Stellungnahme zum Themenkomplex „Social Bots“

Prof. Dr.-Ing. Florian Gallwitz, Technische Hochschule Nürnberg, 14.2.2020

Bevor ich auf Ihre Fragen im Einzelnen eingehe, möchte ich Ihnen zum besseren Verständnis kurz die Hintergründe meiner Beschäftigung mit dem Thema „Social Bots“ darlegen. Weiterhin werde ich die Methoden beschreiben, mit denen ich versucht habe, dem Phänomen auf den Grund zu gehen. Danach fasse ich die zentralen methodischen Probleme der vorliegenden Studien zusammen, auf denen der verbreitete Glaube an „Social Bots“ und auch die anhaltende Falschberichterstattung über „Social Bots“ in den Medien fußt. Abschließend möchte ich meine Einschätzungen zum Thema stichpunktartig zusammenfassen.

Hintergrund

Meine ernsthafte Beschäftigung mit dem Thema „Social Bots“ begann Anfang Dezember 2018. In dieser Zeit wurden die angeblichen Ergebnisse einer unveröffentlichten „Studie“ der Berliner Firma Botswatch in den deutschen Medien verbreitet, wonach 28 Prozent der Tweets in der Debatte zum UNO-Migrationspakt von „Social Bots“ ausgegangen wären.

Aufgrund meiner einschlägigen Fachkenntnisse und Erfahrungen sowohl im Bereich Mustererkennung (die Erkennung von etwaigen „Social Bots“ wäre ein klassisches Mustererkennungsproblem) als auch bei der Entwicklung von natürlichsprachlichen Dialogsystemen („Social Bots“ sollen ja menschliche Nutzer imitieren und mit diesen natürlichsprachlich interagieren) kamen mir erhebliche Zweifel am Wahrheitsgehalt dieser Darstellung. Diese brachte ich schon im Dezember 2018 in mehreren Interviews in Presse und Rundfunk zum Ausdruck. Erhärtet wurde mein Verdacht durch die Weigerung der Fa. Botswatch, auch nur einen einzigen der vermeintlichen „Social Bots“ konkret zu benennen.

In dieser Zeit wurde mir klar, dass nicht nur die Studie der Firma Botswatch, sondern auch die Resultate der akademischen Forschung zum Thema „Social Bots“ höchst zweifelhaft sind. In diesem Zusammenhang stieß ich auf die Ergebnisse des Berliner Datenjournalisten Michael Kreil, dem bereits im Jahr 2017 fundamentale methodische Probleme dieser Forschungsrichtung aufgefallen waren. In einem Vortrag auf dem Chaos Communication Congress im Dezember 2017 hatte er bereits öffentlich hierauf aufmerksam gemacht[1]. Im Dezember 2018 fasste er den aktuellen Stand seiner Erkenntnisse in einem Blog-Artikel[2] zusammen und machte darin deutlich, dass die Ergebnisse der einschlägigen Wissenschaft auf einen kleinen Kreis von Akteuren zurückzuführen sind, die mit fragwürdigen Methoden operieren. Diese Methoden ordnen zahllose Menschen fälschlich als „Social Bots“ ein.

Wie sich im Verlauf der Untersuchungen, die ich in den letzten 13 Monaten teils in enger Zusammenarbeit mit Michael Kreil durchgeführt habe herausgestellt hat, sind die angeblichen Funde von „Social Bots“ sogar *vollständig* durch solche Fehlerkennungen erklärbar.

Methoden

Um die Frage zu klären, ob „Social Bots“ überhaupt existieren, habe ich versucht, dem Phänomen mit sehr unterschiedlichen Ansätzen weiter auf den Grund zu gehen:

---

[1] https://media.ccc.de/v/34c3-9268-social_bots_fake_news_und_filterblasen#t=2627

[2] https://blog.info.graphics/social-bot-research-is-flawed/

- Sichtung der wissenschaftlichen Studien und kritische Prüfung der Methodik
- Überprüfung der Hintergründe zahlreicher Medienberichte in englischer, deutscher und zum Teil auch in spanischer Sprache über angebliche „Social-Bot-Attacken“ aus aller Welt, auf der Grundlage entsprechender Stichwortsuchen u.a. in Google News.
- Kontaktaufnahme mit einschlägigen Wissenschaftlern, stets mit der Bitte um ein konkretes Beispiel für einen „Social Bot“
- Durchführung von zum Teil sehr umfangreichen empirischen Untersuchungen in Zusammenarbeit mit Michael Kreil, u.a. der Versuch der vollständigen Replikation einer Studie, die zahllose „Social Bots“ unter den Followern der Twitter-Accounts deutscher Parteien ausgemacht haben wollte.
- Recherchen zur Identität von als „Social Bot“ verdächtigten Menschen bzw. direkte Kontaktaufnahme mit angeblichen „Social Bots“
- Herstellung von öffentlicher Aufmerksamkeit für meine Suche nach einem Exemplar eines „Social Bot“ durch wiederholte Aufrufe auf Twitter, u.a. durch einen Tweet von 18.12.2018, der alleine über 80.000 „Views“ verzeichnete, und durch einen Artikel im Tagesspiegel vom 3.6.2019 (gemeinsam mit Michael Kreil), der online in deutscher[3] und englischer[4] Sprache veröffentlicht wurde und der auf Twitter große Aufmerksamkeit erhielt.

Diese Untersuchungen und ihre Ergebnisse habe ich auf meinem Twitter-Account @FlorianGallwitz laufend öffentlich dokumentiert.

Im Ergebnis habe ich bis heute weltweit nicht ein einziges Beispiel für einen Twitter-Account gefunden, der auch nur im Ansatz den gängigen Vorstellungen und Definitionen eines „Social Bots" entspricht. Jeder einzelne Fund eines angeblichen „Social Bot“ stellte sich spätestens bei genauerem Hinsehen als ein von einem Menschen geführter Account heraus.

Bemerkenswert ist, dass *keiner* der von mir kontaktierten Autoren von wissenschaftlichen Studien zu „Social Bots“, in denen oft von zahllosen „Social-Bot-Funden“ berichtet worden war, mir auch nur ein *einziges* konkretes Exemplar eines „Social Bots“ benennen konnte. Die einschlägigen Studien dieser Forschungsrichtung werden nämlich – entgegen den Regeln guter wissenschaftlicher Praxis – stets ohne Rohdaten und fast immer ohne konkrete Accountnamen veröffentlicht. Auch die von mir kontaktierten „Social-Bot-Experten“, die sich in der Vergangenheit in der Presse und sogar im Rahmen von parlamentarischen Anhörungen über dieses Phänomen geäußert hatten, konnten mir kein einziges konkretes Exemplar eines „Social Bots“ nennen.

Fundamentale methodische Probleme der „Social-Bot-Forschung“

Studien, die quantitive Aussagen über „Social Bots“ machen, basieren fast immer auf einem oder mehreren der folgenden methodischen Fehler:

1. „Social Bots“ werden häufig von „menschlichen Accounts“ dadurch unterschieden, dass sie eine willkürlich vorgegebene Schwelle an Aktivität überschreiten. Üblich ist hier das vom Oxford-Professor Philip Howard popularisierte Kriterium „50 Tweets pro Tag“. Zahllose menschliche Nutzer überschreiten aber diese Schwelle problemlos und regelmäßig, zum Teil sogar um ein Vielfaches, darunter z.B. der Bundestagsabgeordnete Johannes Kahrs oder der Meteorologe Jörg Kachelmann. Alle Accounts, die diese Schwelle überschreiten, werden jedoch in einschlägigen Studien ohne weitere Sichtung oder Prüfung konsequent als „Bots“ gezählt und bezeichnet.

[3] https://background.tagesspiegel.de/digitalisierung/die-maer-von-social-bots

[4] https://background.tagesspiegel.de/digitalisierung/the-social-bot-fairy-tale

2. Alternativ werden „Social Bots“ in der neueren „wissenschaftlichen“ Literatur zum Thema heute meist dadurch identifiziert, dass sie mittels eines automatischen Tools, meist „Botometer“, als „Bot“ klassifiziert werden. Die Bewertungen des „Botometer“ sind in der Praxis jedoch kaum von Zufallszahlen zu unterscheiden. Zahllose Journalisten, Bundestagsabgeordnete, amerikanische Kongressabgeordnete, ja sogar Nobelpreisträger werden als „Bots“ erkannt. Wirklich automatisierte Accounts, wie der Bot `@big_ben_clock`, der seit 10 Jahren die Glockenschläge des Londoner Wahrzeichens verkündet, werden vom „Botometer“ dagegen nicht als Bot erkannt. Auch in diesen Studien wird regelmäßig und vermutlich sehr bewusst darauf verzichtet, die zahllosen erkannten „Bots“ auch nur flüchtig darauf zu überprüfen, ob es sich tatsächlich um „Social Bots“ handeln könnte. Bei einem Versuch, eine fragwürdige Studie nachzuvollziehen, die nach dieser Methode zahllose „Social Bots“ unter den Followern der deutschen Parteien gefunden haben wollte, identifizierte „Botometer“ 270.000 von 520.000 Accounts als „Bots“. Bei einer manuellen Überprüfung einer repräsentativ ausgewählten Stichprobe von 109 dieser 270.000 angeblichen Bots fand sich jedoch kein einziger Account, der auch nur entfernt einen automatisierten Eindruck gemacht hätte, geschweige denn ein „Social Bot“.[5] Mit der gleichen bizarren Methodik ließe sich nachweisen, dass es sich bei einem erheblichen Teil der Passanten auf dem Times Square in New York um Yetis handelt.[6] Auf die gleiche Weise wurde beispielsweise auch die Behauptung belegt, die kürzlich weltweit durch die Presse ging, dass „Bots“ sich in die Diskussion um die Ursachen der Buschbrände in Australien eingemischt hätten. Für das Vorhandensein von „Bots“ ließ sich bei näherem Hinsehen kein Hinweis finden.
3. Als weiteres angebliches Erkennungszeichen für Bots wird immer wieder eine achtstellige Ziffernfolge am Ende des Twitter-Handles gewertet. Auf dieser Basis[7] veröffentlichte etwa die Londoner Times im November 2019 einen alarmistischen Artikel mit dem Titel *„Army of Twitter bots follow top politicians such as Nicola Sturgeon, John Swinney, Jo Swinson“*. Tatsächlich hängt Twitter seit einiger Zeit an die meisten neu angelegten Accounts automatisch eine achtstellige Ziffernfolge an. Gibt man bei der Account-Eröffnung etwa die Vornamen-Nachnamen-Kombination Hans Meier an, so erhält man ganz automatisch einen Twitter-Handle wie etwa `@Hans28345136`. Dieser lässt sich zwar nachträglich noch ändern, allerdings ist das vielen Twitter-Neuankömmlingen nicht bewusst.
4. Schließlich wird das gehäufte Auftreten von neuen Accounts im Kontext politischer Unruhe gerne als Beleg für das Vorhandensein von „Social Bots“ gewertet. So wurde etwa im November 2019 im Zusammenhang mit dem Sturz von Evo Morales in Bolivien vielfach der Vorwurf laut, hier seien „Bots“ im Spiel, die von den Gegnern von Evo Morales kontrolliert würden. Viele der verdächtigen Accounts trugen dann auch noch eine 8-stellige Nummer im Twitter-Handle (s.o.). Tatsächlich ließ sich das gehäufte Auftreten dieser neuen Accounts dadurch erklären, dass unter den Gegnern von Evo Morales auf Facebook Anleitungen kursierten, die dazu aufforderten, einen Twitter-Account zu eröffnen und dort entsprechend aktiv zu werden, um die Weltöffentlichkeit besser erreichen zu können.[8]

[5] https://twitter.com/FlorianGallwitz/status/1141442509467541506?s=20

[6] https://twitter.com/FlorianGallwitz/status/1222177040914374658?s=20

[7] https://twitter.com/SashaTalavera/status/1193963253438844928?s=20

[8] https://twitter.com/FlorianGallwitz/status/1197635228933804032?s=20

Einen ausgezeichneten, ausführlichen und aktuellen Überblick über die methodischen Probleme der „Social-Bot-Forschung“ in englischer Sprache mit zahlreichen Beispielen hat Michael Kreil unter der Folgenden URL zusammengestellt: https://michaelkreil.github.io/openbots/.

Fazit

- Die Idee, dass von sinistren Mächten gelenkte „Social Bots“ in den Sozialen Medien unterwegs sind, um Einfluss auf die öffentliche Meinungsbildung zu nehmen, ist eine Verschwörungstheorie ohne Grundlage in der Realität.
- In einer ganzen Reihe von Aspekten ist die Social-Bot-Verschwörungstheorie der bizarren „Chemtrail“-Verschwörungstheorie nicht unähnlich. Sie greift aber auch Vorstellungen auf, wie sie aus der mittelalterlichen Hexenverfolgung bekannt sind. Hierzu gehören insbesondere die zahlreichen Anleitungen, die im Internet kursieren, welche Merkmale auflisten, anhand derer man angeblich „Social Bots“ von „echten Menschen“ unterscheiden können soll.
- Nach meiner Überzeugung würde keine einzige der vorliegenden empirischen Studien, in denen quantitative Aussagen über „Social Bots“ gemacht werden, einer Überprüfung nach den in der Wissenschaft sonst üblichen Maßstäben standhalten.
- Zu welchen Anteilen lediglich die Inkompetenz der beteiligten Wissenschaftler und zu welchen Anteilen eine bewusste Täuschungsabsicht die Ergebnisse dieser „Forschungsrichtung“ erklären können, kann ich nicht abschließend beurteilen. Es gibt jedoch Indizien, die aus meiner Sicht zumindest bei einem Teil der Veröffentlichungen stark für eine Täuschungsabsicht sprechen. Dazu gehört, dass es Autoren einschlägiger Studien bei entsprechenden Nachfragen vollkommen bewusst zu sein scheint, dass es sich bei keinem der von ihnen in ihren Veröffentlichungen als „Social Bots“ gezählten Accounts tatsächlich um „Social Bots“ handelt.
- Dass die Social-Bot-Verschwörungstheorie sogar an angesehenen Universitäten Fuß fassen konnte und einschlägige, methodisch hanebüchene Studien z.T. in renommierten wissenschaftlichen Fachzeitschriften platziert werden konnten, lässt sich aus meiner Sicht nur durch falsche Anreizsysteme in der Wissenschaft und systematische Schwächen des Peer-Review-Verfahrens erklären.
- Begünstigt wurde die Verbreitung der Social-Bot-Verschwörungstheorie durch realitätsferne Vorstellungen von den Möglichkeiten sogenannter „Künstlicher Intelligenz“. Zum einen wird der Stand der Technik im Bereich natürlichsprachlicher Mensch-Maschine-Interaktion maßlos überschätzt. Zum anderen fehlt vielen „Social-Bot-Forschern“ das grundlegendste Verständnis für die Funktionsweise von maschinellem Lernen und ganz generell für die wissenschaftliche Methode. Nur so ist es zu erklären, dass mit Tools wie „Botometer“ auf bizarre Weise versucht wird, den vermeintlichen „Social Bots“ auf die Schliche kommen, und dass dessen erratische Ausgaben anschließend auch noch ungeprüft als Wahrheit akzeptiert werden.
- Die große öffentliche Resonanz auf Medienberichte über angebliche „Social-Bot-Armeen“ zeigt, dass es einem überschaubaren Kreis von Forschern innerhalb von wenigen Jahren gelungen ist, diese Verschwörungstheorie in den Köpfen von vielen Journalisten, Politikern und einem großen Teil Netzöffentlichkeit zu verankern. Das belegt auch eine amerikanische Umfrage aus dem Jahr 2018.[9]
- Der Begriff „Bot“ wird (in Ermangelung tatsächlicher Bot-Sichtungen) zunehmend für menschliche Twitter-Nutzer verwendet, als eine Art Schimpfwort. „Bot-Erkennungs-Tools“ wie „BotSentinel“ behaupten im Kleingedruckten gar nicht mehr, „Bots“ im Sinne von automatisierten Accounts zu

[9] https://www.journalism.org/2018/10/15/most-americans-have-heard-about-social-media-bots-many-think-they-are-malicious-and-hard-to-identify/

finden, wie der Name des Dienstes vermuten lässt (übrigens anders als „Botometer“). Stattdessen geht es „BotSentinel“ laut Eigendarstellung um das Aufspüren von Menschen, die ein „Troll-artiges“ Verhalten zeigen sollen. Gemeint sind dabei offenbar in erster Linie Accounts mit relativ wenigen Followern, die mittels Reply auf Tweets von bekannten Politikern oder großen Medien-Accounts reagieren.

- Der Begriff „Bot“ dient heute in der öffentlichen Diskussion überwiegend als Kampfbegriff, mit dem Menschen mit einer politischen Meinung, die von der eigenen abweicht, in der politischen Debatte entmenschlicht und diskreditiert werden sollen.
- Als ganz aktuelles Beispiel für den Einsatz des Bot-Vorwurfs als politische Waffe soll an dieser Stelle ein Tweet der SPD-Landtagsfraktion NRW (`@spd_fraktion_nw`) dienen, die am 13.2.2020 twitterte:
  „Die #NoAfD versucht gerade, uns mit hundebebilderten Fake-Bots anzugreifen. Auch hier gilt wieder: Wir lassen uns von Faschisten nicht einschüchtern. Mehr dazu gleich um 10 Uhr in der Aktuellen Stunde. Ganz liebe Grüße, Eure SPD. #KlareKanteGegenRechts #WirSindMehr“
  Auf Twitter konnte ich keinen einzigen „hundebebilderten“ Account finden, der mit dem SPD-Account in den vorangegangenen Stunden und Tagen Kontakt hatte. Auf Facebook konnte ich einen(!) kritischen Kommentar unter einem aktuellen Post der SPD-Fraktion finden, der von einer Nutzerin stammt, die einen Hund als Profilbild verwendet. Weitere Belege für einen „Bot-Angriff“ oder dafür, dass der „Angriff“ (also das Hinterlassen des kritischen Kommentars) durch einen Bot oder durch die AfD erfolgt sei, konnte oder wollte die SPD-Fraktion NRW mir auch auf mehrfache Nachfrage nicht liefern.
- Ich halte die Tatsache, dass die Bot-Verschwörungstheorie von weiten Teilen der Öffentlichkeit unkritisch akzeptiert wird und von Politikern wie Medien ohne Hemmungen verbreitet und instrumentalisiert wird, für überaus bedenklich. Sie steht gleichermaßen einem rationalen Diskurs *über* die Sozialen Medien im Wege, wie sie die politische Debatte *in* den Sozialen Medien vergiftet. Twitter-Nutzer verdächtigen sich gegenseitig als „Bots“, Teilnehmer an politischen Debatten werden durch den Bot-Vorwurf auf perfide Weise entmenschlicht und politische Fehlentwicklungen werden als Folge des Wirkens angeblicher „Bot-Armeen“ abgetan, statt sich mit dem inhaltlichen Gehalt kritischer Äußerungen oder den tatsächlichen Ursachen solcher Entwicklungen zu befassen.

Zu Ihren Fragen im Einzelnen

Frage 1: Wie definieren Sie einen „Social Bot“?

Bei meiner Suche nach einem „Social Bot“ habe ich mich an der Definition orientiert, die das Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag in seiner Vorstudie zu „Social Bots“ von April 2017 gewählt hat, und die den gängigen Vorstellungen in der Literatur entspricht:

„Social Bots sind Computerprogramme, die eine menschliche Identität vortäuschen und für manipulative Zwecken eingesetzt werden, indem sie wie Menschen im Internet kommunizieren. Menschen, die mit Social Bots interagieren, nehmen diese nicht als durch Algorithmen gesteuerte automatische Kommunikation, sondern als menschliche Internetteilnehmer wahr und sind sich der Manipulation nicht bewusst.“

Kernmerkmale wären also

1. Automatisierte Interaktion mit Nutzern (als Minimalanforderung an einen sehr primitiven „Social Bot“ würde ich bereits das Retweeten oder „Liken“ von Tweets mit bestimmten Stichwörtern gelten lassen.)
2. Vortäuschung einer realen Person
3. Der Versuch, zu manipulieren bzw. Einfluss auf die Meinungsbildung zu nehmen

Weil der Großteil der wissenschaftlichen Literatur und fast alle Medienberichte über angebliche „Social Bots“ den Kurznachrichtendienst „Twitter“ als vermeintlichen Lebensraum von „Social Bots“ ausgemacht haben, beschränke ich mich in meiner Darstellung im Folgenden auf Twitter.

Zur Abgrenzung noch zwei Beispiele zum Begriff der automatisierten Interaktion:

- Der frühere NRW-Landtagsabgeordnete Daniel Schwerd (@netnrd) sendet schon seit einigen Jahren einen Teil seiner Tweets automatisiert, nämlich Links auf Meldungen des NRW-Landtags sowie Links auf Artikel des Satiremagazins „Der Postillon“, die er aus den entsprechenden RSS-Feeds direkt auf seinen Twitter-Account weiterleitet. Die Meldungen erscheinen dort in Tweets unter seinem Namen, auf den ersten Blick so, als hätte er selbst manuell die entsprechenden Artikel empfohlen. Hierfür verwendet er den Web-Dienst ITTT. Dies wird von Twitter auch unter den entsprechenden Tweets kenntlich gemacht, wenn man diesen direkt anklickt. Zwar handelt es sich hier um eine (in diesem Fall ohnehin unproblematische) Teilautomatisierung; eine automatische Interaktion mit anderen Twitter-Nutzern ist dies jedoch nicht.
- Im Mai 2018 berichtete BuzzFeed über ein Netzwerk von rund 70 Fake-Accounts auf Twitter, die sich gegenseitig folgen, und die sich positiv über Homöopathie äußern. Auffällig ist dabei, dass immer wieder die gleichen Formulierungen auf unterschiedlichen Accounts gepostet werden. Nach den Recherchen von BuzzFeed wurden oder werden dieses Accounts wohl von einer einzigen Person betrieben, möglicherweise teilautomatisiert. Eine automatisierte Interaktion der Accounts mit echten Nutzern hat hierbei jedoch offenbar nicht stattgefunden.

Frage 2: Welche Technologie steckt in „Social Bots“ - wie viel KI steckt in „Social Bots“ (Wissen / Erfahrungswerte)

Alle angeblichen Sichtungen von „Social Bots“, denen ich nachgegangen bin, haben sich als Falschmeldungen herausgestellt. Es handelte sich stets um von Menschen bediente Accounts. Daher muss ich mich darauf beschränken, mir hypothetisch vorzustellen, wie es wäre, einen manipulativen „Social Bot“ zu entwickeln, der versucht, in den Sozialen Medien wie ein Mensch zu wirken.

Aus meiner Erfahrung bei der Entwicklung von natürlichsprachlichen Dialogsystemen sowohl in der akademischen Forschung als auch im kommerziellen Umfeld kann ich hierfür relevante Erfahrungen beisteuern: Die Entwicklung von Chatbots im Kundenservice ist nach wie vor aufwendig und erfordert viel Handarbeit. Sobald man offene, natürlichsprachliche Eingaben des Nutzers zulässt, tut sich ein enormer Raum an Möglichkeiten auf, der sich selbst für thematisch sehr eng begrenzte Anwendungen nur in einem aufwendigen, iterativen Entwicklungsprozess mit zwischenzeitlichen Nutzertests einigermaßen in den Griff bekommen lässt. Maschinelle Lernverfahren helfen hierbei kaum.

Die Komplexität der Anwendungen, die sich so realisieren lassen, ist überschaubar, etwa die Zuordnung des Kunden zu einem spezialisierten Kundenberater und/oder das Ausfüllen von Formularwerten (Kundennummer, Flugnummer o.ä.). Eine sinnvolle politische Diskussion ist beim Stand der Technik natürlichsprachlicher Dialogsysteme bzw. Chatbots fernab der Realität.

Wenn es tatsächlich Chatbot-artige „Social Bots“ gäbe, wären diese beim heutigen Stand der Technik wie Chatbots im Kundenservice oder die Sprachassistenten Siri und Alexa auf ein festes Inventar von starren, fest

einprogrammierten Interaktionsmustern eingeschränkt und leicht daran zu erkennen, dass sie außerhalb eines sehr engen thematischen Kontextes nicht in der Lage wären, auf Äußerungen sinnvoll zu reagieren. Eine Ausdehnung auf breitere Themenfelder ist nur mit ungeheurem Aufwand möglich. Der in den letzten 5 Jahren jeweils als weltbester natürlichsprachlicher Chatbot ausgezeichnete Chatbot Mitsuku ist von seinem Programmierer über einen Zeitraum von 13 Jahren mit 350.000 Programmzeilen Code gefüttert worden, um die Illusion eines menschlichen Gesprächspartners wenigstens kurzzeitig zu erwecken. Trotzdem wird meist schon bei der ersten inhaltlich etwas tiefergehenden Nachfrage deutlich, dass man es hier nicht mit einem Menschen zu tun hat. Und Mitsuku ist nur auf harmlose Plaudereien ausgelegt, nicht auf politische Debatten. Für die Betreiber der Plattform Twitter wäre ein gehäuftes Auftreten von Mitsuku-ähnlichen Bots problemlos daran zu erkennen, dass sich dessen Äußerungen schnell wortgleich wiederholen, was sich mittels sogenannter Hash-Tabellen sehr einfach überprüfen lässt.

Die geschilderten Grenzen der Chatbot-Technologie gelten umso mehr, wenn es um tagesaktuelle politische Diskussionen geht. Die notwendigen wochenlangen Entwicklungszyklen würden dazu führen, dass die entsprechende Diskussion auf Twitter bei Fertigstellung des thematisch spezialisierten Bots längst beendet ist.

Auf dieses Problem hingewiesen, weichen die Vertreter der Social-Bot-Hypothese üblicherweise in zwei diametral entgegengesetzte Richtungen aus:

- Auf der einen Seite wird, etwa von Prof. Dirk Helbing von der ETH Zürich, die Hypothese in den Raum gestellt, dass die unbekannten Betreiber der bislang unerkannt gebliebenen „Social Bots“ über einen enormen technologischen Vorsprung gegenüber dem allgemein bekannten Stand der Technik und der Technologie der amerikanischen Internet-Riesen verfügen könnten („Die Geheimdienste können es besser“[10]). Diese Hypothese bedarf m.E. keiner ernsthaften Diskussion. In die gleiche Richtung geht auch die immer wieder gehörte Entschuldigung für das Fehlen glaubwürdiger Bot-Sichtungen, „Social Bots“ seien mittlerweile „evolviert“ und inzwischen so intelligent, dass man sie praktisch nicht mehr von echten Menschen unterscheiden könne.
- Andere „Social-Bot-Forscher“ sind dagegen so realistisch einzuräumen, dass es wohl tatsächlich keine interaktiven, Chatbot-artigen „Social Bots“ gebe. Sie verbinden dies jedoch gerne mit der Behauptung, dass es aber wenigstens „simple“ politische Bot geben müsse, welche automatisiert und in hoher Frequenz vordefinierte Botschaften verbreiten oder politische Botschaften anderer Accounts durch Retweeten „amplifizieren“ würden.

Auch auf solche „simplen“ politischen Bots bin ich allerdings bei meiner Suche nach einem „Social Bot“ nicht gestoßen. Die Erklärung, warum auch diese im politischen Umfeld offenbar nicht existieren, liegt auf der Hand:

- Reichweite lässt sich auf Twitter vor allem durch zwei Faktoren herstellen: Durch eine hohe Zahl von Followern und/oder durch besonders interessante, witzige, originelle oder provokante Tweets. Solche Tweets werden über den Retweet-Mechanismus geteilt und erreichen oft ein Vielfaches der eigenen Follower-Zahl, selten sogar Hunderttausende oder gar Millionen von Nutzern („viraler Tweet“).
- Eine hohe Follower-Zahl erhöht die Chance, dass sich ein besonders interessanter, witziger, origineller oder provokanter Tweet an sehr viele Nutzer verbreitet.
- Ein langfristig hoher Anteil interessanter, witziger, origineller oder provokanter Tweets und Retweets erhöht die Chance, dass man neue Follower dazugewinnt und erhöht somit die Reichweite.

---

https://twitter.com/DirkHelbing/status/1126063443314905088?s=20

- Vorhersehbare, unoriginelle, langweilige, bedeutungsleere oder repetitive Tweets oder Retweets, wie sie sich mit einem Computerprogramm erzeugen ließen, führen dagegen zum Verlust von Followern und damit zum Verlust von Reichweite.
- Eine hohe Tweet-Frequenz beschleunigt den Follower-Verlust weiter: Niemand möchte in seiner Timeline mit großen Mengen wahllos ausgewählter Retweets und automatisch generierter Botschaften bombardiert werden, zwischen denen die Tweets der anderen Accounts untergehen, denen man folgt.

Die (Teil-)Automatisierung von Accounts würde also zum Verlust von Followern und somit zum Verlust von Reichweite führen. Ein sinistrer Akteur, der durch die Kontrolle von Twitter-Accounts auf die politische Stimmung eines ganzen Landes Einfluss zu nehmen sucht, müsste also genau den gegenteiligen Weg gehen. Er müsste versuchen, durch Beschäftigung von besonders witzigen oder originellen menschlichen Autoren eine hohe Followerzahl aufzubauen und durch sorgfältig formulierte, wohldosierte Tweets einen Einfluss auf die öffentliche Debatte zu gewinnen. Die Konkurrenz zum Teil sehr talentierter Nutzer auf Twitter, die ohne böse Hintergedanken mit genau der gleichen Strategie um Aufmerksamkeit ringen, ist allerdings groß. Im Februar 2020 meldete Twitter 152 Millionen Nutzer, die täglich aktiv sind.

Bleibt noch die immer wieder vorgebrachte Idee, durch eine große Zahl von automatisierten Accounts ohne menschliche Follower, die hochfrequent politische Botschaften mit bestimmten Stichwörtern retweeten, vielleicht wenigstens einen Einfluss auf die Sortierung von Tweets in der nicht-chronologischen Timeline oder auf die sogenannten „Twitter-Trends“ zu nehmen. Doch diese Strategie scheitert allem Anschein nach schon daran, dass die entsprechenden Algorithmen von Twitter (ähnlich wie längst auch die Google-Suche) ausgefeilt genug sind, um gegen solch simple Manipulationsversuche robust zu sein, etwa dadurch, dass in Ranking-Entscheidungen die Followerzahl oder Reputation der Nutzer einbezogen wird, die mit bestimmten Trends oder Tweets interagiert haben.

Frage 3: Was ist aus Ihrer Sicht der wissenschaftliche Stand zu der Frage, ob Social Bots (in nennenswertem Umfang) existieren?

Die Aussagen zur Existenz bzw. zur Häufigkeit von „Social Bots“ aus der sogenannten „Social-Bot-Forschung“ beruhen auf groben methodischen Fehlern und/oder auf bewusster Irreführung. Nach über einem Jahr ebenso intensiver wie vergeblicher Suche nach auch nur einem einzigen Exemplar eines „Social Bot“ gehe ich davon aus, dass „Social Bots“ auf der Plattform Twitter überhaupt nicht existieren, ganz sicher nicht in nennenswertem Umfang.

Frage 4: Wenn ja, woran kann man „Social Bots“ einigermaßen zuverlässig erkennen bzw. woran erkennt man, dass die bisher benannten „Social Bots“ menschlicher Natur sind?

In den allermeisten Fällen, in denen angebliche „Social Bots“ konkret benannt wurden oder von den einschlägigen Tools („Botometer“) identifiziert wurden, hat es sich als sehr einfach herausgestellt, zu erkennen, dass die „Social Bots“ in Wahrheit menschliche Nutzer sind. Nur in sehr wenigen Fällen war eine tiefergehende Recherche erforderlich:

- Meist genügt schon ein Blick auf die Tweets eines Accounts, idealerweise auf die Interaktionen mit anderen Nutzern, um zu erkennen, dass es sich um Menschen handelt. Verständnis für andere Benutzeräußerungen, inhaltliche Kommentare zu geteilten Links, Ironie, Sarkasmus und sinnvolle Kommentare zu tagesaktuellen Fragen sind als Gesamtbild meist fernab jeder Automatisierbarkeit.
- Die verschiedentlich von „Bot-Forschern“ vorgebrachte Behauptung, bestimmte Accounts würden nur retweeten und seien deshalb vermutlich Bots, lässt sich erfahrungsgemäß schon durch einen

kurzen Blick auf den „Tweets und Antworten“-Tab, in dem auch Interaktionen mit Nutzern sichtbar werden, oder durch eine spezialisierte Suche über das Twitter-Interface widerlegen, bei der man Retweets herausfiltert und so die selbst formulierten Tweets sichtbar macht.

- Oft ist die Kontaktaufnahme mit den angeblichen „Social Bots“ erfolgreich, entweder direkt über Antworten auf Twitter oder über Direktnachrichten. Auf diese Weise wurde ich beispielsweise darauf aufmerksam, dass eine im kanadischen Wahlkampf in Presse und Fernsehen als „Bot“ beschuldigte Twitter-Nutzerin sogar einen Youtube-Kanal betreibt, in dem sie selbst zu sehen ist und dort auch die entsprechenden Vorgänge um den Botvorwurf ironisch kommentierte.
- Manche der angeblichen „Social Bots“ werden sogar unter dem Klarnamen der Nutzer betrieben. Beispielsweise stellte sich ein weiterer im kanadischen Wahlkampf als „Bot“ beschuldigter Account, der durch eine zeitweise extrem hohe Zahl an Tweets und Retweets aufgefallen war, nach kurzer Recherche als authentischer Account eines akademisch gebildeten kanadischen Rentners heraus, bei dem Name und Lebenslauf im Twitterprofil sogar völlig korrekt wiedergegeben war.

Frage 5: Welche Relevanz haben „Social Bots“ in Social-Media-Diskursen?

„Social Bots“ haben keine Relevanz in Social-Media-Diskursen. Großen Schaden richtet allerdings der Irrglaube an, dass „Social Bots“ sich an solchen Diskursen beteiligen würden.

Frage 6: Können Sie aktuelle Beispiele dafür benennen, dass ein Bot in die politische Meinungsbildung eingegriffen hat? Gibt es Belege für eine Beeinflussung der politischen Debatte und/oder von Wahlen durch „Social Bots“ in Deutschland oder in anderen Ländern?

Nein.

Frage 7: Welche Maßnahmen erachten Sie für effektiv, um das Verhalten von Bots zu analysieren und angemessen zu reagieren, falls negative Auswirkungen zu erwarten sind?

Es sind keine Maßnahmen notwendig.

Frage 8: Inwiefern ist eine Kontrolle überhaupt möglich?

Das automatisierte Betreiben von Twitter-Accounts, besonders in größerer Zahl, lässt sich am einfachsten durch Twitter selbst unterbinden, etwa durch Captcha-Tests, Abgleich von IP-Adressen, automatische Test auf Dopplungen von Tweets, das Hinterlegen eindeutiger Telefonnummern bei der Account-Eröffnung etc.

Der im Vergleich zum Medium E-Mail recht geringe Anteil an kommerziellem Spam, dem man als Twitter-Nutzer begegnet zeigt, dass solche Maßnahmen von Twitter bereits mit einem gewissen Erfolg umgesetzt werden. Dennoch sind zweifellos noch zahlreiche kommerzielle Fake- und Spam-Accounts, vermutlich teils mit einem hohen Grad an Automatisierung, auf Twitter zu finden. Dazu gehören Accounts mit fiktiven Frauennamen und leichtbekleideten Frauen im Profilbild[11], die Kunden für pornographische Angebote gewinnen wollen und Accounts, die Twitter-Nutzer in betrügerische Geschäfte mit kryptographischen Währungen verwickeln wollen. Zumindest die genannten Fake-Accounts mit Frauennamen scheinen von Twitter sogar in großer Zahl toleriert zu werden. Die meisten dieser Accounts sind schon seit vielen Jahren

[11] Beispiel `@Gretchen_67897` mit folgendem Text im Profil: „*Ich möchte neue Leute kennenlernen, mir folgen, und ich werde dir meine Fotos schicken.....*“

auf Twitter angemeldet, was auch ein Indiz dafür sein könnte, dass das Neuanlegen von Fake- und Spam-Accounts vor einigen Jahren noch leichter war als heute.

Frage 9: Erachten Sie eine Regulierung von Social Bots, etwa ein Verbot von Social Bots im Rahmen von Wahlkämpfen in Deutschland / Europa für erforderlich?

Nein.

Frage 10: Was könnte eine Kennzeichnungspflicht für „Social Bots“ bewirken?

Die Idee einer Kennzeichnungspflicht von „Social Bots“ ist aus mehreren Gründen absurd:

- „Social Bots“ existieren allem Anschein nach überhaupt nicht. Insofern erscheint eine Kennzeichnungspflicht für „Social Bots“ ähnlich sinnvoll wie eine Regulierung der Haltungsbedingungen von Yetis oder vorgeschriebene Flugkorridore für UFOs.
- Selbst wenn es „Social Bots“ gäbe, denen ja per definitionem eine Täuschungsabsicht zugrunde liegen soll, wäre eine solche Kennzeichnungspflicht ebenso wenig erfolgversprechend wie eine Kennzeichnungspflicht für Taschendiebe oder für E-Mail-Spam.
- Eine Kennzeichnungspflicht für „Social Bots“ würde der für den politischen Diskurs überaus schädlichen Social-Bot-Verschwörungstheorie zusätzliche Nahrung verleihen.