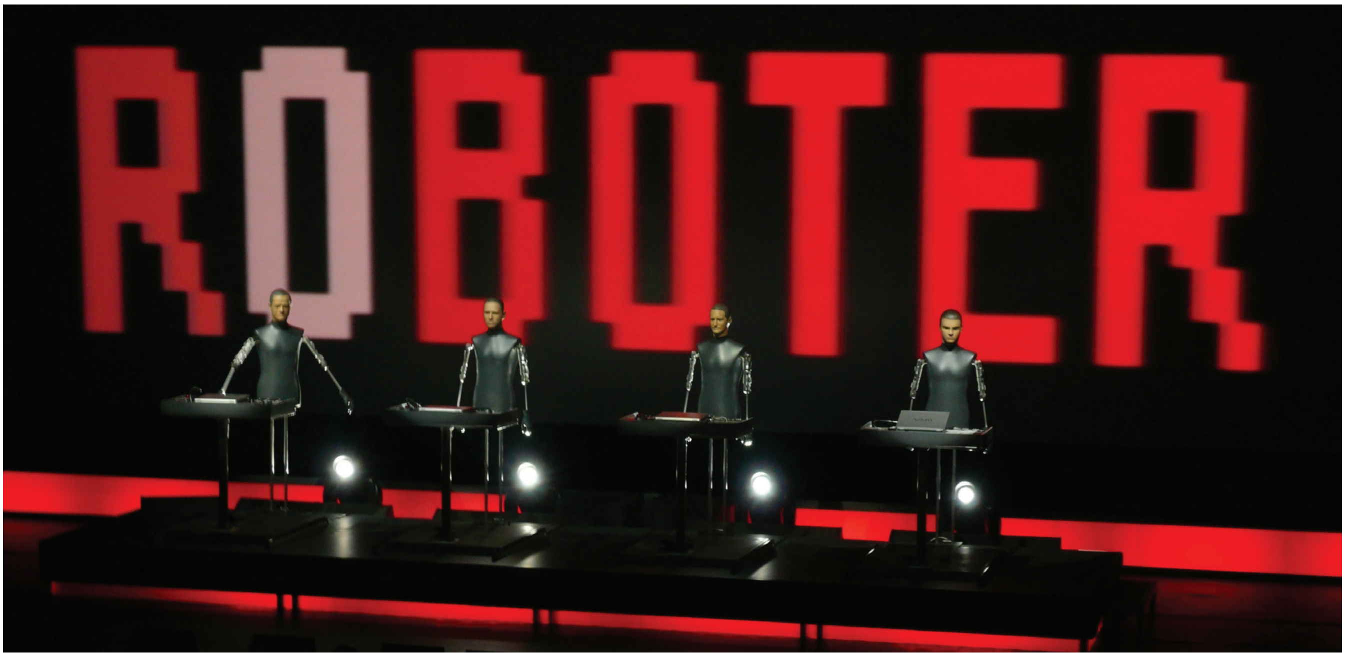
Facts & Findings

Konrad Adenauer Stiftung

SEPTEMBER 2016
NO. 221



Invasion of the social bots

Simon Hegelich

Key Points

- Social bots influence opinion. Social networks are their preferred sphere of action. They have already interfered with the opinion-forming process during debates about Brexit, in connection with events such as Russia's annexation of the Crimean Peninsula, and most recently in the election campaign involving Trump and Clinton.

- Social bots are software robots programmed by humans. They collect information and data, but also systematically introduce trends and key topics into social media without users being aware of it. The influencing potential – the so-called "bot effect" – is theoretically very large, but difficult to prove empirically.

- It is becoming ever more difficult to distinguish bots from real people. Bots operate with fake profiles, pretending to be human. They become involved in public debate on social media in a controlled manner. If bots were to proliferate on social media disproportionally, they could have a disruptive impact on existing communication platforms such as Twitter as users may no longer see any point in communicating on a platform that is mainly populated by machines.

- Political groups and parties in Germany can also utilise social bots for their own ends. The successful deployment of bots in the USA will move across to Germany as well and result in intensive use of social bots in the area of political communication.

- Transparency and measures to trigger an open debate can be instrumental in raising awareness and promoting competent use of social bots. The idea that quantity necessarily equates quality does not apply on the Internet.

CONTENTS

Preface

Be it Brexit, the Russia-Ukraine conflict or the US presidential election campaign: there are increasing instances of social bots operating on social media with the aim of influencing political debate. The following text describes what bots are, how social bots work and the threats they entail. In addition, some examples of the use of social bots for political purposes are provided. In conclusion, a prognosis is given as to what developments can be expected in this area in the near future.

What are bots?

The term "bot" as an abbreviation of robot developed as a description of programs operating autonomously on the Internet. However, this description is not entirely clear and a more detailed definition is required. The term is often applied to the scripts used by search engines such as Google to scour the Internet as well as to computers that have been infected by malicious software and subsequently lead a life of their own. But when we talk about bots these days, the term more likely refers to automated accounts in the social networks, which perform routine tasks as so-called "chat bots" or as a simple form of artificial intelligence, or which act as "social bots", disguising their true identity and pretending to users that they are real people. Over two decades ago, the scientist Roger Clark already called attention to the threats that could arise from an "active digital persona".[1]

How do social bots work?

Social bots are increasingly deployed in a political context

While bots are currently being talked up by technology companies such as Facebook, Google and IBM as a new trend, which will help to make apps and websites redundant because users can interact directly with the bot assistant, social bots are increasingly used in a political context. The aim is to influence the public or specific target groups through automatically generated contents and interactions.[2]

In terms of technology, social bots are very easy to create nowadays. Three elements are necessary: user accounts registered on the respective social network, access to an automated interface (API) of that network and a program controlling the bot accounts automatically. The registered user accounts are generally purchased on the Internet. Suppliers of fake social media accounts either generate them manually or directly in automated operation, and some offer logon data for hacked accounts as well. Depending on the quality, 1,000 fake accounts can currently be bought for between $45 (simple Twitter accounts) and $150 ("aged" Facebook accounts). The sellers generally operate from abroad (frequently Russia).

Social bots can be created using simple means

As a rule, the social networks make the APIs available free of charge to attract developers for their platform. However, there are large differences with respect to the registration process and the user-friendliness of the APIs, resulting in networks such as Twitter and Instagram being infested with far more bots than, say, Facebook – simply because it is easier to access the API there.

The software for controlling the bots is also available to purchase. A very high-quality piece of software that can be used to control 10,000 Twitter accounts costs around $500. Bots can also be programmed easily on the basis of existing software libraries (a very basic bot requires a mere 15 lines of code). There are bots of greatly differing types. In the simplest case, the bot's autonomous action is limited to sending ready-made messages. But there are also bots (albeit rarer) that are capable of interacting with real users and generating new texts independently. As normal communication on the social networks is generally not particularly complex, even primitive bots frequently don't arouse suspicion. A typical bot on Twitter, for instance, could independently generate messages based on preselected websites, automatically follow other users and send ready-made propaganda messages, embellished with key terms and hashtags that happen to be popular at the time, either in response to "the click of a button" or according to a random schedule. Where technology is concerned, one needs to bear in mind that these bots are fundamentally freely scalable. Anybody who has a program suitable for controlling one bot can use it to control a whole army of bots.

Once programmed, one bot can be expanded into a whole army

Threats posed by social bots

The most significant threat currently arises from the sheer volume of messages that can be disseminated via a botnet. Bots manipulate trends in social networks and these trends affect political and economic decision-making processes. Under the buzzword "Big Data", more and more companies in diverse sectors put great stock in analysing user behaviour in social networks to obtain insights into how well their own brand is doing as well as into the behaviour of potential customers or to uncover social trends. Such analyses are also already being used in the political sphere.[3] While there is still considerable caution in evidence in Germany in this area,[4] political social media analysis has already developed into a significant market internationally, where actors such as Civics Analytics (currently active on behalf of Hillary Clinton) and Cambridge Analytica (engaged by Donald Trump) operate. But if trends are manipulated by bots on a grand scale and bots muscle in on all important debates (see next chapter), these analyses are just inaccurate at best. At worst, they may induce politicians to pander to such trends in their statements or even in their policies with the result that the position pushed by the bots may potentially receive a level of support that the bots by themselves could never have achieved.

Bots manipulate trends on a grand scale

Secondly, there is a risk of the bots influencing the opinions of specific groups. One can probably safely assume that the bot posts do not effect the manipulation directly. All studies indicate that people don't change their political conviction simply because they see messages on social media. But a more subtle type of manipulation is very probably at play. If bots are, for instance, used to disseminate extreme contents in a discussion context on a large scale (e.g. in a Facebook group or under a topic-specific hashtag), this will generally result in people with moderate views withdrawing from the particular discussion. People who have a position that is radically opposed to the bot messages will feel challenged to actively oppose the content, which in turn will bring those who share the opinion promoted by the bots onto the barricades. This creates a heated debating atmosphere where people who are fundamentally inclined towards radical positions feel encouraged.

Social bots can create a heated debating atmosphere

Thirdly, bots can also be used for specific purposes in a cyberwar scenario. The strategies range from the infiltration of social networks for spying on users to the purposeful dissemination of misinformation (e.g. in crisis situations) to cyberattacks through the dissemination of malicious software or the organisation of so-called DDoS attacks[5].

Bots in the wild – the "bot effect"

The following examples demonstrate that the above-mentioned strategies are already being used on a large scale. An analysis of these bot deployments also illustrates how they can potentially be countered and which risks are, in fact, realistic.

30 per cent of the Twitter followers of both candidates in the US presidential election campaign aren't human

In the current US presidential election campaign, one can assume that social bots constitute a substantial proportion of the candidates' followers. The online magazine "vocativ" reports that the proportion of real Twitter followers is around 60 per cent, both for the Democrat candidate Hillary Clinton and the Republican candidate Donald Trump. And the proportion of Trump's fake followers has apparently risen strongly compared to analysis figures from the summer of 2015.[6]

During the 2012 US presidential election campaign, there had already been a sudden increase in the number of the followers of the then challenger in evidence, which was found to be due to the use of fake followers.[7] Huge numbers of fake users have also been identified in connection with the political parties in Switzerland.[8] In the run-up to the Italian parliamentary elections in 2013, the Twitter followers of one candidate were analysed using a "bot detection algorithm" and it was found that over half were fake followers.[9] In its current ranking of the Twitter accounts of heads of state and political leaders, the "Twiplomacy Study" places the Venezuelan President second in terms of the number of posts on Twitter and third in the retweet statistics.[10] The noticeable thing here is that his tweets are favorited much less frequently, which indicates that some of his followers are fake users.

"likes" and "shares" are the key figures targeted in trend manipulations

This simplest method to manipulate social networks, where bots produce pure volume while not generating new content, may initially appear to be a relatively harmless form of manipulation, but its consequences are not insignificant.[11] Added to this is the fact that social networks are controlled via algorithms that give preference to popular contents. Accounts that have a large following are treated more favourably by the social networks and consequently reach more genuine users. One common method of manipulating trends on the social networks is to purposefully target the key figures that are purely quantitative, such as "likes" and "shares" on Facebook and the frequency of hashtags on Twitter.

In the course of the Brexit debate, scientists found out that a very large number of the tweets with hashtag "#Brexit" originated from bots.[12] Hashtags linked to the Remain campaign (such as "#StongerIn") were used much less frequently by the bots. This example also shows, however, that it is easy to overestimate the risk bots pose. Theoretically, one might have gained the impression that the Brexit campaign was clearly ahead in terms of popular support. This could potentially have caused Remain supporters not to cast their vote because the outcome appeared to be predetermined. However, as we know, the general feeling was that there would be a (probably close) victory for Remain. Nor did the bots appear to have a noticeable effect on voter behaviour. In the UK, Twitter is used almost exclusively by well-educated younger people. But it was precisely this demographic group that voted against Brexit. In addition, the analysed figures indicate that a large proportion of the bots used both pro-Brexit and Remain hashtags. This was probably because many of them were not actually political bots but simply advertising spam where those hashtags are used that happen to be on trend at the time. This example shows that even large-scale attempts to influence trends using bots do not necessarily equate to effective manipulation.

A considerably more complex botnet has been uncovered in the context of the Ukraine conflict.[13] This botnet involves some 15,000 Twitter accounts on which an average 60,000 messages a day are posted. The contents of the tweets are chosen to match the presumed interests of young Ukrainian men. The bots talk a great deal about football, tell sexist jokes and disseminate links for the illegal downloading of current American movies. However, in between these tweets, propaganda messages of the "Right Sector" – an ultranationalist Ukrainian confederation with a paramilitary wing – are disseminated systematically. There are different manipulation strategies in evidence here. For one, this is also about distorting trends by popularising certain hashtags. But in addition to that, the bots also purposefully link buzzwords such as "Maidan" and "Euromaidan" to the "Right Sector" hashtag to induce Twitter's algorithms to offer users who are searching for "Maidan" "Right Sector" contents as well. Another strategy that has come to the fore in this context is the dissemination of misinformation. The botnet thus spread the message that the separatists had obtained missiles from Russia and would now aim them at Kiev. In addition, the bots systematically follow Ukrainian politicians to expand their own reach. This is effective because even if the politicians do not get taken in by the bots and pass on their messages knowingly or inadvertently, Twitter will be more likely to present the bots' tweets to other users who follow the same politicians. The Ukrainian bots also have an entire arsenal of tricks at their disposal for evading classic bot detection algorithms. They follow each other and consequently have a balanced ratio of friends and followers, they follow a schedule in their posting of tweets that simulates break and sleeping periods yet appears to be random, and they are capable of making slight modifications to the tweets so that the message remains identical, but automatic programs will not recognise the texts as identical copies.

Tricks used by bots: simulation of sleeping periods, balanced ratio of friends and followers

There are reports emerging from Russia about the completion of a major project to set up an infrastructure for the manipulation of social media. After the so-called "troll factory" in St. Petersburg that has been producing all manner of opinions for quite some time, social bots are the logical next thing. Particularly in connection with the war in Ukraine, Russophile comments were clearly predominating on the German-language Internet, contrary to figures from surveys and the opinion of the journalists focusing on the subject and political representatives.

There is a botnet concerned with Donald Trump operating, which is technically more primitive, but very ambitious in terms of the strategic objectives. The bots, exclusively purporting to be good-looking young women and men, are specialised in disseminating jokes. Many of the tweets, however, involve blatant racism or anti-Semitism. They are interspersed with derogatory tweets about Donald Trump. The makers of this botnet probably assume that Trump supporters will not even realise that their candidate is being insulted on the Internet. The racist jokes are, in fact, intended to penetrate the filter bubble so that the discrediting messages can take effect. The fact that this will allow racist propaganda of the worst kind to permeate the Internet as collateral damage does not seem to greatly bother the botnet makers.

Racist jokes used to penetrate the filter bubble

However, bots need not be limited to disseminating messages. Their impact can go much further and extend into the area of cyber warfare. This involves so-called social engineering strategies, which aim at influencing users by means of psychological tricks, such as suggestion, to achieve the desired effect.

At the “Black Hat” hacker conference, there was a presentation on a concept involving the use of artificial intelligence to infect individual users with malicious software via bots.[14] The hackers developed a program that can generate a virtual perfect social media friend for any user. The program analyses the user’s posts and then attempts to automatically generate posts that may be of great interest to that user. A link connecting to a website with malicious software is then incorporated into these posts. In tests, half the subjects did click on that link. This method represents a perfidious subcategory of so-called phishing. The standard method simply involves the dissemination of posts on a massive scale in the hope that the subject will pique many people’s interest (“You’re a winner!”, “Hello, remember me?”, etc.). However, not even 5 per cent of users generally click on links incorporated into such messages. In the case of what is referred to as spear phishing, the messages are modified to match the individual users, making use of information collected from the social networks. The possibility of this being automated by means of software and combined with a botnet, referred to as Automated Spear Phishing, means that in principle everybody can be targeted by a bot with customised messages.

Fake follower accounts now also figure in foreign policy strategies

One must also assume that the massive use of fake follower accounts controlled by algorithm and the associated managed dissemination of contents on social networks has become a strategy used in foreign politics. There have been reports since 2011 about the U.S. Air Force having developed something called “persona management”.[15] This involves software that allows social bots to be generated rapidly en masse, disguised in a way to enable them to infiltrate terror cells on social networks. But this software is apparently also capable of performing other types of tasks.[16]

What does the future hold?

Social bots are here to stay. While the methods to detect bots are improving all the time, the same applies to the bots themselves. It is the case for both sides that new methods can be analysed relatively quickly and suitable countermeasures are developed. Overall, the proportion of bots among social network users will probably even out at a relatively high level.

That said, huge distortions can take place in the meantime, particularly if bot activities suddenly increase in volume or if there is a new quality surge in bot technology. The former is of particular concern with respect to specific events, such as elections or crisis situations. In these scenarios, bots may actually have a large short-term impact because the manipulation is not uncovered until the event itself is in the past. And there are already clear indications of a quality surge in bot technology. An increasing number of excellent development environments for areas of artificial intelligence involving the understanding and generation of text (natural language processing, natural language generation) are currently being made available for free because corporations such as Google, Facebook and IBM hope that this will produce significant developmental advances for their own technology. Equipped with these tools, bot developers are now working on a new generation of bots that normal users will find impossible to spot.

At the same time, economic interest in the use of bots is also resulting in a radical change of the rules: bots are virtually being legalised. They are no longer seen as a manipulative threat, but rather as helpful assistants in everyday life. This frequently entails bots being identified as such (for instance in the Slack and Telegram networks). Although the rule that bots need to identify themselves can be bypassed, the incentive to do so will probably be much smaller in future than is the case currently, with bots acting automatically in a grey zone. The legal proliferation of bots will probably also change awareness among users. Anybody interacting with bots in everyday life will no longer be surprised by them and people will be more likely to wonder if a message is being sent by a human or a machine. Generally speaking, development in social media is so rapid and disruptive in many areas that we all have to learn to deal with this tool afresh all the time. Bots represent one example illustrating that digitisation invalidates a fundamental truth that has applied virtually universally to date: quantity is ultimately an indication of quality. That no longer applies today as a message that is disseminated by millions of posts can definitely be untrue.

A new principle applies on the Internet: quantity is not an indication of quality

The important thing now is for both users and political actors to become sensitised to this method of communication and to the associated agenda setting and to find adequate means to deal with them. The greatest challenge will lie in analysing how users can recognise bots more easily and promoting digital media competence in socio-political circles.

[1] *Clarke, Roger 1994. The Digital Persona and its Application to Data Surveillance. In: The Information Society, pp. 77-92, here p. 68f.: "In extreme cases, an active agent may be capable of autonomous behavior. It may be unrecallable by its originators (as was the case with the Cornell worm). It may, by accident or design, very difficult to trace to its originator."*

[2] *Boshmaf, Yzan, Ildar Muslukhof, Konstantin Beznosov, Matei Pipeanu 2013. Design and analysis of a social botnet. In: Computer Networks 57, pp. 556-578, here p. 556: "This is achieved by either simply mimicking the actions of a real OSN [abbreviation of Online Social Network, author's comment] user or by simulating such a user using artificial intelligence, just as in social robotics."*

[3] *Hegelich, Simon/Shahrezaye, Morteza 2015. The Communication Behavior of German MPs on Twitter. Preaching to the Converted and Attacking Opponents. In: European Policy Analysis 1.*

[4] *Jungherr, Andreas/ Schoen, Harald/ Güldenzopf, Ralf et al. 2016. Twitter als politische Informationsquelle. In: Konrad-Adenauer-Stiftung – Schlaglicht, www.kas.de/politischekommunikation.*

[5] *"distributed denial of service"*

[6] *Beckler, Ryan 2015. Which Presidential Candidates Have The Most Fake Twitter Followers? Four months ago, Donald Trump had the best real-to-fake Twitter follower ratio. What happened?. Online: http://www.vocativ.com/239402/which-presidential-candidates-have-the-most-fake-twitter-followers/. Accessed: 05 Jul 2016: "Most interestingly though, is the fact that the 'real' proportion of Donald Trump's 4.43 million followers plummeted from 90 percent to 61 percent in just four months."*

[7] *Considine, Austin 2012. Buying Their Way to Twitter Fame. Online: http://www.nytimes.com/2012/08/23/fashion/twitter-followers-for-sale.html. Accessed: 05 Jul 2016.*

[8] *Schuppisser, Raffael 2016. Die Märchen der Roboter. Auf Facebook und twitter verbreiten Maschinen Hassbotschaften und Falschmeldungen. Wir werden von ihnen manipuliert, ohne es zu merken. Denn Roboter tarnen sich als Menschen – auch in den US-Wahlen. In: Schweiz am Sonntag, pp. 46-47.*

[9] *Squires, Nick 2012. Human or 'bot'? Doubts over Italian comic Beppe Grillo's Twitter followers. A bearded comic who has been hailed as a powerful new force in Italian politics faces claims that more than half his followers on Twitter simply do not exist. Online: http://www.telegraph.co.uk/technology/twitter/9421072/Human-or-bot-Doubts-over-Italian-comic-Beppe-Grillos-Twitter-followers.html. Accessed: 05 Jul 2016.*

[10] *Lüfkens, Matthias 2015. Twiplomacy Study 2015. Accessed: 05 Jul 2016.*

[11] *Cresci, Stefano/ Di Pietro, Roberto/ Petrocchi, Martinella/ Spognardi, Angelo/ Tesconi, Maurizio 2015. Fame for sale: efficient detection of fake Twitter followers. In: Decision Support Systems 80, pp. 56-71, here p. 58: "At a first glance, acquiring fake followers could seem a practice limited to foster one's vanity—a maybe questionable, but harmless practice. However, artificially inflating the number of followers can also be finalized to make an account more trustworthy and influential, in order to stand from the crowd and to attract other genuine followers. Recently, banks and financial institutions in U.S. have started to analyze Twitter and Facebook accounts of loan applicants, before actually granting the loan. Thus, to have a "popular" profile can definitely help to augment the creditworthiness of the applicant."*

[12] *Howard, Philip/ Bence, Kollanyi 2016. Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum. In: Comprop, Research Note 2016 1.*

[13] *13| Hegelich, Simon 2015. Are social Bots on Twitter Political Actors? Empirical Evidence from a Ukrainian/ Russian Social Botnet. MPSA annual conference 2015; Hegelich, Simon 2016. Decision Trees and Random Forests. Machine Learning Techniques to Classify Rare Events. In: European Policy Analysis 2.*

[14] *Seymour, John and Tully, Philip 2016. Weaponizing Data Science for Social Engineering: Automated E2E Spear Phishing on Twitter. Online: https://www.blackhat.com/docs/us-16/materials/us-16-Seymour-Tully-Weaponizing-Data-Science-For-Social-Engineering-Automated-E2E-Spear-Phishing-On-Twitter.pdf. Accessed: 30 Aug 2016.*

[15] *Gehl, Robert W. 2013. The Computerized Socialbot Turing Test. New Technologies of Noopower. In: SSRN Electronic Journal.; Webster, Stephen 2011. Exclusive: Military's 'persona' software cost millions, used for 'classified social media activities'. Online: http://www.rawstory.com/2011/02/exclusive-militarys-persona-software-cost-millions-used-for-classified-social-media-activities/. Accessed: 05 Jul 2016.*

[16] *Finger, Lutz/ Dutta, Soumitra 2014. Ask, measure, learn. Using social media analytics to understand and influence customer behavior. Beijing: O'Reilly, here p. 173: "While the planned applications for this software are classified, such tools would enable virtual people to be placed strategically in locations around the world, to influence the public opinion in ways that would benefit the US government.".*

BIBLIOGRAPHY

- *Barberá, Pablo. 2015. Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data. Political Analysis 23: pp. 76–91. DOI: 10.1093/pan/mpu011.*

- *Hegelich, Simon. 2016b. Social Botnets auf Twitter - Der Fall Ukraine. In Media Bias im Internet // Media Bias im Internet - Tendenzfreiheit und Vielfalt von Medien(inhalten). Tendenzfreiheit und Vielfalt von Medien(inhalten) // Lecture event organised jointly by the Institut für Rundfunkrecht and the Institut für Rundfunkökonomie of the University of Cologne on 19 June 2015, Institut für Rundfunkrecht of the University of Cologne and Roland Bornemann (eds.), pp. 127–136. Munich: C.H. Beck; Verlag C. H. Beck*

Author

Prof. Dr. Simon Hegelich, Bavarian School of Public Policy (HfP) at the Technical University of Munich

*Professor Hegelich (*1976) combines political science and computer science in his research field of Political Data Science. This covers topics relating to digitisation, including the study of its political relevance, as well as classic questions of political science, using methods such as machine learning, data mining, computer vision and simulation. Simon Hegelich studied political science at the University of Münster, where he also gained his PhD and post-doctoral lecturing qualification. From 2011 to 2016, he headed the interdisciplinary research centre FoKoS at the University of Siegen. In 2016, Simon Hegelich took up the post of Professor for Political Data Science at the Bavarian School of Public Policy (HfP) at the Technical University of Munich.*

Cover page image:
Ronald Preuß (https://commons.wikimedia.org/wiki/File:Wir_sind_die_Roboter.jpg), „Wir sind die Roboter", https://creativecommons.org/licenses/by-sa/2.0/legalcode

Konrad-Adenauer-Stiftung e.V.

Contact:
Daphne Wolter
Coordinator Media Policy
Department Politics and Consulting
Phone: +49(0)30/26996-3607
Email: daphne.wolter@kas.de

Postanschrift: Konrad-Adenauer-Stiftung, 10907 Berlin

ISBN 978-3-95721-246-7

www.kas.de