TAB OFFICE OF TECHNOLOGY ASSESSMENT AT THE GERMAN BUNDESTAG

SOCIAL BOTS – THE POTENTIAL OPINION MAKERS

TAB-Fokus no. 16 regarding Horizon-Scanning no. 3 April 2017

Summary

- Generally, social bots are used to distort the content of discussions and to influence the significance of topics or the popularity of people and products.
- They have the potential to change the political debate culture on the Internet by massively spreading (fake) news and to undermine faith in democracy by »poisoning« a society's climate.
- So far, the still recent phenomenon of social bots has been detected particularly on the platform Twitter. There is only little evidence so far with regard to their actual impact on shaping a (political) opinion.
- The technical possibilities to uncover the impacts of social bots lag behind their rapid development.

What is involved

Social media are enjoying increasing popularity as a source of information. However, messages are not only posted by real people. For some time already, so-called social bots are posting messages as well. Social bots are computer programs developed to automatically generate messages in social networks such as Facebook or Twitter in order to influence or manipulate discourses. They are able to generate meaningful texts (e.g. comments, answers or statements) that are similar to texts written by humans. It is rarely obvious that the messages have not been created by a human, but by a machine.

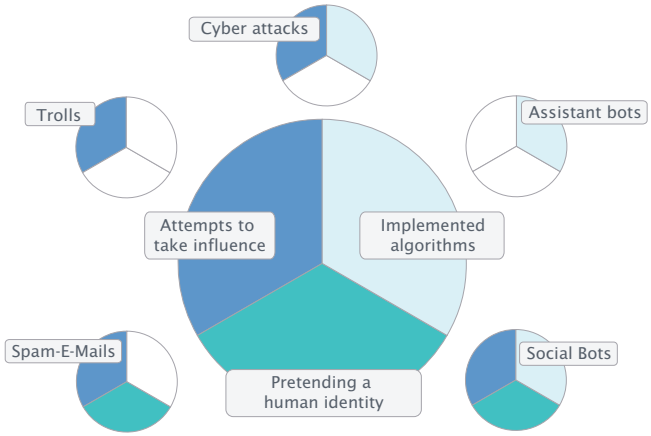
Uncovering fake accounts of social bots, i.e. fake user profiles that do not belong to an authentic person, is impeded by the fact that these fake accounts use real user names and personal information (images, links) of real users. Such fake accounts can be easily multiplied so that thousands of user accounts can be created e.g. on Twitter that will generate tens of thousands of tweets a day. It is assumed and partly proven that social bots are being used deliberately by states, companies and stakeholder groups.

Properties of social bots

Unlike other Internet phenomena such as assistant bots, junk mail, Internet trolls or cyber attacks, social bots are characterised by the combination of three key characteristics (fig. 1):

- Social bots are algorithms implemented in a software.
- They pretend to be a real person.
- Social bots try to influence people shaping an opinion.

Fig. 1 Differentiation between various Internet phenomena



Social bots differ from assistant bots (e.g. chatbots, digital virtual assistants) particularly with regard to their inten-

Client

Committee on Education, Research and Technology Assessment
+49 30 227-32861
bildungundforschung@bundestag.de

tion, even though they have similar technical basics. The purpose of assistant bots, for example, is to send automated messages such as weather forecasts or weather warnings.

Internet phenomena such as trolls (as human actors) as well as junk mail have in common with social bots that they intend to manipulate or make use of disinformation. The common feature of cyber attacks and social bots is their technical basis and, again, the intention to take influence.

Depending on their technical development level, social bots are more or less able to pretend to have a human identity. Simple social bots are able to identify keywords such as e.g. »refugees« and respond to these keywords by posting images from the Internet or retweeting comments. Some imitate the behaviour of human users by pretending a different degree of activity at different times of the day. More complex social bots are able to analyse communication contents and conduct dialogues. Currently, most of the social bots on the Internet are rather simple.

Only little programming skills are required to create a simple social bot. Manuals and instructions for this pur pose are freely available on the Internet. However, the level of difficulty increases significantly with the technical complexity of the bot to be programmed, e.g. if the bot shall carry out language analysis and simulate dialogs.

Considerable development leaps are expected for the field of bot technology in the years to come. The technological maturity of social bots will benefit from progress made in the fields of artificial intelligence, machine learning and big data. This is why, in the future, social bots will be able to show even more »human-like« behaviour and will be more difficult to detect. The technical opportunities to detect them follow this development, but – just like antivirus software – always lag a little bit behind.

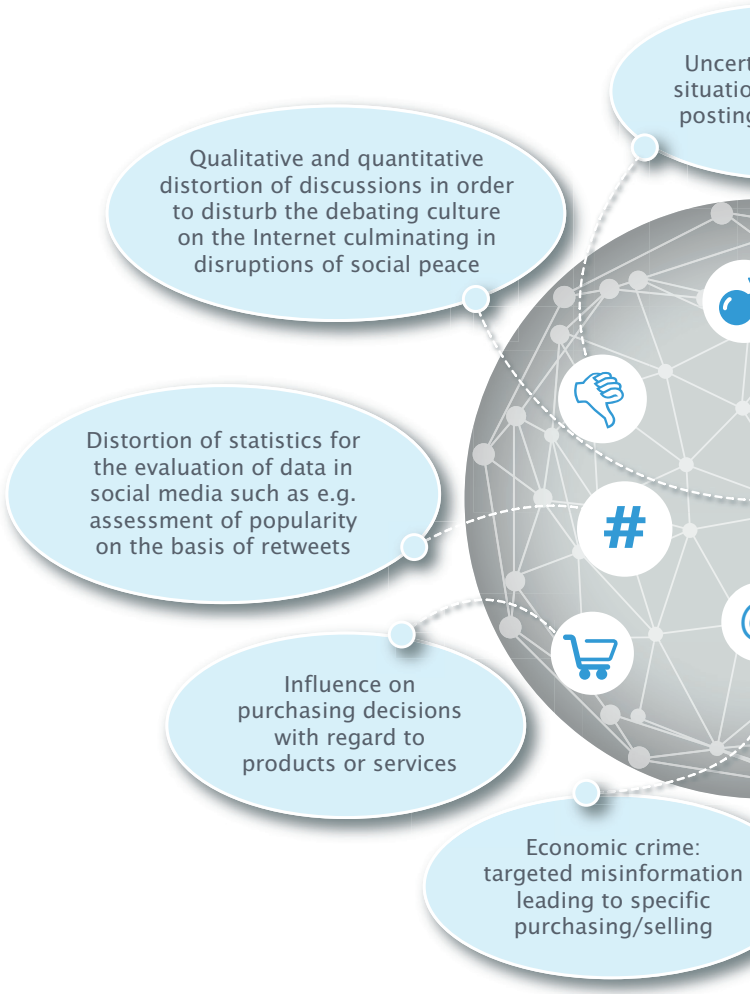
Influence and effectiveness of social bots

There is only a limited number of notable examples showing the influence of social bots that are referred to repeatedly both in the media and in scientific articles. The three examples most frequently mentioned in the press and literature are social bots that have been used during the protest movement in Ukraine, in the course of the Brexit campaign and in the US presidential election campaign 2016.

So far, social bots have been proven primarily on the platform Twitter that offers an easily accessible interface for programmers. There are no scientific studies yet proving that social bots actually succeed in influencing social groups. For this reason, there is little evidence showing the extent of their actual impact.

To date, the main fields of application for social bots are election campaigns, protests or attempts to influence political tendencies (fig. 2). In this context, social bots are used for four purposes:

Fig. 2 Potential impact areas of social bots



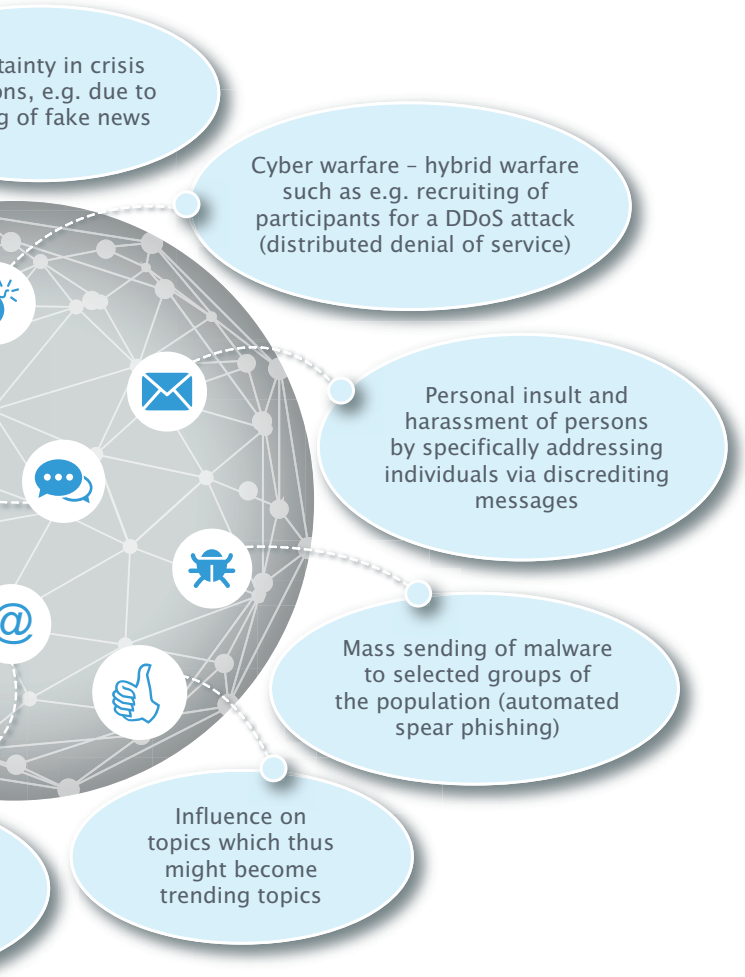
- »Muting« opposing opinions by a flood of hashtags with distracting, polarising or trivial messages,
- Disseminating propaganda and spin,
- Artificially generating high follower counts on Twitter that shall emphasise the significance of the own position,
- Discrediting or insulting people or tempting them to buy payable services on the Internet.

Social bots seem to be potentially influential in the context of political culmination if narrow majorities are at stake in political decision-making processes as it could be observed

in the election campaign between Hillary Clinton and Donald Trump or during the Brexit campaign. Television debates in election campaigns offer an opportunity to disseminate opinions during or immediately after the debates.

Future impact potential and possible applications of social bots

Experts consider the potential of social bots regarding **political processes** to be predominantly high. Social bots can



be used for disseminating news on the Internet in order to manipulate tendencies or to influence political debates and discourses. In particular, there is a potential danger if social bots disseminate masses of fake news in crisis situations such as e.g. after attacks. Thus, social bots can contribute to changing the political debate culture on the Internet and involve disinformation and a »climate poisoning« in public discourse.

Economic processes are another sphere of influence of social bots. Social bots bear the risk of influencing the customer and buying behaviour of individuals (via so-called influencer marketing) and of manipulating even entire markets such as stock exchange trading.

In view of **IT security** in the context of Industry 4.0, the Internet of Things and the associated increase of networked devices, a risk due to social bots currently still seems to be unlikely, as social bots do not attack the hardware or software of IT systems directly. Against the background of rapid developments and of devices becoming more and more intelligent on the one hand and in view of the future capabilities of (social) bots on the other hand, future risks – such as e.g. highjacking of devices for malicious purposes – are difficult to assess. Social bots can represent a risk for IT security particularly if they target humans as potentially weak points of IT security and exploit them for attacks (e.g. by sending links that will install malware).

Business models of social networks are primarily based on sales of advertising and/or user data. These models can only work with humans acting on the platform and making purchasing decisions. In the long term, social bots represent a threat for the business model of social networks. Some users might turn away from them, because they lose confidence in the authenticity of the messages. Moreover, as a consequence, investors might withdraw from the social networks.

However, the use of social bots does not necessarily have to be associated with negative intentions. Possible **positive applications** include artistic and creative applications as well as approaches using social bots as a »honeypot« or as a countermeasure (so-called counter-speech campaigns) in order to fight fake news. Moreover, they could be used for positively influencing human behaviour (nudging). In the latter case, however, this would only be ethically acceptable if the principles of informational self-determination are observed.

Options for action

The findings available so far suggest that knowledge regarding the quality and reliability of sources on the one hand and basic knowledge of IT-related contexts on the other hand are decisive for confidently dealing with propaganda or fake news. **Children, young people and even adults should be encouraged and strengthened with regard to their media literacy in terms of so-called digital literacy.** A basic understanding of IT-related functionalities and contexts – e.g. regarding the question of how messages are becoming a trend – should be imperatively included in school education. It should be clarified whether the topic of social bots can be dealt with educationally within a larger framework for »correctly« using networked/digital and social media or whether this should be addressed specifically.

Another particular target group for corresponding measures are **journalists**, because **as multipliers** they are obliged to be especially diligent with regard to choosing their sources. Established media also increasingly fall back on contents from social media and consider them as a basis with regard to assessing the relevance of topics and news. Due to the fact that they can be easily manipulated, journalists should not rely only on the indicators commonly used in social media, such as e.g. the number of retweets, but carefully check these sources for credibility and authenticity – as it is done for assessing visual material.

The existing legal framework does not offer any justification to prohibit social bots and their use for manipulative purposes. At this point in time, compulsory **labelling of bots** seems to be rather unsuitable i.a. due to difficulties with regard to a reliable detection of bots, a lack of sanctioning possibilities as well as conflicts with data privacy. Instead, **social media should commit themselves increasingly** and implement measures against the dissemination of social bots on their platforms. In case of unlawful practices committed by social bots, it should be considered to prosecute the initiators or programmers. Only in exceptional cases, it is likely to be possible to identify initiators of social bots who operate internationally or from third countries. **Self-commitment of companies and civil-society organisations** would at least make it possible to put an end to a further dissemination of social bots.

Though it is indispensable to develop **detection systems**, no definitive technical solution has been found yet. As the majority of social bots uses the short message service Twitter, which is characterised not only by a simple content-related message structure that can be easily generated even by machines, but also by an easily controllable interface (application programming interface [API]), this interface represents a potential defence mechanism against social bots. For this reason, there are considerations to implement an identification of the accessing algorithm at this interface. Thus, it would be possible to determine how the algorithm works, what are its effects etc. Such a measure would ensure that only desirable algorithms would be granted access, while undesirable algorithms could be blocked. However, there is some doubt even among experts as of whether such a mechanism can actually be effective and whether it has a chance of being implemented.

Dealing with the relatively young phenomenon of social bots illustrates that many questions still remain unanswered. However, there are only very few cases where a direct impact of social bots and their messages could be proven. Further research and **investigations are required to enable a comprehensive clarification** and evaluation of the risk potential as well as of the technical and legal challenges involved. Only with an enhanced knowledge base it is possible to answer the question whether social bots are a potential threat to democracy or just an annoying marginal phenomenon.

Horizon-Scanning no. 3
Social Bots
Sonja Kind, Tobias Jetzke, Sebastian Weide, Simone Ehrenberg-Silies, Marc Bovenschulte

Website of the project
www.tab-beim-bundestag.de/en/research/uV005.html

Project manager and contact
Dr. Sonja Kind
+49 30 310078-283
sonja.kind@vdivde-it.de

The Office of Technology Assessment at the German Bundestag (TAB) is an independent scientific institution which advises the German Bundestag and its committees on questions of scientific and technological change. TAB has been operated by the Institute for Technology Assessment and Systems Analysis (ITAS) of the Karlsruhe Institute of Technology (KIT) since 1990. It has been cooperating with the Helmholtz Centre for Environmental Research – UFZ, the IZT – Institute for Futures Studies and Technology Assessment and VDI/VDE Innovation + Technik GmbH since September 2013. The Committee for Education, Research and Technology Assessment decides on TAB's work programme, which also includes subjects proposed by other parliamentary committees. The standing »TA Rapporteur Group« consists of one member from each of the parliamentary parties: Dr. Philipp Lengsfeld (CDU/CSU), René Röspel (SPD), Ralph Lenkert (Die Linke), and Harald Ebner (Bündnis 90/Die Grünen) and the Chairwoman of the Committee, Praticia Lips (CDU/CSU).