

Simon Kruschinski, Pascal Jürgens, Birgit Stark, Marcus Maurer, Christian Schemer

In Search of the Known Unknowns. The Methodological Challenges in Developing a Heuristic Multi-Feature Framework for Detecting Social Bot Behavior on Facebook

Abstract

Recent political events such as the 2016 U.S. election and the UK Brexit campaign have shown that automated activities by social bot accounts are no longer a marginal phenomenon. Researchers and companies try to detect the opaque and dynamic computational propaganda efforts by programming or using detection methods which draw on arbitrary social bot features often in combination with machine learning algorithms. Furthermore, studies about social bots have concentrated almost exclusively on Twitter.

On the one hand, this reveals the existence of a research gap regarding social bot activities on Facebook which is the more relevant information intermediary for public, media, and political actors alike. On the other hand, this poses several methodological challenges for social bot detection since the used methods are dependent on the platform's contexts and the availability and maintenance of a high-quality gold standard of human-annotated social bots.

In this article, we want to shed light on the challenges of social bot detection on Twitter and derive methodological implications for researching automated behavior on Facebook. By drawing on a large-scale study of Facebook bots in the German federal election 2017, we propose an alternative, more theory-driven detection approach that focuses on text duplication as the primary social bot feature to investigate digital astroturfing campaigns which spread political content in high volume and frequency. In conjunction with large-scale data collection, stronger data sharing, and replication attempts, such analytical strategies could serve to establish long-term criteria for the identification of problematic use of platforms – whether by social bots or humans.

Keywords: social bots, computational methods, digital astroturfing, social network sites, German federal election, locality sensitive hashing.

Dieser Beitrag kann wie folgt zitiert werden:

Kruschinski, Simon, Pascal Jürgens, Birgit Stark, Marcus Maurer, Christian Schemer: In Search of the Known Unknowns. The Methodological Challenges in Developing a Heuristic Multi-Feature Framework for Detecting Social Bot Behavior on Facebook. In: Müller, Philipp; Stefan Geiß; Christian Schemer; Teresa K. Naab; Christina Peter (Hrsg.): Dynamische Prozesse der öffentlichen Kommunikation. Methodische Herausforderungen (Methoden und Forschungslogik der Kommunikationswissenschaft, Band 15). Köln [Herbert von Halem] 2019, S. 102-143

Weitere Informationen zum Buch erhalten Sie unter:

<https://www.halem-verlag.de/dynamische-prozesse-der-oeffentlichen-kommunikation/>

DOI: 10.1453/heuristic-multi-feature-framework_9783869624044.pdf