bot kódot írni.

***K1.1** Mik a nyílt forráskódú Twitter botok fejlesztésének és a botok forráskódjának megosztásának legfontosabb gyakorlatai? Mik a legfontosabb okai, hogy ebben a folyamatban a fejlesztők GitHub platformot használják?*

Ezek a kérdések arra összpontosítanak, hogy a fejlesztők hogyan használják a GitHubot, például milyen gyakran frissítik a kódjukat, mennyi információt osztanak meg más fejlesztőkkel a repozitóriumokban, kapnak-e támogatást más fejlesztőktől, stb. A kérdések megválaszolásához a GitHub hozzáférést biztosít szinte az összes fontosabb metaadathoz (a repozitóriumokra vonatkozó adatokhoz). A GitHub használatának okaira és a Twitter bot fejlesztésének motivációira vonatkozó kérdésekben azonban kérdőíves adatokra támaszkodtam.

***K1.2** Hogyan sajátítják el a fejlesztők a Twitter botok fejlesztéséhez szükséges készségeket? Mennyiben határozza meg, illetve mennyire könnyíti meg a programozással kapcsolatos általános tudás a Twitter botok fejlesztését?*

A korábbi kutatásaimra támaszkodva, itt abból indultam ki, hogy a botok fejlesztéséhez programozási ismeretek szükségesek, ugyanakkor a botkódokat egészen változatos fejlesztői háttérrel rendelkező fejlesztők hozzák létre. A fejlesztők között találhatunk informatikus háttérrel rendelkező programozókat, társadalomtudósokat, újságírókat és művészeket egyaránt. Ugyanakkor a botok fejlesztése, mint általában a nyílt forráskódú szoftverek programozása, egyre inkább decentralizált, a szereplők a konkrétabb igényekhez igazítható, újra felhasználható kódokra épülő kódbázisra és megosztott tudásra épül. Ezt a feltételezést szerettem volna alaposabban megvizsgálni, amikor a fenti kutatási kérdésre kerestem a választ. Számos különböző információforrás áll rendelkezésre a fejlesztők számára a Twitter botok létrehozásáról szóló blogbejegyzésektől az egyetemi kurzusokon át a GitHubon elérhető botkódokig. Melyek a legfontosabb információforrások a botfejlesztéshez, és mit tesz a botfejlesztő, ha egy váratlan problémával találja szemben magát? A kutatási kérdés megválaszolásakor elsősorban a kérdőíves vizsgálat eredményeire támaszkodtam.



**K1.3** *Létezik-e egy valódi bot-fejlesztői közösség a GitHubon vagy a különféle botokhoz tartozó kódot magányos fejlesztők hozzák létre?*

Bár a GitHub lehetővé teszi, hogy több fejlesztő működjön együtt egy adott szoftver projektben, korábbi kutatásaim alapján megállapítottam, hogy a GitHubon elérhető botkódok közel 90 százalékát csak egy fejlesztő hozta létre.

A GitHubon található repozitóriumok nemcsak nyomon követik a kódhoz hozzáféréssel rendelkező fejlesztők által végrehajtott összes változtatást, hanem azt is rögzítik, hogy ki járult hozzá a projekthez egy kódrészlet javasolásával vagy csak egy probléma (bug) bejelentésével. Mindezek az adatok nyilvánosan elérhetők az általam vizsgált repozitóriumokkal kapcsolatban a GitHubon keresztül, és fontos betekintést nyújtanak abba, hogy mennyire elterjedt a botfejlesztési együttműködés, és létezik-e egy aktív fejlesztői közösség a GitHubon, amely a botokra összpontosít.

## 3.2. Kutatási kérdések a Twitter botokkal kapcsolatban

**K2.1** *Hogyan hoznak létre, dolgoznak fel és tesznek közzé tartalmakat a nyílt forráskódú botok a Twitteren?*

Mit tudhatunk meg a Twitter botokról, ha párhuzamosan megvizsgáljuk a GitHubon található forráskódokat és a Twitteren használt botok tevékenységét? E két adatforrás (a GitHub és a Twitter) újszerű kombinálásával leírhatom a nyílt forráskódú Twitter botok működését. Az elemzés során egyaránt támaszkodhatok a Twitter fiókok leírásaira, valamint a GitHubon közzétett kódokra. A fenti adatok mindegyike összegyűjthető a GitHub és a Twitter különböző API-jain keresztül.

Hoznak-e létre önállóan tartalmat a Twitter botok? Ha igen, hogyan generálják ezek a fiókok a tartalmat, melyek a felhasznált fő információforrások? Egyes Twitter botok áttekinthető módon bemutatják azokat a forrásokat, amelyekre támaszkodnak, akár a bot bemutatkozó szövegében (bio) a Twitteren, akár magukban a tweetekben. Ez utóbbira jó példa egy a tweetben található link, amely a GitHub oldalra vagy az információ forrására mutat. Ezeknek a forrásoknak a megismerése és annak megértése, hogy a bot hogyan hoz létre tartalmakat gyakran csak korlátozott mértékben megvalósítható vagy akár lehetetlen anélkül, hogy megvizsgálnánk a bot mögött futó kódot.

A kutatási kérdés megválaszolásához azt is megpróbálom számszerűsíteni, hogy a nyílt forráskódú botok által a Twitteren generált forgalom mekkora része eredeti tartalom, és mekkora tartalom származik pusztán retweetekből vagy más Twitter-fiókok által közölt

tartalmak idézéséből.

**K2.2** Milyen hosszú a nyílt forráskódú botok életciklusa, és mekkora forgalmat generál ezalatt egy átlagos bot a Twitteren? Mik a legfontosabb nehézségei egy Twitter bot hosszabb ideig tartó működtetésének, és miért tiltják ki a botokat a Twitterről vagy válnak azok inaktívvá?

Egy Twitter bot életciklusa többféleképpen is meghatározható. Megnézhetjük a fiók által generált első és utolsó tweetek között eltelt időt Twitter adatok segítségével, vagy kiszámolhatjuk a bot forráskódjának fejlesztésére és alkalmankénti frissítésére fordított időt a GitHubon. A legtöbb esetben ezek az időszakok átfedik egymást, de a bot-fiókokat gyakran felfüggeszti a Twitter vagy maga a programozó kapcsolja le a botot és használja fel a Twitter profilt vagy a akár magát a kódot egy másik projektben. Ha egy bot élettartama hosszabb, a kódfejlesztő(k)nek olyan problémákkal kell foglalkoznia/niuk, mint például a Twitter API-k működésének vagy a bot (adott esetben Twitteren kívüli) információforrásainak változása. Azt is próbálom megérteni, hogy egyes botok miért szüntetik be a tevékenységüket. A legfontosabb nehézségek és kihívások felderítéséhez egyrészt a Twitteren még (részben) elérhető, de nem működő vagy inaktív botok tanulmányozására építék, részben pedig a felmérésben magukat a fejlesztőket kérdezem meg.

## 4. Az értekezés tudományos eredményei

**A1.** A GitHubon közzétett forráskódok elemzése alapján megállapítható, hogy a Twitter botok a legkülönbözőbb feladatok végzésére programozhatóak be. A botok mögött futó kód ráadásul viszonylag könnyen újra felhasználható, adaptálható új feladatokhoz. Kérdés azonban, hogy azok akik nem rendelkeznek magas szintű programozási ismeretekkel, vagy nem tudnak fizetni drága IT-szolgáltatásokért, mennyire könnyen tudnak a szabadon elérhető kódok alapján saját botot fejleszteni, illetve a botokat üzemeltetni a Twitteren. A kutatásom eredményei alapján, annak ellenére, hogy a botokkal kapcsolatos kód és ismeretek mindenki számára elérhetőek, nem beszélhetünk a bot technológia valódi demokratizálódásáról. A felmérés eredményei azt mutatják, hogy a botfejlesztők túlnyomó többsége továbbra is erős programozói háttérrel rendelkezik. A botfejlesztők közel 78 százaléka rendelkezik formális programozási vagy számítástechnikai képesítéssel (vagy vesz részt ilyen jellegű képzésben). Egy bot fejlesztése sokszor nem túl bonyolult feladat, a kódrészletek rendelkezésre állnak, de mások kódjának megértése mégis igényel bizonyos szintű programozási ismereteket.

**A2.** A dolgozatomban bemutatott eredmények alapján elmondható, hogy a GitHubon található botkódok túlnyomó többségén csupán egy-egy szerző dolgozott és a kódot rövid idő alatt hozta létre. A bot repozitóriumok több mint 40 százaléka csupán egy egynapos projekt volt. Bizonyos esetben a GitHubon elérhető adatok megtévesztőek, mert a szóban forgó botokat a GitHubon kívül is fejleszthették – ilyen esetekben a szerző csak a végső kódot töltötte fel a platformra, a GitHubot úgymond kódtárolóként (code dumpster) használva. Mind a rövid fejlesztési idő, mind a közös fejlesztés hiánya azt jelzi, hogy a platformot a bot fejlesztők bot projektjeik során jellemzően nem használják a különböző tudással rendelkező fejlesztők összekapcsolására vagy az eltérő földrajzi helyekről dolgozó fejlesztők közös munkafelületeként.

**A3.** Érdekes módon az egynél több fejlesztő által fejlesztett projekteket hosszabb ideig karbantartják, illetve maga a fejlesztés is tovább tart. Ezek a projektek gyakrabban és jellemzően nagyobb mennyiségben kapnak visszajelzéseket más fejlesztőktől – például ritkábban ismerik el a fejlesztők munkáját más GitHub felhasználók egy-egy csillaggal. Az egy szerző által fejlesztett bot kódok közül mindössze 10 esetből 1-szer készül olyan másolat amellyel más programozók dolgoznak tovább a kódon. (Ezt “fork”-nak nevezik a platform szóhasználata szerint.) Ezzel szemben a több szerző által gondozott projektek fele rendelkezett legalább egy forkkal. Az egyszemélyes projektek átlagosan csupán 1 csillaggal

kaptak, míg a többszerzős repozitóriumok átlagosan 9 csillagot kaptak. Az egyszerűs projektet átlagosan 134 napig, míg a több szerző által fejlesztett / gondozott bot repozitóriumokat 428 napig tartották karban. Ez egyaránt magyarázható a kód összetettségével (a többszereplős projektek jellemzően bonyolultabbak, így több a hibalehetőség is a kódokban) és az ezekre a repozitóriumokra irányuló fokozott figyelemmel.

**A4.** Az esteleges iskolai programozási ismereteken és más GitHubon elérhető bot kódok tanulmányozásán túl egy tipikus fejlesztő számos más forrásból is tájékozódik a botfejlesztés közben. A fejlesztők számára a legfontosabb információforrásokat a Q&A (kérdések és válaszok) oldalak (pl. egy adott probléma megoldása) és a témával kapcsolatos blogbejegyzések (pl. egy bot fejlesztésének lépésről lépésre történő leírása) jelentik. Az a tény, hogy inkább egyéni tájékozódással, információ kereséssel oldják meg a fejlesztők a problémáikat, nem pedig a GitHub fejlesztői közösséghez fordulnak, jelentős részben összefügg a GitHub felépítésével és működésével. Bár a platform célja az együttműködés és a csapatmunka ösztönzése, az egyéni fejlesztőknek nem igazán vannak jó lehetőségeik arra, hogy kikérjék más fejlesztők véleményét vagy segítségért folyamodjanak. A GitHubon a botfejlesztők körében végzett felmérésből az is kiderült, hogy a botfejlesztők jellemzően nem arra használják a platformot, hogy mások (bot) repozitóriumaikhoz járuljanak hozzá vagy ilyen jellegű segítséget kérjenek más fejlesztőktől. A kollaborációnál sokkal fontosabb a fejlesztők számára a GitHub egyik alapfunkciójának, a verziókezelő rendszernek a használata.

**A5.** A kérdőíves vizsgálat eredményei alapján az is elmondható, hogy a botfejlesztők sok esetben a GitHubon elsősorban azért teszik közzé bot kódjaikat, hogy ezzel ismertségre tegyenek szert és így építsék karrierjüket. Különösen igaz ez a fiatal fejlesztőkre. Számukra az elismerés és a karrierépítés a GitHub platform használatának legfőbb motivációja. Amellett, hogy a fejlesztők megvalósítanak egy-egy ötletet, ami kipattant a fejükből, a botok programozása, fejlesztése sok fejlesztő számára egyfajta tanulási folyamat. Amikor megkérdeztem a fejlesztőket a bot létrehozásának legfontosabb okáról, a felmérésben résztvevők többsége azt az opciót választotta, hogy a Twitter platform megértése, az API-ok működésének megismerése „nagyon fontos” vagy „fontos” szempont volt a botprojekt elindításában. Érdekes módon az önkifejezés sokkal kevésbé volt fontos, és a túlnyomó többség elutasította azt a válaszlehetőséget, hogy a botot egy politikai ügy támogatására fejlesztené. A fejlesztőkkel folytatott háttérbeszélgetésekből egyértelműen kiderült, hogy a fejlesztők gyakran azért tesznek közzé botkódokat GitHub-profiljukon, hogy jelezzék,

személyes projektjeik is vannak, és a munkájukon kívül is érdeklődnek a programozás iránt.

A felmérésben részt vevők által megadott Twitter-felhasználónevek és a GitHubon megtalálható bot repozitóriumok adatai alapján sikerült összeállítanom egy listát, amely 321 párosított bot repozitóriumot tartalmaz. (A párosítás itt arra vonatkozik, hogy a bot repozitórium és a Twitteren futó bot között egyértelmű kapcsolat tárható fel.) Az egyes Twitter (bot) profilokon megjelent tartalmakból a Twitter REST API-ján keresztül közel 500 000 tweetet tudtam letölteni az elemzésemhez. Először ezt a Twitter-adatbázist hasonlítottam össze a GitHubról gyűjtött adatokkal. A Twitteren lévő botok regisztrációs ideje és a botok által generált tartalom időbeli megoszlása jól illeszkedik a GitHubon található bot repozitóriumok számának időbeli változásához. Ez azt is sugallja, hogy a felmérés válaszadói jól reprezentálják a nyílt forráskódú botfejlesztők egészét.

**B1.** A GitHubon található szabad-forráskódú botok számának változása és a Twitteren vizsgált botok által generált tartalom időbeli alakulása egyaránt arra utalnak, hogy az elmúlt hat évben dinamikusan megnőtt a Twitter botok népszerűsége a fejlesztők körében.

**B2.** A tweetek időbeli eloszlásának modellezésével becsléseket tudtam adni a nyílt forráskódú botok által generált forgalom mennyiségére vonatkozóan is. Ez egy módszertani újdonság, és hasonló körülmények között használható más automatizált fiókok által generált tartalom elemzésére.

**B3.** A vizsgált nyílt forráskódú botok által a Twitteren közzétett tartalom túlnyomó többsége eredeti tartalom volt, nem pedig a platformon már meglévő tartalom további felhasználása (lásd retweetek vagy idézetek). Ez talán az egyik legérdekesebb megállapítása a Twitter adatokon végzett elemzésnek. Bár a bot repozitóriumok módszeres elemzése beazonosított olyan Twitter botokat, amelyek vagy csak felerősítik más Twitter felhasználók tartalmait (automatikusan újratweetelve azok üzeneteit), vagy a Twitteren korábban publikált tartalmakat hasznosítják újra valamilyen kiválasztási szempont mentén (pl. csak bizonyos kulcsszavakat tartalmazó vagy csak bizonyos népszerűségű tartalmakat retweetelnek). Mégis, a legtöbb szabad forráskódú bot saját tartalmat állít elő, illetve az előállított tartalmak nagy része ilyen botoktól származik. Míg a generatív botok egy kód alapján állítják elő a tartalmaikat (például egy algoritmus alapján hoznak létre képeket), addig a “továbbítók” más webes platformokról (gyakran API-ok segítségével) válogatnak tartalmakat, majd az így kiválasztott információt teszik közzé a Twitteren.

## 5. A témakörrel kapcsolatos saját publikációk jegyzéke

### 5.1. Folyóirat cikkek

Marchal, Nahema, Lisa-Maria Neudert, Bence Kollanyi, and Philip N. Howard.

‘Investigating Visual Content Shared over Twitter during the 2019 EU Parliamentary Election Campaign’. *Media and Communication* 9, no. 1 (3 February 2021): 158–70.  
<https://doi.org/10.17645/mac.v9i1.3421>.

Bradshaw, Samantha, Philip N. Howard, Bence Kollanyi, and Lisa-Maria Neudert. ‘Sourcing and Automation of Political News and Information over Social Media in the United States, 2016-2018’. *Political Communication* 37, no. 2 (3 March 2020): 173–93.  
<https://doi.org/10.1080/10584609.2019.1663322>.

Neudert, Lisa-Maria, Philip Howard, and Bence Kollanyi. ‘Sourcing and Automation of Political News and Information During Three European Elections’. *Social Media + Society* 5, no. 3 (July 2019): 205630511986314.  
<https://doi.org/10.1177/2056305119863147>.

Kollanyi, Bence. ‘Where Do Bots Come from? An Analysis of Bot Codes Shared on GitHub’. *International Journal of Communication* 10 (2016): 4932–51.

### 5.2. Konferencia prezentáció és publikáció

Machado, C., Kira, B., Narayanan, V., Kollanyi, B., & Howard, P. (2019, May). A Study of Misinformation in WhatsApp groups with a focus on the Brazilian Presidential Elections. In Companion Proceedings of The 2019 World Wide Web Conference (WWW '19), Ling Liu and Ryen White (Eds.). ACM, New York, NY, USA.

Gorwa, R., Kollanyi, B., & Howard, P. (2018, May). A critical analysis of bot detection methodologies In: Methodological Challenges to Studying Misinformation and Disinformation in Data-Driven Politics: Fake News, Bots, and Digital Campaigns, Panel Discussion at the ICA 2018, Prague.

Kollanyi, B. (2016, June). How to write a Twitter bot? Bot Codes Shared on GitHub. Algorithms, Automation and Politics, Preconference of the ICA 2016, Fukuoka.

## 6. Főbb hivatkozások

- Assenmacher, Dennis, Lena Clever, Lena Frischlich, Thorsten Quandt, Heike Trautmann, and Christian Grimme. 'Demystifying Social Bots: On the Intelligence of Automated Social Media Actors'. *Social Media + Society* 6, no. 3 (July 2020): 205630512093926. <https://doi.org/10.1177/2056305120939264>.
- Bastos, Marco T., & Dan Mercea. 'The Brexit Botnet and User-Generated Hyperpartisan News'. *Social Science Computer Review* 37, no. 1 (February 2019): 38–54. <https://doi.org/10.1177/0894439317734157>.
- Bessi, Alessandro, & Emilio Ferrara. 'Social Bots Distort the 2016 US Presidential Election Online Discussion'. *First Monday* 21, no. 11–7 (2016). <https://doi.org/10.5210/fm.v21i11.7090>.
- Brachten, Florian, Milad Mirbabaie, Stefan Stieglitz, Olivia Berger, Sarah Bludau, & Kristina Schrickel. 'Threat or Opportunity? - Examining Social Bots in Social Media Crisis Communication'. In *Australasian Conference on Information Systems 2018*. University of Technology, Sydney, 2018. <https://doi.org/10.5130/acis2018.bo>.
- Bright, Jonathan, Scott Hale, Bharath Ganesh, Andrew Bulovsky, Helen Margetts, & Phil Howard. 'Does Campaigning on Social Media Make a Difference? Evidence From Candidate Use of Twitter During the 2015 and 2017 U.K. Elections'. *Communication Research* 47, no. 7 (October 2020): 988–1009. <https://doi.org/10.1177/0093650219872394>.
- Broniatowski, David A., Amelia M. Jamison, SiHua Qi, Lulwah AlKulaib, Tao Chen, Adrian Benton, Sandra C. Quinn, and Mark Dredze. 'Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate'. *American Journal of Public Health* 108, no. 10 (October 2018): 1378–84. <https://doi.org/10.2105/AJPH.2018.304567>.
- Bruns, Axel, & Jean Burgess. 'Twitter Hashtags from Ad Hoc to Calculated Publics'. *Hashtag Publics: The Power and Politics of Discursive Networks*, 2015, 13–28.
- Caldarelli, Guido, Rocco De Nicola, Fabio Del Vigna, Marinella Petrocchi, & Fabio Saracco. 'The Role of Bot Squads in the Political Propaganda on Twitter'.

*Communications Physics* 3, no. 1 (December 2020): 81.

<https://doi.org/10.1038/s42005-020-0340-4>.

Chavoshi, Nikan, Hossein Hamooni, & Abdullah Mueen. 'DeBot: Twitter Bot Detection via Warped Correlation'. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 817–22. Barcelona, Spain: IEEE, 2016.

<https://doi.org/10.1109/ICDM.2016.0096>.

Chen, Zhouhan, & Devika Subramanian. 'An Unsupervised Approach to Detect Spam Campaigns That Use Botnets on Twitter'. *ArXiv:1804.05232 [Cs]*, 14 April 2018.

<http://arxiv.org/abs/1804.05232>.

Chen, Zhouhan, Rima S. Tanash, Richard Stoll, & Devika Subramanian. 'Hunting Malicious Bots on Twitter: An Unsupervised Approach'. In *Social Informatics*, edited by Giovanni Luca Ciampaglia, Afra Mashhadi, & Taha Yasseri, 10540:501–10. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017. [https://doi.org/10.1007/978-3-319-67256-4\\_40](https://doi.org/10.1007/978-3-319-67256-4_40).

Chu, Zi, Steven Gianvecchio, Haining Wang, & Sushil Jajodia. 'Who Is Tweeting on Twitter: Human, Bot, or Cyborg?' In *Proceedings of the 26th Annual Computer Security Applications Conference on - ACSAC '10*, 21. Austin, Texas: ACM Press, 2010. <https://doi.org/10.1145/1920261.1920265>.

Chu, Zi, Steven Gianvecchio, Haining Wang, & Sushil Jajodia. 'Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg?' *IEEE Transactions on Dependable and Secure Computing* 9, no. 6 (November 2012): 811–24.

<https://doi.org/10.1109/TDSC.2012.75>.

Cresci, Stefano. 'A Decade of Social Bot Detection'. *Communications of the ACM* 63, no. 10 (23 September 2020): 72–83. <https://doi.org/10.1145/3409116>.

Cresci, Stefano, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi, & Maurizio Tesconi. 'The Paradigm-Shift of Social Spambots: Evidence, Theories, and Tools for the Arms Race'. *Proceedings of the 26th International Conference on World Wide Web Companion - WWW '17 Companion*, 2017, 963–72.

<https://doi.org/10.1145/3041021.3055135>.



- Dabbish, Laura, Colleen Stuart, Jason Tsay, & Jim Herbsleb. 'Social Coding in GitHub: Transparency and Collaboration in an Open Software Repository'. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work - CSCW '12*, 1277. Seattle, Washington, USA: ACM Press, 2012.  
<https://doi.org/10.1145/2145204.2145396>.
- Diakopoulos, Nicholas. *Automating the News: How Algorithms Are Rewriting the Media*. Cambridge, Massachusetts: Harvard University Press, 2019.
- Edwards, Chad, Autumn Edwards, Patric R. Spence, & Ashleigh K. Shelton. 'Is That a Bot Running the Social Media Feed? Testing the Differences in Perceptions of Communication Quality for a Human Agent and a Bot Agent on Twitter'. *Computers in Human Behavior* 33 (April 2014): 372–76.  
<https://doi.org/10.1016/j.chb.2013.08.013>.
- Efthimion, Phillip George, Scott Payne, & Nicholas Proferes. 'Supervised Machine Learning Bot Detection Techniques to Identify Social Twitter Bots' 1, no. 2 (2018): 71.
- Efthimion, Phillip George, Scott Payne, & Nicholas Proferes. 'Supervised Machine Learning Bot Detection Techniques to Identify Social Twitter Bots' 1, no. 2 (2018): 71.
- Ferrara, Emilio, Onur Varol, Clayton Davis, Filippo Menczer, & Alessandro Flammini. 'The Rise of Social Bots'. *Communications of the ACM* 59, no. 7 (24 June 2016): 96–104. <https://doi.org/10.1145/2818717>.
- Ferrara, Emilio. 'Disinformation and Social Bot Operations in the Run up to the 2017 French Presidential Election', 2017, 33.
- Floridi, Luciano. *Information: A Very Short Introduction*. OUP Oxford, 2010.
- Forsgren, N., G., Ccarelli, D., Fordi, V., Gennarell, Y., Huang, & Zimmerman, T. (2020). 2020 State of the Octoverse: Empowering healthy communities.  
<https://octoverse.github.com/>.

- Gaffney, Devin, & Cornelius Puschmann. 'Data Collection on Twitter'. *Twitter and Society* 55 (2014): 67.
- Geiger, R. Stuart. 'Bot-Based Collective Blocklists in Twitter: The Counterpublic Moderation of Harassment in a Networked Public Space'. *Information, Communication & Society* 19, no. 6 (2 June 2016): 787–803. <https://doi.org/10.1080/1369118X.2016.1153700>.
- Gilani, Zafar, Reza Farahbakhsh, Gareth Tyson, Liang Wang, & Jon Crowcroft. 'Of Bots and Humans (on Twitter)'. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 349–54. Sydney Australia: ACM, 2017. <https://doi.org/10.1145/3110025.3110090>.
- Gilani, Zafar, Reza Farahbakhsh, Gareth Tyson, and Jon Crowcroft. 'A Large-Scale Behavioural Analysis of Bots and Humans on Twitter'. *ACM Transactions on the Web* 13, no. 1 (20 February 2019): 1–23. <https://doi.org/10.1145/3298789>.
- Gorwa, Robert, & Douglas Guilbeault. 'Unpacking the Social Media Bot: A Typology to Guide Research and Policy'. *Policy & Internet* 12, no. 2 (June 2020): 225–48. <https://doi.org/10.1002/poi3.184>.
- Gousios, Georgios, Bogdan Vasilescu, Alexander Serebrenik, & Andy Zaidman. 'Lean GHTorrent: GitHub Data on Demand'. In *Proceedings of the 11th Working Conference on Mining Software Repositories - MSR 2014*, 384–87. Hyderabad, India: ACM Press, 2014. <https://doi.org/10.1145/2597073.2597126>.
- Haustein, Stefanie, Timothy D Bowman, Benoît Macaluso, Cassidy R Sugimoto, & Vincent Larivière. 'Measuring Twitter Activity of ArXiv E-Prints and Published Papers', 3, 2014. <https://doi.org/10.6084/m9.figshare.1041514>.
- Hofeditz, Lennart, Christian Ehnis, Deborah Bunker, Florian Brachten, & Stefan Stieglitz. 'Meaningful Use of Social Bots? Possible Applications in Crisis Communication during Disasters', 17. Stockholm & Uppsala, Sweden, 2019.
- Howard, Philip N., Samuel Woolley, & Ryan Calo. 'Algorithms, Bots, and Political Communication in the US 2016 Election: The Challenge of Automated Political Communication for Election Law and Administration'. *Journal of Information*

- Technology & Politics* 15, no. 2 (3 April 2018): 81–93.  
<https://doi.org/10.1080/19331681.2018.1448735>.
- Hwang, Tim, Ian Pearce, & Max Nanis. ‘Socialbots: Voices from the Fronts’. *Interactions* 19, no. 2 (March 2012): 38–45. <https://doi.org/10.1145/2090150.2090161>.
- Jensen, Jens F. ‘Interactivity: Tracking a New Concept in Media and Communication Studies’ 12 (1998): 20.
- Jones, Bronwyn, & Rhianne Jones. ‘Public Service Chatbots: Automating Conversation with BBC News’. *Digital Journalism* 7, no. 8 (2019): 1032–53.  
<https://doi.org/10.1080/21670811.2019.1609371>.
- Kalliamvakou, Eirini, Kelly Blincoe, Leif Singer, Daniel M German, & Daniela Damian. ‘The Promises and Perils of Mining GitHub (Extended Version)’, 39, 2014.
- Kollanyi, Bence. ‘Where Do Bots Come from? An Analysis of Bot Codes Shared on GitHub’. *International Journal of Communication* 10 (2016): 4932–51.
- Kudugunta, Sneha, & Emilio Ferrara. ‘Deep Neural Networks for Bot Detection’. *Information Sciences* 467 (October 2018): 312–22.  
<https://doi.org/10.1016/j.ins.2018.08.019>.
- Lazer, D., A. Pentland, L. Adamic, S. Aral, A.-L. Barabasi, D. Brewer, N. Christakis, et al. ‘SOCIAL SCIENCE: Computational Social Science’. *Science* 323, no. 5915 (6 February 2009): 721–23. <https://doi.org/10.1126/science.1167742>.
- Lazer, David M. J., Alex Pentland, Duncan J. Watts, Sinan Aral, Susan Athey, Noshir Contractor, Deen Freelon, et al. ‘Computational Social Science: Obstacles and Opportunities’. *Science* 369, no. 6507 (28 August 2020): 1060–62.  
<https://doi.org/10.1126/science.aaz8170>.
- Lee, Kyumin, James Caverlee, & Steve Webb. ‘The Social Honey-pot Project: Protecting Online Communities from Spammers’. In *Proceedings of the 19th International Conference on World Wide Web - WWW '10*, 1139. Raleigh, North Carolina, USA: ACM Press, 2010. <https://doi.org/10.1145/1772690.1772843>.

- Lee, Kyumin, Brian David Eoff, & James Caverlee. 'Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter'. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 8, 2011.
- Lima, Antonio, Luca Rossi, & Mirco Musolesi. 'Coding Together at Scale: GitHub as a Collaborative Social Network'. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 10, 2014.
- Lin, Po-Ching, & Po-Min Huang. (2013, January). A study of effective features for detecting long-surviving Twitter spam accounts. In *Proceedings of the 15th International Conference on Advanced Communication Technology* (pp. 841–846). Pyeonchang, Korea: IEEE.
- Lokot, Tetyana, & Nicholas Diakopoulos. 'News Bots: Automating News and Information Dissemination on Twitter'. *Digital Journalism* 4, no. 6 (17 August 2016): 682–99. <https://doi.org/10.1080/21670811.2015.1081822>.
- Luceri, Luca, Ashok Deb, Silvia Giordano, and Emilio Ferrara. 'Evolution of Bot and Human Behavior during Elections'. *First Monday*, 31 August 2019. <https://doi.org/10.5210/fm.v24i9.10213>.
- Lumezanu, Cristian, Nick Feamster, & Hans Klein. '#bias: Measuring the Tweeting Behavior of Propagandists'. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media*, 8. Dublin, Ireland, 2012.
- Manovich, Lev. *The Language of New Media*. MIT press, 2002.
- Marres, Noortje. 'Foreword by Noortje Marres'. In Snee, Helene, Christine Hine, Yvette Morey, Steven Roberts, and Hayley Watson, eds. *Digital Methods for Social Science*. London: Palgrave Macmillan UK, 2016. <https://doi.org/10.1057/9781137453662>.
- Maus, Gregory, & Onur Varol. 'A Typology of Socialbots', 2017, 8.
- Metaxas, Panagiotis Takis, & Eni Mustafaraj. 'From Obscurity to Prominence in Minutes: Political Speech and Real-Time Search', 7. Raleigh, NC, USA, 2010.
- Minnich, Amanda, Nikan Chavoshi, Danai Koutra, & Abdullah Mueen. 'BotWalk: Efficient Adaptive Exploration of Twitter Bot Networks'. In *Proceedings of the 2017*

*IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 467–74. Sydney Australia: ACM, 2017.

<https://doi.org/10.1145/3110025.3110163>.

Murthy, Dhiraj. *Social Communication in the Twitter Age (Digital Media and Society)*. Cambridge: Polity Press, 2013.

Mustafaraj, Eni, & Panagiotis Takis Metaxas. ‘What Edited Retweets Reveal about Online Political Discourse’. In *Conference: Analyzing Microtext, Papers from the 2011 AAAI Workshop*, 6. San Francisco, California, USA, 2011.

Oentaryo, Richard J., Arinto Murdopo, Philips K. Prasetyo, & Ee-Peng Lim. ‘On Profiling Bots in Social Media’. In *Social Informatics*, edited by Emma Spiro and Yong-Yeol Ahn, 10046:92–109. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016. [https://doi.org/10.1007/978-3-319-47880-7\\_6](https://doi.org/10.1007/978-3-319-47880-7_6).

Puschmann, Cornelius, & Jean Burgess. ‘The Politics of Twitter Data’. *SSRN Electronic Journal*, 2013. <https://doi.org/10.2139/ssrn.2206225>.

Ratkiewicz, J, M D Conover, M Meiss, B Goncalves, A Flammini, & F Menczer. ‘Detecting and Tracking Political Abuse in Social Media’. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 8, 2011.

Rebillard, Franck, & Annelise Touboul. ‘Promises Unfulfilled? “Journalism 2.0”, User Participation and Editorial Policy on Newspaper Websites’. *Media, Culture & Society* 32, no. 2 (March 2010): 323–34. <https://doi.org/10.1177/0163443709356142>.

Rogers, Richard. *Digital Methods*. Cambridge, Massachusetts: The MIT Press, 2013.

Rogers. ‘Political Research in the Digital Age’. *International Public Policy Review* 8, no. 1 (2014): 73-87.

Schultz, Tanjev. ‘Interactive Options in Online Journalism: A Content Analysis of 100 U.S. Newspapers’. *Journal of Computer-Mediated Communication* 5, no. 1 (23 June 2006): 0–0. <https://doi.org/10.1111/j.1083-6101.1999.tb00331.x>.

- Song, Jonghyuk, Sangho Lee, & Jong Kim. 'Spam Filtering in Twitter Using Sender-Receiver Relationship'. In *Recent Advances in Intrusion Detection*, edited by Robin Sommer, Davide Balzarotti, & Gregor Maier, 6961:301–17. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.  
[https://doi.org/10.1007/978-3-642-23644-0\\_16](https://doi.org/10.1007/978-3-642-23644-0_16).
- Spyridou, Paschalia-Lia, & Andreas Veglis. 'Exploring Structural Interactivity in Online Newspapers: A Look at the Greek Web Landscape'. *First Monday*, 2008.
- Stieglitz, Stefan, Florian Brachten, Björn Ross, & Anna-Katharina Jung. 'Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts', 2017, 11.
- Storey, Margaret-Anne, Leif Singer, Brendan Cleary, Fernando Figueira Filho, & Alexey Zagalsky. 'The (R) Evolution of Social Media in Software Engineering'. In *Future of Software Engineering Proceedings*, 100–116. Hyderabad India: ACM, 2014.  
<https://doi.org/10.1145/2593882.2593887>.
- Stukal, Denis, Sergey Sanovich, Richard Bonneau, & Joshua A. Tucker. 'Detecting Bots on Russian Political Twitter'. *Big Data* 5, no. 4 (December 2017): 310–24.  
<https://doi.org/10.1089/big.2017.0038>.
- Suchacka, Grazyna, & Jacek Iwański. 'Identifying Legitimate Web Users and Bots with Different Traffic Profiles — an Information Bottleneck Approach'. *Knowledge-Based Systems* 197 (June 2020): 105875.  
<https://doi.org/10.1016/j.knosys.2020.105875>.
- Sundar, S. Shyam. 'Theorizing Interactivity's Effects'. *The Information Society* 20, no. 5 (November 2004): 385–89. <https://doi.org/10.1080/01972240490508072>.
- Takhteyev, Yuri, & Andrew Hilt. 'Investigating the Geography of Open Source Software through GitHub', 2010, 10.
- Tao, Y., Dang, Y., Xie, T., Zhang, D., & Kim, S. (2012, November). How do software engineers understand code changes? An exploratory study in industry. In *Proceedings of the ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering* (pp. 1-11).

- Tarte, S., Willcox, P., Glaser, H., & De Roure, D. (2015, June). Archetypal narratives in social machines: approaching sociality through prosopography. In Proceedings of the ACM web science conference (pp. 1-10).
- Toffler, Alvin. *The Third Wave*. A Bantam Book. New York Toronto London Sydney Auckland: Bantam Books, 1990.
- Tromble, Rebekah. 'Thanks for (Actually) Responding! How Citizen Demand Shapes Politicians' Interactive Practices on Twitter'. *New Media & Society* 20, no. 2 (February 2018): 676–97. <https://doi.org/10.1177/1461444816669158>.
- Varol, Onur, Emilio Ferrara, Clayton A Davis, Filippo Menczer, & Alessandro Flammini. 'Online Human-Bot Interactions: Detection, Estimation, and Characterization'. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, 10, 2017. <https://arxiv.org/abs/1703.03107>.
- Wang, Alex Hai. 'Don't Follow Me: Spam Detection in Twitter'. In *2010 International Conference on Security and Cryptography (SECRYPT)*, 10. Athens, 2010.
- Woolley, Samuel C., & Philip N. Howard. (2016). Social media, revolution, and the rise of the political bot. *Routledge handbook of media, conflict, and security*. New York, NY: Routledge, 282-292.
- Woolley, Samuel C. 'Automating Power: Social Bot Interference in Global Politics'. *First Monday*, 21, no. 4 (2016). <https://doi.org/10.5210/fm.v21i4.6161>.
- Wu, Yu, Jessica Kropczynski, Patrick C. Shih, & John M. Carroll. 'Exploring the Ecosystem of Software Developers on GitHub and Other Platforms'. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW Companion '14*, 265–68. Baltimore, Maryland, USA: ACM Press, 2014. <https://doi.org/10.1145/2556420.2556483>.
- Zagalsky, Alexey, Joseph Feliciano, Margaret-Anne Storey, Yiyun Zhao, & Weiliang Wang. 'The Emergence of GitHub as a Collaborative Platform for Education'. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1906–17. Vancouver BC Canada: ACM, 2015. <https://doi.org/10.1145/2675133.2675284>.