(Un)Trendy Japan: Twitter bots and the 2017 Japanese general election[1]

JOZEF MICHAL MINTAL AND RÓBERT VANCEL

sciendo

Politics in Central Europe (ISSN: 1801-3422)
Vol. 15, No. 3
DOI: 10.2478/pce-2019-0027

Abstract: *Social networking services (SNSs) can significantly impact public life during important political events. Thus, it comes as no surprise that different political actors try to exploit these online platforms for their benefit. Bots constitute a popular tool on SNSs that appears to be able to shape public opinion and disrupt political processes. However, the role of bots during political events in a non-Western context remains largely under-studied. This article addresses the question of the involvement of Twitter bots during electoral campaigns in Japan. In our study, we collected Twitter data over a fourteen-day period in October 2017 using a set of hashtags related to the 2017 Japanese general election. Our dataset includes 905,215 tweets, 665,400 of which were unique tweets. Using a supervised machine learning approach, we first built a custom ensemble classification model for bot detection based on user profile features, with an area under curve (AUC) for the test set of 0.998. Second, in applying our model, we estimate that the impact of Twitter bots in Japan was minor overall. In comparison with similar studies conducted during elections in the US and the UK, the deployment of Twitter bots involved in the 2017 Japanese general election seems to be significantly lower. Finally, given our results on the level of bots on Twitter during the 2017 Japanese general election, we provide various possible explanations for their underuse within a broader socio-political context.*

Keywords: ***bots, Twitter, Japan, general election, bot detection***

1 The research was approved by the Matej Bel University (MBU) Ethics Committee (Reference no.: 1113-2017-FF). All methods were performed in accordance with MBU ethics guidance and regulations. Support for this research was provided by IBM under an IBM Faculty Award, and by the VEGA grant project n. 1/0433/18. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors, and do not necessarily reflect the views of the funders. The authors thank Karol Fabián for valuable insights, Stefano Cresci and colleagues for providing the training dataset for the bot detection model. We are also grateful to Jaroslav Ušiak, Tetiana Kyselova, Lucia Husenicová, Peter Terem, Ján Farkaš, Michal Farkaš, Vladimír Müller, and Bram Meens for their feedback and valuable suggestions.

Introduction

Social networking services (SNSs), such as Facebook or Twitter, represent an important new area for civic engagement and political participation (Bode – Varga – Borah – Shah 2013). SNSs and social media in general have been shown to influence political self-expression, information seeking, and real-world voting behaviour (Bond et al. 2012). Thus, it comes as no surprise that SNSs represent an attractive platform for actors who aim to abuse their functionalities (Shirky 2011). A popular tool used on SNSs to manipulate public opinion or disrupt political communication is bots (Woolley 2016), as seen in the Syrian civil war (Abokhodair – Yoo – McDonald 2015), the UK Brexit referendum (Howard – Kollannyi 2016), the 2016 US presidential election (Bessi – Ferrara 2016; Kollanyi – Howard – Woolley 2016), or the 2017 French presidential election (Ferrara 2017).

Although various studies have been published on the role of automation in SNSs during political events in Western democracies, the role of bots during political events in non-Western contexts—in countries such as Japan—remains largely under-studied. The first and, at the time of the writing this paper, only research we are aware of that has touched on the topic of bots during political events in Japan was conducted by Schäfer et al. (2017). However, in their paper, the focus lies mainly on right-wing internet activism facilitated by political bots in Japan. Furthermore, the authors did not elaborate further on the volume of bots used during political events in Japan. Thus, this is—to our knowledge—the first study that empirically demonstrates that bots are not often used in electoral processes in Japan compared to those in Western societies. More importantly, this article also explores the reasons for the underuse of bots within a wider socio-political context of Japan. To achieve the above aim, the presented article first offers background information on the 2017 Japanese general election. Then, it introduces methods and results of the empirical study. Finally, based on the findings of the study, we explore in the article the reasons behind the underuse of bots in the 2017 elections.

The backstory of the election

The Japanese Prime minister Shinzo Abe called a snap election in late September 2017, with the *dai-yonjūhachikai Shūgiin giin sōsenkyo* (48th general election of members of the House of Representatives) election day set for October 22, 2017. The official reasons given for the snap election were the search for public support for Abe's announced policies, such as the constitutional revision, and Abenomics[2]. According to Abe (2017a), these policies address the most severe

2 Abenomics—a comprehensive policy package of PM Shinzo Abe and his government unveiled after he became PM in late 2012. A goal of the policy is to revive the Japanese economy from two decades of deflation, all while maintaining fiscal discipline.

threats facing Japan—the ageing population and the problem of North Korea. However, the hastily called election, together with a relatively short campaign period, given the political situation in Japan at that time, have put the reasons behind the snap election into a peculiar light. Other alternative explanations emerged; among other things, it was mentioned that by calling a snap election, Abe tried to capitalize on the internal turmoil in the greatest opposition party, *Minshintō* (Democratic Party). The internal crisis in this party led to a split and the creation of two new parties, with one of the parties, led by the populist politician Yuriko Koike, being created only a few hours before the snap election announcement.

With the election, Abe might have tried to utilize a trend of increasing public support for his government, mostly in the context of an escalated North Korean threat. This surge of public support came after months of low voter preferences for Abe due to an unfolding political scandal. The Prime Minister and his wife were under accusations of corruption linked with Moritomo Gakuen, a private school operator. In addition to the possibly speculative sale of land in Toyonaka for 14% of the estimated value, Moritomo Gakuen also sparked controversy by implementing a nationalistic curriculum and a daily recitation of the Imperial Rescript on Education from 1890 (Repeta 2017). Nationalism and pre-war sentiment are considered sensitive issues in Japanese society, formed by decades of following the Yoshida Doctrine and the policy of antimilitarism (Berger 2003). This was not the first scandal affecting the Prime Minister's credibility. Another scandal emerged, with the government even being alleged of misleading Parliament as well as the public, over the dangerous situations into which the members of the Japan Self-Defence Forces were placed while taking part in the UN peacekeeping mission in South Sudan.[3] Despite the scandals of Abe's administration, it seems that the crisis on the Korean peninsula helped Abe, with his stable and hard-line attitude, to recover from the previous slump.

The newly created political parties *Kibō no Tō* (Party of Hope) and *Rikken Minshutō* (Constitutional Democratic Party), led by Yuriko Koike and Edano Yukio, respectively, were considered to be the most important challengers of the *Jiyū-Minshutō* (Liberal Democratic Party). Initial estimates even indicated a close competition between *Jiyū-Minshutō* and *Kibō no Tō* (Asahi Shimbun 2017). Although the public support for the *Kibō no Tō* peaked approximately 13% in the opinion polls[4], the party ultimately gained 17% in the elections. Despite the initial lag, Edano Yukio and the *Rikken Minshutō* were able to win 55 out of 465 seats in the Lower House. This result positioned them as the strongest opposition party. *Kibō no Tō* won 50 seats, lagging behind initial expectations.

3 In this context, it is necessary to emphasize strict conditions under which Japan Self-Defence Forces may engage in overseas missions (Calder 2018).

4 According to a nationwide telephone survey conducted on Sept. 26–27, 2017 (Asahi Shimbun 2017).

The snap election resulted in a landslide victory for the coalition of Shinzo Abe's *Jiyū-Minshutō* and *Kōmeitō;* he managed to secure a two-thirds majority, with 313 seats in the Lower House of the National Diet.

Our approach

To study bots on SNSs during the 2017 Japanese general election, we chose to conduct our research on Twitter, as it is the most widely used social networking service in Japan (Nussey – Ingram 2018). We base our research on collecting and analysing vast amounts of Twitter data related to the 2017 Japanese general election. In doing so, we build on research practices used to collect Twitter data in other countries during political events (Brachten et al. 2017; Bessi – Ferrara 2016). In this paper, we adopt a wider definition of bots as automated SNS accounts (Ferrara – Varol – Davis – Menczer – Flammini 2016; Brachten – Stieglitz – Hofeditz – Kloppenborg – Reimann 2017). To study bots, it is fundamental first to detect bots. Various approaches have been proposed in the academic literature to detect bots, ranging from defining a high level of automated accounts as accounts producing at least 50 tweets a day using pre-defined hashtags (Howard – Kollannyi 2016), corpus-linguistics approaches of detection (Schäfer – Evert – Heinrich 2017), to various approaches using machine learning (Yang – Harkreader – Gu 2013; Subrahmanian et al. 2016). Among the methods of bot detection on a mass scale, those utilizing machine learning are thought to be the most precise/cost-efficient. However, available methods using machine learning are usually volume-limited (Davis – Varol – Ferrara – Flammini – Menczer 2016) or publicly unavailable (Subrahmanian et al. 2016) for general users. This issue led us to build a custom model, which is, in terms of area under curve (AUC) for the test set, comparable to leading bot detection models based on user profile features. Our approach of detecting bots involved in the 2017 Japanese general election-related discussions on Twitter is based on using supervised machine learning algorithms to produce a custom ensemble classification model. Following the results on the volume of bots involved in the election-related discussion on Twitter, we provide in the discussion section of this article various possible explanations of our findings, which we base within a broader context of Japanese political behaviour, demography and foreign affairs.

Methods

Data collection

For the data collection of tweets, we chose a topic-based sampling method. The approach is based on the collection of tweets containing pre-identified hashtags[i] (Gerlitz – Rieder 2013). Hashtag identification was performed using a hybrid-

-snowball sampling method. The starting hashtags were manually selected based on trending election-related hashtags in Japan during the 2017 election. During the data collection period, we also monitored trending hashtags and the most retweeted posts that were based in Japan as well as written in Japanese. Hashtag identification and monitoring was performed using the Trendsmap Analytics platform (Trendsmap Pty Ltd, Chelsea, VIC, and Australia).

We collected the Twitter data between October 10, 2017, and October 23, 2017, using the Twitter Search API utilizing multiple credentials, and data collection was set to continuous auto-updating. We chose to use the Twitter Search API over the Twitter Stream API, as advised by Bessi and Ferrara (2016). After the collection period, we merged the individual datasets into one containing 905,215 tweets. We checked the data for incompleteness, incorrectness, inaccuracies, and irrelevance. Duplicate record detection by the *id_str* of tweets resulted in the identification of 239,815 duplicates that were removed from the dataset. The final dataset thus contained 665,400 unique tweets. From this dataset, we extracted a list of 96,917 unique Twitter accounts. The data were cleaned and analysed using IBM Watson Analytics and IBM SPSS Modeler v18.0 (IBM, Armonk, NY, USA). To obtain data on the identified Twitter accounts, we used a "Get a User's Info on Twitter" script on Blockspring (Grafly Inc., San Francisco, CA, USA). We were able to collect 126 data fields on 94,451 Twitter accounts. On the remaining 2,466 accounts, we obtained an error code, indicating that the accounts were either suspended or deleted.

Building the MVBDM bot detection model

For the bot detection model, we used a supervised machine learning approach to build an ensemble classification model. As the training data for our model, we used five separate datasets of annotated Twitter accounts provided to us under licence from Stefano Cresci and colleagues (Cresci – Pietro – Petrocchi – Spognardi – Tesconi 2017; Yang et al. 2013). The datasets used were (1) the *genuine accounts* dataset, containing 3,474 genuine Twitter users; (2) the *social spambots #1* dataset, containing 991 Twitter bots; (3) the *social spambots #2* dataset, containing 3,457 Twitter bots; (4) the *social spambots #3* dataset, containing 464 Twitter bots; and (5) the *traditional spambots #1*, containing 1,000 Twitter bots. We merged the five datasets into one final dataset containing 9,386 Twitter accounts and 53 data fields in total. Field storage types and data types for the data fields were set to be automatically determined. As the target data field for prediction, we chose the *bot* variable. We then checked the data for incompleteness, incorrectness, inaccuracies, and irrelevance. This resulted in the omission of 17 data fields due to their irrelevance as predictors. We partitioned the cleaned data into a training set and a testing set with a ratio of 80:20. In the training set, we executed an analysis of possible modelling methods. As the target data field was binary, nine modelling methods were identified to be suitable. For building the ensemble model, namely, the MVBDM bot detection model,[ii] we used the three models with the highest overall accuracy based on the testing set, with the other models set to be discarded. Confidence-weighted voting for the whole ensemble model was set to the highest level of confidence. The final ensemble model was then analysed for overall accuracy. Data preparation and model building were performed using IBM SPSS Modeler v18.0 (IBM, Armonk, NY, USA).

Application of the MVBDM bot detection model

We prepared the dataset containing the collected Twitter accounts data. The data were checked for incompleteness, incorrectness, and inaccuracies. We then applied our MVBDM bot detection model, which classified the accounts as either 0 or 1, with 0 meaning that it did not forecast the account to be a bot and 1 meaning that it did forecast the account to be a bot. Every forecast was also assigned a confidence coefficient *($XFC-bot*), representing the level of confidence the model had in its prediction. We then performed a basic descriptive statistical analysis to determine the volume of tweets in our dataset that were generated by accounts classified by our model as bots. We also performed an analysis of tweets generated by the 2,466 accounts in our dataset that were suspended or deleted and thus could not be categorized by our model.

Results

MVBDM bot detection model

The models with the highest overall accuracy based on the testing set were models using the C.5, CHAID, and decision list.[iii] The C.5 model worked by splitting the sample based on the field that provides the maximum information gain. The accuracy of this approach yielded 99.053% overall accuracy for the training set and 98.302% for the testing set. The area under the curve (AUC) for bot classification *($XF-bot* variable) was 0.992 for the training set and 0.986 for the testing set. AUC was calculated as the area under a receiver operator characteristic (ROC) curve. The second model, the CHAID model, worked by using chi-square statistics to identify optimal splits. The overall accuracy of this model was 98.44% (training set) and 97.454% (testing set). The AUC (training set) was 0.988, and the AUC (testing set) was 0.993. The last model, namely, the decision list, identified subgroups or segments that show a higher or lower likelihood of a given binary outcome relative to the overall sample. The overall accuracy of this model was 94.214% (training set) and 93.369% (testing set). The AUC (training set) was 0.955, and the AUC (testing set) was 0.945. These three models were used together to produce the ensemble model. The accuracy analysis of the ensemble model showed that the prediction accuracy for the training set was 99%, and for the testing set, it was 98.14% (Table 1.). The area under the curve (AUC) for bot classification (*$XF-bot* variable) was 0.999 for the training set and 0.998 for the testing set. The Gini coefficient evaluation yielded a result of 0.997 for the training set and 0.996 for the testing set. Gini was calculated as Gini=2AUC-1. The confidence value report for bot classification *($XF-bot* variable) in the testing set is shown in Table 2.

Table 1: Ensemble model accuracy – comparing $XF-bot with bot

	Training set	**Testing set**
Correct	7,426 (99%)	1,850 (98.14%)
Wrong	75 (1%)	35 (1.86%)
Total	7,501 (100%)	1,885 (100%)

Table 2: Confidence Values Report for $XFC-bot

'Partition' = Testing set	**Confidence**
Range	0.473 – 1.0
Mean Correct	0.984
Mean Incorrect	0.828
Always Correct Above	0.993 (55.07% of cases)
Always Incorrect Below	0.474 (0.05% of cases)
98.14% Accuracy Above	0.0
2.0 Fold Correct Above	0.933 (99.33% of cases)

In total, 94,451 Twitter accounts were classified by our model. The model assigned 2,411 accounts for a *$XF-bot* score of 1. The remaining 92,040 accounts were assigned a *$XF-bot* score of 0, with 0 meaning that the model did not forecast the account to be a bot and 1 meaning that the model did forecast the account to be a bot. The confidence coefficient *($XFC-bot*) for all predictions in total yielded a mean of 0.944 and a range of 0.544, with 1 for absolute confidence and 0 for zero confidence in the model's prediction (Table 3.).

Table 3: Statistics of confidence coefficient- $XFC-bot

'Dataset – Japan Twitter accounts'	Confidence
Mean	0.944
Min	0.450
Max	0.994
Range	0.544
Median	0.943

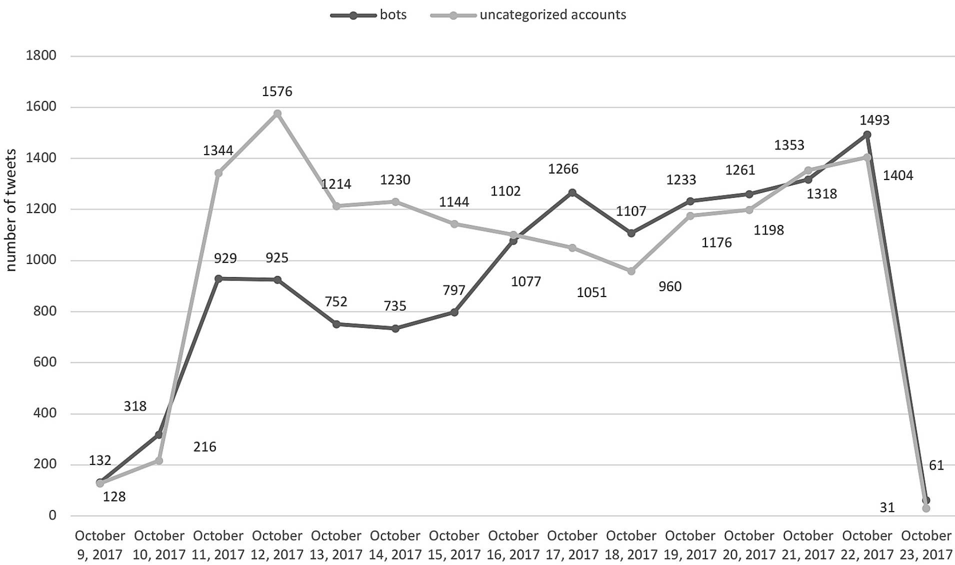
The analysis of tweets in our Japanese general election dataset resulted in the identification of 13,404 tweets generated by accounts classified by our model as bots (Table 4.). The overall tweeting volume of the identified bots increased from 318 tweets on October 10, 2017, to 1,493 tweets on October 22, 2017 (election day). After the election, the tweeting volume dropped to 61 tweets on October 23, 2017 (Figure 1). Due to time zone differences and the technical setup during the data collection phase, we also observed 132 tweets in the dataset from October 9, 2017.

Table 4: Twitter bot accounts – number of tweets

'Dataset – Bot accounts'	Tweets
Mean	5.590
Min	1
Max	336
Median	1

The second analysis of tweets in our Japanese general election dataset was set to identify tweets generated by the 2,466 uncategorized accounts. This analysis resulted in the identification of 15,127 tweets generated by such accounts. The tweeting volume of the uncategorized accounts increased from 216 tweets on October 10, 2017, to a maximum of 1,576 tweets on October 12, 2017. On election day, 1,404 tweets were generated. After the election, the volume dropped to 31 tweets on October 23, 2017 (Figure 1). Due to time zone differences and the technical setup during the data collection phase, we also observed 128 tweets in the dataset from October 9, 2017.

Figure 1: Time series of tweets generated by bots and uncategorized accounts



Trending hashtags and retweets

The analysis of trending Twitter hashtags and most retweeted statuses in Japan on the election day was automatically performed by Trendsmap Analytics (Trendsmap Pty Ltd, Chelsea, VIC, and Australia) on a dataset of 59,829,000 tweets. From this dataset, Trendsmap Analytics classified 19,391,000 tweets as retweets and 14,485,100 as replies. The most retweeted post on the election day was by *@McDonaldsJapan* with 166,450 retweets and 41,008 likes, followed by a tweet *by @Tekken_omachi* with 71,115 retweets and 72,058 likes. The third most-shared tweet on election day was by *@711SEJ* with 40,830 retweets and 3,861 likes. The top three hashtags shared on election day were #立憲民主党, *#tspook, and #*比例は共産党. [iv]

Discussion

The 2017 Japanese general election seemed to keep Japan out of the contemporary struggle of most of the Western democracies—Twitter bots trying to massively shape public opinion during critical political events. Our research showed a relatively low presence of bots during the 2017 Japanese general election, as only 2,411 of 94,451 accounts in our dataset were classified as bots by our model. This number of bots accounted for a relatively low number of bot-generated tweets—13,404 of the 665,400 collected tweets in total. Even if accounting for the 15,127 tweets generated by suspended or deleted accounts, the total possible number of bot-generated tweets was roughly 4% of the entire collected conversation. In our results of tweet sharing in Japan, we documented a negligible interest in sharing or even creating election-related content on Twitter. The presented results are based on the collection and processing of vast amounts of Twitter data and the use of a supervised machine learning approach to build a custom bot classification ensemble model. The classification model was trained on 9,386 Twitter accounts and has an AUC (testing set) of 0.998.

However, the observed low number of Twitter bots involved in 2017 Japanese general election-related discussions on Twitter raises an interesting question. Why is Japan—the third-largest economy in the world, a country with a considerable defence capability, an influential global player, and one of the most important difference-makers in the East Asian regional order—not dragged into a complex bot battle on Twitter? At first sight, our results seem to be contradictory to those of other widely cited research on the presence of bots during elections in other countries. For example, Bessi and Ferrara (2016) have documented that approximately 400,000 bots were engaged in political discussions about the 2016 presidential election in the U.S., with approximately 3.8 million tweets generated by bots, representing approximately one-fifth of the entire conversation. Similar results in the range of 12.5% to 25% bot-generated tweets were also found by other studies examining this phenomenon in France, the US and the UK (Ferrara 2017; Kollanyi et al. 2016; Gallacher – Kaminska – Kollanyi – Howard 2017). However, the presence of bots during political events seems to be a rather complex issue influenced by various factors. Distinct methods of bot usage vary by country and political instance (Wolley 2016). In the case of Japan, we propose various inter-combinable explanations for the low number of bots involved in the discussions related to the Japanese election on Twitter. This account, however, should not be considered to be an explicit list of all possible explanations; instead, it should serve as a thought base for further research.

One of the explanations for the low number of observed Twitter bots during the election-related discussions in Japan might stem from the specific demography of the island nation. Demographic factors play an essential role in shaping the political campaign and shaping the means of communication and the media market (Sciubba 2012; Tepe – Vanhuysse 2009). The share of the population over 65 years is considerably high in Japan (26.6% in 2015), with the share of the Japanese population over 80 years reaching 8% in 2015 (He – Goodkind – Kowal 2016). According to projections made by the U.S. Census Bureau, the share of elderly citizens in Japan will grow even more significantly in the future. It is estimated that the share of Japanese citizens older than 65 years will reach 40.1% in 2050. Japan continuously ranks as having one of the highest life expectancies in the world. This trend will most likely continue, as the life expectancy projections for Japanese individuals born in 2050 are 91.6 years for both sexes (He et al. 2016). Japan also has the highest percentage of centenarians globally. In September 2017, the number of centenarians rose to 4.8 per 10,000 citizens (Matsukawa 2017). Japanese demography might not only affect voting behaviour but also the behaviour of political institutions, the media environment, and the perception of new means of social media, as well as the will of politicians and campaign managers to accept and use such new means. Among other things, this notion seems to be apparent in the share of traditional media in Japan, as traditional media such as newspapers and TV broadcasts enjoy a sizable audience (NHK Broadcasting Culture Research Institute 2016; NSK 2016).

The Japanese newspaper market has one of the highest circulations in the world, with 44 million daily newspapers (WAN-IFRA 2016). This amount of newspaper circulation is further backed by a robust delivery system, as 95% of newspapers bought in Japan are delivered (NSK 2016). For the Japanese television market, according to the NHK Broadcasting Culture Research Institute, 85% of Japanese households watched TV in 2015, with the average time spent watching TV being 203 minutes per day (NHK Broadcasting Culture Research Institute 2016). According to Ogashara (2018), the position of traditional forms of media is further solidified by the high level of trust that the Japanese have in newspapers and TV broadcasts, mostly due to the older population and a low level of media polarization. Since almost the entire post-war era in Japan has been ruled by *Jiyū-Minshutō*, the media lacks the motivation to be excessively critical. Being neutral helps the media attract a larger audience (Ogasahara 2018). However, the strong popularity of traditional forms of media in Japan might also be a reflection of the nation's cultural nature. The Japanese are considered to be naturally passive, slowly adapting to changes. This characteristic is noticeable, for example, in Japanese political debates, where Japanese politicians commonly address forces beyond their reach, such as *hitsuzen no ikioi* (inevitable momentum) (Pyle 2006; Benedict 2014; McFarlane 2009).

The high share of traditional media in Japan is, of course, reflected in the relatively low share of new media. According to the 2016 report by Poushter (2016), internet penetration in Japan was 69% among adults and 91% overall (Newman – Fletcher – Levy – Nielsen 2016). This figure is far behind the percentage of adults using the internet in advanced economies, such as the US, Canada, or South Korea. Statistics on the percentage of adults owning a smartphone show that Japan is 39% behind the global average of 43%. The share of Japanese citizens over 35 years of age owning a smartphone device was considerably low: 31% compared to 77% among millennials (Poushter 2016). These statistics seem to be reflected in the relatively low usage of social networking services in Japan. Twitter, the most popular social networking service in Japan, has more than 45 million users in Japan, which is approximately one-third of the country's entire population (Nussey – Ingram 2018). In the 20 to 29 age group, the usage rate of Twitter was 30.19%, followed by 22.14% in the 13 to 19 age group. Only 3.55%

of people older than 60 years use Twitter (MIC 2016). Statistics published by the Reuters Institute for the Study of Journalism document a lower number of interactions with news on social networking services in Japan than in the US or South Korea. In 2016, only 9% of the Japanese population shared a news story on a social networking service. In the US, this proportion was unambiguously higher at 25%; in Taiwan, it was 34%. A similar pattern emerges for commenting on news stories on social networking services. The Reuters study shows that only 6% of Japanese individuals commented on news stories via social networking services in 2016, compared with 22% of individuals in the US or 28% in Spain. These results, according to Newman et al. (2016), document the passivity of Japanese citizens as well as the specific state of the media environment.

A combination of the abovementioned factors seems to not only influence the media market but also political campaigns. Countless politicians in Japan run their campaigns in a more traditional manner, with the campaigns being based more on direct communication with citizens via public speaking in front of supermarkets or at subway stations. This approach to political marketing seems to further solidify the traditional, unattractive aura of political campaigns in the eyes of the younger generation of voters. Despite the significant support for Shinzo Abe and *Jiyū-Minshutō*, the percentage of the young Japanese population participating in elections seems to be inadequate. During the House of Councillors election in 2016, only 46.78% of young citizens turned out to vote, compared to the overall voter turnout of 54.70% (Nippon.com 2017). A possible explanation for this observed lack of interest, according to Shin (2016), seems to be precisely the "unattractiveness" of Japanese politicians to young audiences, mostly on social networking services. Traditional means of communication and political campaigns do not appear to be appealing to Japanese youth.

However, the distinctiveness of Japanese political campaigns might be the result of not only an ageing population but also strict laws. According to Plasser (2002), Japan has one of the most regulated campaigning practices in the world. The Public Offices Election Act from 1950 even determined details such as the dimensions of political posters. Regarding online campaigning, the Public Offices Election Act was amended only in 2013, allowing Japanese politicians for the first time in history to utilize the internet for political campaigning. Some commentators from the academic sphere expected an enormous change in the way campaigns were run. This optimism was transformed into the popular phrase *net senkyo* (net election). However, *net senkyo* optimism never met expectations that predicted a radical change in the means of political communication. In comparison with the U.S., South Korea, or Taiwan, Japan is the least developed country in terms of online campaign deployment (Kiyohara – Maeshima – Owen 2018). Politicians, among others, seem to lack experience in running online campaigns. Instead of broadening their potential electorate by using social networking services, they focus mostly on presenting their activities via public speaking. Despite the relatively poor utilization of the online environment, Fahey (2017) argues that there is a noticeable increase in Japanese politicians engaging in social networking services. His observation is based on the increased number of tweets between 2013 and 2017 by the members of the Japanese Diet.

However, the low number of observed Twitter bots during election-related discussions in Japan could also be caused by factors other than a distinct manner of political campaigning coupled with a more traditional media market, legislative regulations, and older voter demography. The low presence of bots during the election campaign in Japan may also be the result of an apparent lack of interest of other states in changing the political scene in Japan. Japanese relations with China and Russia, two countries frequently accused of using Twitter bots to shape public opinion, are complex and have several friction points, mostly in the form of territorial disputes. However, it is likely that both China and Russia had no relevant interest in influencing the 2017 Japanese general election. Shinzo Abe and his cabinet are relatively predictable in terms of foreign policy and attitudes towards other regional players. The continuity of relations among leaders of the nations that neighbour Japan is crucial in the long-term evolution of the East Asian order. In light of recent political development in China, the stability of the Japanese political scene might be beneficial for both states. It has already been reported that the Chinese leader Xi Jinping and Shinzo Abe underwent a reconciliation of mutual relations. According to Abe, *a new start for Japan-China relations should* lead to high-level mutual visits (Abe 2017b). The same goes for Japan-Russia relations. Russian president Vladimir Putin and the Japanese Prime Minister are, among other things, working closely on a final solution to the Northern Territories/Kuril Islands dispute, as well as on potentially deepening the economic cooperation between the two states (The Japan Times 2018).

In discussing possible explanations for the low number of observed Twitter bots during the 2017 Japanese general election, we should not forget the relatively short time period between Abe's announcement of the snap election and election day, as well as the language issue. Any potential agent seeking an opportunity to influence the Japanese election campaign via Twitter bots might face time pressure in devising a comprehensible strategy. The language of the island nation might also play a role. In programming effective human-like bots, natural language processing (NLP) is fundamental. However, as Halpern (2006) notes, the complexity of the Japanese language presents particular challenges to developers of NLP, especially in the area of word segmentation, information retrieval, named entity extraction, and machine translation.

All the possible explanations presented for the relatively low presence of Twitter bots during the 2017 Japanese general election are interesting topics for future research. Future studies might also consider examining the real-world impact of Twitter bots during political events in Japan. Despite the rigorous research methods implemented in this study, as with other research regarding the presence of Twitter bots during political events, some limitations apply. One of the most discussed limitations of studying bots during political events is whether the sample data downloaded through the Twitter API is a valid representation of the overall activity on Twitter (Morstatter – Pfeffer – Liu – Carley 2013). By using the Twitter Search API over the Twitter Stream API in our study set to perform continuous auto-updating, as advised by Bessi and Ferrara (2016), this approach should yield a complete sample of all the tweets containing a pre-defined hashtag and thus avoid the valid representation limitation. Another widely discussed limitation arises from using machine learning. Although bot detection mechanisms utilizing machine learning are considered to be the most precise en masse, they still face some restrictions. Bot detection solutions using supervised machine learning techniques are based on the underlying assumption that the difficulty and characteristics of the bots contained in the training dataset and those contained in the real-world environment at a specific time and space are generally alike. However, this assumption reduces the ability of bot detection models utilizing such techniques to correctly spot bots that are unlike the bots that the model was trained on. Therefore, to assure the maximum detection range of our model and to avoid the above limitation as much as possible, we trained our model on a pre-annotated dataset of traditional bots as well as on a new generation of bots. We also conducted exploratory data analyses of the datasets we were working with to detect anomalies between training and testing datasets, which could possibly decrease the accuracy of our model. This approach yielded a model with an AUC of 0.998. However, active deployment of our model in the future would require more testing and constant recalibration.

The power of Twitter bots attempting to shape or disrupt public opinion during defining social moments in Western countries should not be ignored. This activity has been used in massive attempts to influence various elections in the United States, France, and the United Kingdom, among others. However, given the results of this study, Japan seems to maintain its distance from this contemporary struggle in most Western democracies. Although this result is surprising at first glance, to adequately interpret it, one must consider the wider socio-political context of Japan discussed in this paper. In analysing these particularities and clarifying the issues presented in this study, future research might look deeper into the utilization rates of bots for Japanese political communication in general, cross platform coordination, as well as empirically test the links between the Japanese particularities mentioned and bot activity on social networking sites.

Data availability

The data that support the findings of this study, except the training data for the bot detection model, are available in the Zenodo repository: doi: 10.5281/zenodo.1413589 The training data for the bot detection model were used under license from Cresci et al. (2017), the data are available on request at mib.projects.iit.cnr.it/dataset.html.

Endnotes:

i The pre-identified hashtags were as follows: (1) 自民党 (Liberal Democratic Party); (2) この┌を守り┌く (Protect this country); (3) 安倍晋三 (Shinzo Abe); (4) 衆院選 (House of Representatives election); (5) 小池百合子 (Yuriko Koike); (6) 日本に希望を (Hope for Japan); (7) 希望の党 (Party of Hope); (8) 立憲民主党 (Constitutional Democratic Party); (9) 枝野幸男 (Edano Yukio); (10) 枝野出る (Edano Wake Up).

ii The complete technical build settings of the ensemble model (MVBDM.gm file) are available in the Zenodo repository: doi: 10.5281/zenodo.1413589

iii A detailed description of the algorithms used can be found at <ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/18.0/en/AlgorithmsGuide.pdf >.

iv The relevance of the top ten retweeted tweets and the top ten trending hashtags to the election was manually determined by two independent coders. Statistics for selected days prior to election day are available in the replication dataset for this study in the Zenodo repository: doi: 10.5281/zenodo.1413589

References

Abe, Shinzo (2017): Press Conference by Prime Minister Shinzo Abe Following His Attendance at the APEC Economic Leaders' Meeting, ASEAN-related Summit Meetings, and Other Related Meetings, (14 November), Cabinet Public Relations Office, available at https://japan.kantei.go.jp/98_abe/statement/201711/_00007.html.

Abe, Shinzo (2017): Press Conference by Prime Minister Shinzo Abe, (15 September), Cabinet Public Relations Office: available at https://japan.kantei.go.jp/97_abe/statement/201709/_00011.html.

Abokhodair, Norah – Yoo, Daisy – McDonald, David W. (2015): Dissecting a social botnet: Growth, content and influence in Twitter. In *CSCW '15: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 839–851. available at doi:10.1145/2675133.2675208

Asahi, Shimbun (2017): Asahi poll: 70% not buying Abe's justification for calling snap poll', The Asahi Shimbun (28 September): available at http://www.asahi.com/ajw/articles/AJ201709280047.html.

Benedict, Ruth (2006): *The Chrysanthemum and the Sword*, Houghton Mifflin Harcourt.

Berger, Thomas U. (2003): *The Cultures of Antimilitarism. National Security in Germany and Japan*, John Hopkins University Press.

Bessi, Allesandro – Ferrara, Emilio (2016): Social bots distort the 2016 U.S. Presidential election online discussion, First Monday, 21: available at doi:10.5210/fm.v21i11.7090.

Bode, Leticia – Vraga, Emily, K. – Borah, Porismita – Shah, Dhavan, V. (2013): A New Space for Political Behavior: Political Social Networking and its Democratic Consequences, 414–429, *Journal of Computer-Mediated Communication* 19.

Bond, Robert, M. – Fariss, Christopher, J. – Jones, Jason, J. – Kramer, Adam. D. I. – Marlow, Cameron, – Settle, Jamie – Fowler, James, H. (2012): A 61-million-person experiment in social influence and political mobilization, 295–298, Nature, 489.

Brachten, Florian – Stieglitz, Stefan – Hofeditz, Lennart – Kloppenborg, Katharina – Reimann, Annette. (2017). Strategies and Influence of Social Bots in a 2017 German state election – A case study on Twitter, Australasian Conference on Information Systems, available at arxiv.org/abs/1710.07562.

Calder, Kent, E. (2018): Japan in 2017: Political Consolidation amid Global Volatility, 43–54, *Asian Survey* 58.

Cresci, Stefano – Pietro, Roberto – Petrocchi, Marinella – Spognardi, Angelo – Tesconi, Maurizio (2017): The Paradigm-Shift of Social Spambots, 963–972, Proceedings of the 26th International Conference on World Wide Web Companion.

Davis, Clayton – Varol, Onur – Ferrara, Emilio – Flammini, Allessandro – Menczer, Fillippo (2016): 'BotOrNot: A System to Evaluate Social Bots', 273–274, Proceedings of the 25th International Conference on World Wide Web Companion.

Ferrara, Emilio – Varol, Onur – Davis, Clayton – Menczer, Fillippo – Flammini, Allessandro (2016): The Rise of Social Bots, 96–104, Communications of the ACM, 59.

Ferrara, Emilio (2017): Disinformation and social bot operations in the run up to the 2017 French presidential election, First Monday, 22, available at doi:10.5210/fm.v22i8.8005.

Gallacher, John. D. – Kaminska, Monica – Kollanyi, Bence – Howard, Philip N. (2017): Junk News and Bots during the 2017 UK General Election: What Are UK Voters Sharing Over Twitter? Data Memo 2017.5, Project on Computational Propaganda, University of Oxford, available at https://www.oii.ox.ac.uk/blog/junk-news-and-bots-during-the-2017-uk-general-election/

Gerlitz, Carolin – Rieder, Bernhard (2013): 'Mining One Percent of Twitter: Collections, Baselines, Sampling, ' M/C Journal 16, available at://journal.media-culture.org.au/index.php/mcjournal/article/view/620

Halpern, Jack (2006): 'The Contribution of Lexical Resources to Natural Language Processing of CJK Languages', Chinese Spoken Language Processing – 2006, 768–780, Springer

He, Wan – Goodkind, Daniel – Kowal, Paul – U.S. Census Bureau (2016): 'International Population Reports, P95/16-1, An Aging World: 2015', U.S. Government Publishing Office.

Howard, Philip N. – Kollanyi, Bence (2016): 'Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum, Research Note 2016. 1. ', Project on Computational Propaganda, University of Oxford, available at arxiv.org/abs/1606.06356

Fahey, Rob (2016): Election 2017: Candidates Take to Twitter. *Tokyo Review*, available at http://www.tokyoreview.net/2017/10/twitter-japan-election-campaign

Kiyohara, S., Maeshima, K., & Owen, D. (2018). *Internet Election Campaigns in the United States, Japan, South Korea, and Taiwan.* London: Palgrave Macmillan.

Kollanyi, Bence – Howard, Philip N. – Woolley, Samuel C. (2016): 'Bots and Automation over Twitter during the U.S. Election, Data Memo 2016'. Project on Computational

Propaganda, University of Oxford, available at http://blogs.oii.ox.ac.uk/politicalbots/wp--content/uploads/sites/89/2016/11/Data-Memo-US-Election.pdf

Macfarlane, Alan. (2009): *Japan Through the Looking Glass,* Profile Books.

Matsukawa, N (2017): '100-year-olds in Japan hit record high population of 67,824', The Asahi Shimbun (15 September): available at http://www.asahi.com/ajw/articles/AJ201709150056.html.

Ministry of Internal Affairs and Communications (MIC). (2016): 'Jōhō tsūshin media no riyō ji-kan to jōhō kōdō ni kansuru chōsa hōkoku-sho' (Survey on utilization time and information behavior of information and communication media) ', Research institute, Ministry of Internal Affairs and Communications, available at http://www.soumu.go.jp/iicp/chousakenkyu/data/research/survey/telecom/2016/02_160825mediariyou_houkokusho.pdf

Morstatter, Fred – Pfeffer, Jurgen – Liu, Huan – Carley, Kathleen M. (2013): 'Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose', 400–408, Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media

Newman, Nic, Fletcher, R., Levy, D. A. L., & Nielsen, R. K. (2016): 'The Digital News Report 2016',Reuters Institute for the Study of Journalism.

NHK Broadcasting Culture Research Institute. (2016): NHK Data Book 2015 National Time Use Survey, NHK Publishing.

Nihon Shinbun Kyokai (NSK). (2016): 'Percentage of Home Delivery', Nihon Shinbun Kyokai (The Japan Newspaper Publishers & Editors Association), available at http://www.pressnet.or.jp/english/data/circulation/circulation03.php.

Nippon.com (2017): 'Japan Teenagers to Vote in General Election for First Time (News)', Nippon Communications Foundation (10 October): available at https://www.nippon.com/en/genre/politics/l10544/.

Nussey, Sam – Ingram, Daniel (2018, February 28): 'In Japan, Twitter sees a surge of users – and revenue'. Reuters (28 February) available at https://www.reuters.com/article/us-twitter-japan/in-japan-twitter-sees-a-surge-of-users-and-revenue-idUSKCN1GC17T.

Ogasahara, Morihiro (2018): 'Media Environments in the United States, Japan, South Korea, and Taiwan. Internet Election Campaigns in the United States, Japan, South Korea, and Taiwan', 79–113, Springer International Publishing AG

Plasser, Fritz – Plasser, Gunda (2002): *Global Political Campaigning: A Worldwide Analysis of Campaign Professionals and Their Practices*, Praeger.

Poushter, Jacob (2016): 'Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies'. Pew Research Center, available at www.pewglobal.org/files/2016/02/pew_research_center_global_technology_report_final_february_22__2016.pdf.

Pyle, Kenneth B. (2006): 'Profound Forces in the Making of Modern Japan', 393–418 *Journal of Japanese Studies*, 32.

Repeta, Lawrence (2017): 'Backstory to Abe's Snap Election – the Secrets of Moritomo, Kake and the "Missing" Japan SDF Activity Logs'. Asia-Pacific Journal: Japan Focus15, available at https://apjjf.org/2017/20/Repeta.html.

Sciubba, Jennifer D. (2012): ' Rhetoric and action on ageing in the world's 'oldest' democracies: party platforms and labor policies in Germany, Italy, and Japan', In Ageing Populations in Post-Industrial Democracies Comparative Studies of Policies and Politics, 54–78, Routledge.

Shin, Sato (2016): 'Teen Voters and Politics in Japan ', Nippon Communications Foundation (19 August): available at https://www.nippon.com/en/currents/d00231/.

Shirky, Clay (2011): 'The Political Power of Social Media: Technology, the Public Sphere, and Political Change', 28–41, Foreign Affairs, 90.

Schäfer, Fabian – Evert, Stefan – Heinrich, Philipp (2017): 'Japan's 2014 General Election: Political Bots, Right-Wing Internet Activism, and Prime Minister Shinzō Abe's Hidden Nationalist Agenda', 294–309, Big Data, 5.

Subrahmanian, V. S. – Azaria, Amos – Durst, Skylar – Kagan, Vadim – Galstyan, Aram – Lerman, Kristina –… Hwang, Tim (2016): The DARPA Twitter Bot Challenge, 839–851, Computer, 49.

Tepe, Markus – Vanhuysse, Pieter (2009): 'Are Aging OECD Welfare States on the Path to Gerontocracy? Evidence from 18 Democracies, 1980–2002', 1–28, Journal of Public Policy, 29.

The Japan Times (2018): Japan might ask Russian President Vladimir Putin to visit in 2019 to discuss isle row ', The Japan Times (6 January): available at https://www.japantimes.co.jp/news/2018/01/06/national/politics-diplomacy/japan-mulls-inviting-russian-president-putin-2019-press-isles-issue/

Woolley, Samuel C. (2016): 'Automating power: Social bot interference in global politics', First Monday, 21, available at doi:10.5210/fm.v21i4.6161.

World Association of Newspapers and News Publishers (WAN-IFRA) (2016): 'World Press Trends Report 2016', World Association of Newspapers and News Publishers.

Yang, Chao – Harkreader, Robert – Gu, Guofei (2013): Empirical Evaluation and New Design for Fighting Evolving Twitter Spammers, 1280–1293, IEEE Transactions on Information Forensics and Security, 8.

Jozef Michal Mintal *a doctoral candidate in International Relations at Matej Bel University, Fellow at the Centre for Media, Data and Society at Central European University, and research Fellow and co-founder of the UMB Data & Society Lab, where he is working under the IBM funded CARWST research project. His research focuses on the digital weaponization of information, communication and technology, as well as on the societal implications of these processes. Email: jozef.mintal@umb.sk.*

Róbert Vancel *is an UMB Data & Society Lab Associated Fellow and PhD Candidate at Matej Bel University. His research is focused on the position of Japan in the East Asia Region, and Japan's foreign and defence policies. In 2016 he participated in MIRAI, an exchange program financed by the Ministry of Foreign Affairs of Japan. Email: robert.vancel@umb.sk.*