Disinformation from the Inside: Combining Machine Learning and Journalism to Investigate Sockpuppet Campaigns

Christopher Schwartz
KU Leuven
Institute of Philosophy
openDemocracy.net
Christopher.Schwartz@kuleuven.be

Rebekah Overdorf
EPFL
Distributed Information Systems Laboratory
rebekah.overdorf@epfl.ch

ABSTRACT

This paper brings together machine learning and investigative journalism to examine sockpuppets accounts, a historical breed of fake accounts that are non-automated and human-controlled. Due to their flexible and human-centered nature, sockpuppets pose a complication for purely technological approaches to detecting and studying fake accounts. We find that as machine learning-based detection methods of bots slowly grow stronger, adversaries engaging in disinformation are turning to such sockpuppets accounts, and in particular a subset of sockpuppets that we call "infiltrators" — those that aim to integrate into a community in order spread disinformation. This represents a new stage in the evolution of the sockpuppet concept: where bots seek to simulate audiences and drown online social media platforms with a particular point of view, infiltrators seek to persuade and assimilate genuine audiences from within. In addition to these insights into infiltrator sockpuppets, combining machine learning and investigative journalism enables learning something more than detection and important patterns of activity: it can also gain a sense of the motivations and reasoning of adversaries who engage in disinformation.

KEYWORDS

Social Networks, Disinformation, Sock Puppets

ACM Reference Format:
Christopher Schwartz and Rebekah Overdorf. 2020. Disinformation from the Inside: Combining Machine Learning and Journalism to Investigate Sockpuppet Campaigns. In *Companion Proceedings of the Web Conference 2020 (WWW '20 Companion), April 20–24, 2020, Taipei, Taiwan.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3366424.3385777

1 INTRODUCTION

Fake social media accounts, or simply "fake accounts", represent a multifaceted threat globally. Although great strides have been made in computer science to detect and study them, ultimately the threat needs to be contended with in an interdisciplinary manner. The problem of fake accounts cuts across multiple disciplines, from military studies to marketing to information systems.

This paper presents one such interdisciplinary response. Our research combines machine learning and investigative journalism to examine sockpuppets, an historical breed of fake accounts that are non-automated and human-controlled [6]. Sockpuppets pose a complication for purely technologically-based approaches to fake accounts as they are, so to speak, analogue in nature.

While traditionally the definition of sockpuppets is sometimes limited to those accounts controlled by someone who is already integrated in the community, we take a broader definition for the purpose of this work and consider a sockpuppet account to be any account that aims to appear like a real user. Further, we call such accounts that also aim to infiltrate a community and spread disinformation *infiltrators*. We have found that as machine learning-based detection methods of bots slowly grow stronger, adversaries engaging in disinformation are turning to the use of infiltrator accounts instead.

Our research seeks to do the following:

- Identify sockpuppets that are disseminating disinformation in the form of "fake news."
- Map sockpuppets' connections to each other and to genuine accounts.
- Monitor and assess the ways in which these sockpuppets are disseminating disinformation.
- Classify and model these interactions.
- When possible, unmask the entities behind them.

Each of these goals is achieved by operationalizing supervised machine learning as a method of journalistic investigation. We have assembled a team of journalists who assist in the supervision of machine learning tasks, including establishing ground truth, harvesting sample sets, derive features, train, and validate results.

Our case study is Kyrgyzstan, a post-Soviet Central Asian republic with unfortunately "ideal" conditions for the success of disinformation. As a key part of this work, we have interviewed a whistleblower source who purports to have been a a former member of a campaign of sockpuppet in Kyrgyzstan headed by a major Kyrgyz politician. It was this individual who led us to focus on sockpuppets to begin with; they have also led us to focus primarily on Facebook, although we have also been researching Twitter.

This work is thus unique in that it is providing a chance at seeing how disinformation works from the inside. By combining journalism and machine learning, we can learn the following:

- *From a journalistic perspective*, what motivates disinformers, and how fake accounts are used to manipulate public opinion and censor opposition, moderate or alternative voices.
- *From a technical perspective*, how what motivates fakeness is expressed in the mechanics of fake accounts and how fakeness relates to other social media attacks.

WWW '20 Companion, April 20–24, 2020, Taipei, Taiwan

ACM ISBN 978-1-4503-7024-0/20/04.
https://doi.org/10.1145/3366424.3385777

- *From a combined perspective*, how insights from the political context can be used to study fake accounts.

This paper will proceed by first reviewing our definition of fake news, then explain our use of Kyrgyzstan as a case study. Following this, we will turn to the investigation itself. We will conclude by offering a description of the infiltrator sockpuppet.

2 DEFINING FAKE NEWS

Many definitions of fake news abound [27]. We understand fake news as the deliberate and, crucially, *malicious* presentation of verifiably misleading or untrue claims as *news-reporting*. Our definition precludes parody news [42] while includes not only fabricated or manipulated content that presents itself as traditional news-reporting from a journalistic agency, but also generally anything published by an online social media account that purports to announce a "truth".

The role and flexibility of the meaning of truth is key to understanding fake news as a form of disinformation. The truth presented by fake news-reports can be *either* objective or subjective:

- "Objective" in the sense of an incontestable fact, even if it turns out that the fact has been misconstrued or invented and hence false.
- "Subjective" in the sense of the perspective of a demographic group. For example, if party X feels itself oppressed by party Y, separate from what either party Y might assert or what third-party Z might be able to say about the interaction between the two parties.

One may argue that we might be better served by sticking with the often preferred and inherently expansive notion of "disinformation" [23]. We initially conceived of this problem in terms of propaganda — biased or untrue information used for political purposes. However, the terminology of fake news captures something important that disinformation alone misses: fake news typically emulates tabloid journalism [11]. As such, it often involves "clickbait" themes of exaggeration, such as enticement, enigma, taboo and transgression [12], essentially weaponizing tabloid news-reporting semantics. This seems an essential ingredient for its virality.

3 KYRGYZSTAN AS CASE STUDY

We chose Kyrgyzstan because in its brief history as a sovereign state — it gained independence from the Soviet Union in December 1991 — the country has found itself as a proving ground for disinformation techniques. When assessing how Kyrgyzstan became such a proving ground, we reasoned from the perspective of an adversary seeking to disinform by using fake news. We concluded that, unfortunately for Kyrgyzstan, the country meets certain important pre-conditions for high susceptibility to disinformation.

3.1 Conditions for "Successful" Disinformation

From the perspective of an adversary, there are at least two ways to assess the potential success of a piece of fake news content: a) from within the content itself, and b) with respect to its target audience. The former first and foremost concerns semantics, the latter social-political conditions. Regarding semantics, as Justin Cheng et al. have so succinctly put it, that "anyone can become a troll" [13], so long as the fake news content's tabloid semantics can manipulate an audience's sense of the taboo [30, 31].

That said, not all audiences are created equally. The success chance of fake news can be dramatically increased if the target audience is either in a state of alarm (e.g., during a natural disaster) [33] or if it feels itself to be disenfranchised or otherwise marginalized from the broader social-political system [17]. This feeling of marginalization can either be issue-specific, or it can be generalized. In other words, the target audience may feel themselves to be full members of society *"except"* about X, or the target audience may feel ostracized *in general*. An example of the first would be voters who may have a specific policy preference that bucks against the majority's preference, while an example of the second would be ethnic minorities in an ethnic nation-state.

Kyrgyzstan meets these conditions with a vengeance. For example, at the time of independence, the titular ethnic group, the Kyrgyz, comprised approximately 52 percent of the nation's 4.25 inhabitants. By 2018, the total population had swollen to 6.27 million, of which approximately 73 percent were ethnic Kyrgyz [1]. This transition has entailed tensions between the Kyrgyz and ethnic minorities, twice erupting in violence in the country's southern regions in 1990 and 2010 [40]. Many ethnic Kyrgyz have felt underrepresented in their "own" nation-state, [28] and yet at the same time, minorities have felt sidelined [21]. Throughout its nearly three decades of existence, Kyrgyzstan has suffered repeated bouts of economic and political upheaval as it lurched from communism to capitalism, and from de-industrialization to a service economy, and meanwhile, from a presidentialist system of government to a parliamentary one. This upheaval has involved recessions, two revolutions, and as recently as August 2019, a failed uprising. A sense of chronic crisis, coupled with "revolution fatigue", pervades Kyrgyzstan's populace [38].

We further posit that a society is most susceptible to disinformation when its social-political and media landscape is generally democratic. To be sure, disinformation can be effective in other contexts, especially depending upon its task and technique. For example, Kyrgyzstan's neighbor Kazakhstan is authoritarian [18], but its population has proven vulnerable to fake news about China-related topics, such as the possible spread of the coronavirus [34]. Nevertheless, insofar that anti-Chinese fake news has sought to disrupt the Kazakh-Chinese geostrategic relationship, it has been far less disturbing to the status quo than similar anti-Chinense fake news in Kyrgyzstan [16], which is a far more open society.

In a region notorious for authoritarianism, Kyrgyzstan boasts a generally free market, as well as parliamentary and presidential elections held more often than not according to transparent schedules and rules, not the whims of ruling elites [29]. The country also stands out among its neighbors as having a genuinely free media landscape, with very little *explicit* censorship [9]. Indeed, the capital, Bishkek, acts as a "haven" for journalists from across Central Asia [20], and the International Exchange and Review Board has assessed Kyrgyzstan's domestic medias landscape as "near sustainable" in its 2019 Media Sustainability Index [22]. Similarly, whereas throughout Central Asia, state control of Internet Service Providers (ISPs) is the norm, whether by overt monopolization or covert means [15], Kyrgyzstan's government not only lacks

such control, but it also has "no functional legal and technical provisions exist[ing] for shared use of existing infrastructure by ISPs, forcing them to build redundant and expensive infrastructure" [4]. Although the state-owned KyrgyzTelecom remains the largest ISP within the Kyrgyzstani market, this situation nonetheless poses quite a hurdle to potential state surveillance and censorship.

At the same time, Kyrgyzstan is beset with many anti-democratic problems. The quality of elections is marred by widespread vote-buying and tampering by the ruling party through state mechanisms. Meanwhile, self-censorship is a rampant practice among local journalists [25], partially due to a history of periodic violence against the profession [43], and partially due to ruling elites exploiting the weak judicial system to target the press with inflated slander lawsuits [10]. It also important to note that as a former Soviet republic, Kyrgyzstan has inherited a deep historical legacy of exposure to propaganda. In the assessment of analyst and journalist Ermek Baisalov, "Kyrgyzstan's vulnerability to the issue of disinformation points [to] several factors, such as weak domestic journalism, undeveloped home television content, demand for the 'tabloid' press, and lack of educational media literacy programs for the younger and older generations" [7].

3.2 Kyrgyzstan's Bouts with Disinformation

These problems combine to augment susceptibility to disinformation within Kyrgyzstan by increasing the sense of disenfranchisement and marginalization across the general population, regardless of ideology and demographics. Indeed, the history of the country over the past decade has been rocked by disinformation, particularly in the form of fake news:

- In August 2019, during a failed uprising launched by Kyrgyzstan's former president Almazbek Atambayev, content attempting to inflame intra-Kyrgyz regional rivalries spread rapidly through online social networks [26]. Fake accounts, and even entire fake community groups, were discovered on Facebook as key elements of the dissemination effort [35].
- Throughout 2018, fake news about Chinese "imperialism" fuelled violent anti-Chinese demonstrations throughout the country, risking another crucial relationship. Demonstrators called for bans preventing Chinese citizens from doing business in Kyrgyzstan and marrying Kyrgyz women [3].
- In 2014 and 2015, fake news about an imminent nuclear war between the United States and Iran, as well as about American support for "gay propaganda" and ethnic Uzbek separatism, rallied public support for Atambayev's pro-Russian government to shut down an American military base and cancel an important 1991 bilateral treaty [24, 32, 36].
- As far back as June 2010, fake news about ethnic Uzbek separatists attempting to secede from the country, as well as about Uzbek men sexually assaulting ethic Kyrgyz women, have been said to have provoked inter-ethnic clashes throughout Kyrgyzstan's southern region [19].

4 HUNTING DOWN SOCKPUPPETS

As noted above, our research seeks to do the following:

- Identify sockpuppets that are disseminating disinformation in the form of fake news.
- Map sockpuppets' connections to each other and to genuine accounts.
- Monitor and assess the ways in which these sockpuppets are disseminating disinformation.
- Classify and model these interactions.
- When possible, unmask the entities behind them.

Each of these goals is achieved by operationalizing supervised machine learning as a method of journalistic investigation. We have assembled a team of journalists who performing the following blocs of machine learning tasks:

- *Establishing ground truth* by manually identifying a) sockpuppets that seed fake news, and b) real accounts that serve as re-propagators. In order to identify sockpuppets, we proceed down to general tracks: using domestically-generated, context-specific fake news as an indicator (cross-checked with other meta-data), and using a whistleblower source. Our focus has been upon Facebook and Twitter, especially the former due to a lead from our whistleblower. Our Grand Unifying Taxonomy has aided us in sifting through the diversity of fake account activity occurring within Kyrgyzstan.
- *Harvesting sample sets*, which normally would mean deploying crawlers, but because we are focusing upon Facebook [8], our journalists must manually go through accounts. They collect both accounts they are certain are sockpuppets and those they are uncertain of, organizing them into separate databases. This has had the happy side-effect of adding an extra level of refinement to our features sets.
- *Deriving features*, *training* and *validating results*, all of which make use of the journalists's context knowledge. Because the average size of a sockpuppet campaign in Kyrgyzstan appears to range between 10 and 100 fake accounts, our sample sets are very small. Consequently, we use k-fold cross-validation. An iterative process thus ensues, in which the journalists refine the feature sets, whereupon the classifiers are retrained and the results re-checked against the journalists' context knowledge.

At the time of this writing, we are harvesting sample sets and deriving features – and what we are finding is provocative, as we will explain in the conclusion below. We are building three databases of fake accounts from three distinct sockpuppet campaigns that have been linked to three different political figures within Kyrgyzstan.

4.1 A Whistleblower Comes Forward

A critical moment in our research came when we interviewed a whistleblower source who purports to be a former member of a campaign of sockpuppets established by a major Kyrgyz politician. This individual has buttressed their claims by doing two things:

- They claimed to have been the originator and holder of an account that we had already identified as fake in April 2018 and which, to the best of our knowledge, they could not have known that we knew was such.
- They provided documents that originated from within the purported campaign. As of this writing, we are still attempting to independently verify the authenticity of the documents. Nevertheless, we consider them likely to be authentic.

The whistleblower first showed the documents to us by logging into the cloud-based server upon which they were being stored and showing them to us. They then subsequently provided us PDF copies. Unfortunately, this means that we have not been able to use meta-data analysis to confirm the provenance of the documents. Nevertheless, internal content analysis suggests they are genuine.

Our codename for our whistleblower source is "Piala", the word for the small Central Asian bowl that serves as the region's traditional tea cup; we also refer to the Kyrgyz politician as "Chainik", a teapot. According to Piala, the sockpuppets involved are "high quality", meaning they are few in number but resource-intensive [37]. Crucially, this particular campaign and many others in Kyrgyzstan have been trained in machine learning-based methods of bot detection, meaning that they are actively attempting to elude and confound purely technological-based attempts at discovery [37].

According to Piala, the original goal of the campaign was to support another politician who was the latter's client. However, by the Autumn of 2018, Chainik considered reorganizing the campaign members themselves into online social network-based "opinion-makers", potentially under their real identities, for the purpose of provoking a "revolution" in Kyrgyzstan [37]. Piala has cited Chainik's shift in strategy as a secondary reason why they came forward [37] (although we must note that we have not yet found any evidence that Chainik pursued this idea).

While Piala was still a campaign-member, they were paid in cash in person by a political operator closely associated with Chainik. The campaign was organized on a cloud-based server, replete with spreadsheets in which campaign members logged their activities, including URLs. A document that Piala has provided us is precisely one such spreadsheet: among other things, it includes both the names of the fake accounts and the names of the real people who controlled them, including complete names.

It is important to note that Piala approached us to begin with, we did not find her. We are not at liberty to explain how they came to be aware of our investigation. Piala claims to have approached us for moral reasons. Specifically, they describe themselves as having been a supporter of the campaign's goals in principle, but came to disagree with its methods, in particular the use of sockpuppets in-themselves. To be clear, Piala has not expressed moral concern about the content promulgated by the sockpuppets – to the contrary, they evince belief in the content's sincerity and do not seem to interpret it as a form of fake news (or, at least, they do not interpret as fake news the content that was promulgated during their time with the campaign) [37]. We should also note that Piala does not seem aware of the specific notion of "sockpuppet", and instead has referred to the human-controlled accounts simply as "fake accounts" [37].

4.2 Chasing the Whistleblower's Leads

As journalists say, we have been "chasing" Piala's leads. First, our focus has been narrowed primarily upon Facebook. The whistleblower claims has been the main locus of sockpuppet activity by Kyrgyz politicians due to the opportunities Facebook presents to have deep and potentially persuasive engagements with audiences [37]. We intend to expand our research into Instagram, where suspected fake accounts of notable Kyrgyz officials have been found [41]. In the meantime, we have looked into Kyrgyzstan's segment of Twitter and indeed found it to be primarily an echo of what occurs within Kyrgyzstan's segment of Facebook.

Piala has been particularly important in narrowing our focus upon sockpuppets. This has seemed logical to us given the adversaries' reasoning to focus upon Facebook: these campaigns are actively seeking to persuade audiences of certain perspectives, not to flood the information space with content. Additionally, Piala claims that Kyrgyzstan's many sockpuppet campaigns are aware of the machine learning-based methods of detecting bots, and that they are actively taking steps to undermine these. Again, choosing sockpuppets seems consistent with this.

4.3 "Every politician has a campaign"

Piala claims that "every politician" has a sockpuppet campaign, not only Chainik [37]. This is an impossible claim to independently verify. However, We have conferred with another confidential source as well as scoured Facebook, identifying possible sockpuppet campaigns of two other politicians. These fake accounts have many of the same behaviors and features as the campaign run by Chainik.

In terms of circumstantial evidence, recent developments appear to have further born out the whistleblower's claim. In November 2019, several local news agencies published a large-scale investigative report into the finances of Rayimbek Matraimov, an ex-customs official and behind-the-scenes "kingmaker". The investigation itself was inspired by a whistleblower – a money-runner for Matraimov – who was assasinated in Istanbul shortly before the reports were published [5]. A new movement called "RE:АКЦИЯ" ("Reation") was launched by small businessmen and intelligentsia to call for government action against Matraimov and his family. RE:АКЦИЯ organized two large protests in the capital, Bishkek, in November and December 2019.

Fake news-generating Facebook accounts and groups linked to Matraimov have been laboring to discredit the news agencies and RE:АКЦИЯ. Facebook group "Real Time 312", purporting to be an actual news agency, has been particularly intrepid at producing fake news-reports purporting evidence proving that the protests are actually part of an elaborate international conspiracy seeking to undermine Kyrgyzstan's sovereignty and culture. The entirety of Real Time 312 appears to be a sockpuppet, as the content is too dynamic to be algorithmically written. Indeed, several of the posts have been attributed to one author: "Vladimir Petrov", possibly named for the Russian antagonist from the television show, *House of Cards*.

4.4 Analysis

4.4.1 *Data Collection and Annotation Methods.*

For this project we utilize two datasources. First, we have been collecting data from Twitter via its API. Second, we have been manually collecting and annotating sample sockpuppets from Facebook. The goal of both of these collection methodologies has been the same: to root out inauthentic activity and study its effects. The methodologies themselves, however, are quite different.

We should first note that in many countries, including Kyrgyzstan, Facebook usage far exceeds that of Twitter. Yet, much of the literature on fake accounts and disinformation focuses on Twitter

because it is an easier platform from which to collect data. Facebook, however, offers richer data, as accounts on that platform are more dynamic. Combining the two platforms has enabled a fuller picture of the phenomenon of sockpuppets.

We have opted for a qualitative approach on Facebook in particular, for two reasons. First, as is well known, Facebook has made it intentionally difficult to collect data using automated means. Second, on philosophical grounds we challenge the notion that the "easiest" way to collect data is the "correct" way. Purely quantitative approaches do not yield the same depth of insight that qualitative approaches can.

While sacrificing speed and volume, manual collection and annotation on Facebook has enabled a more concentrated and targeted approach, with a more reliable ground truth than any automated method can offer. Meanwhile on Twitter, we have sacrificed the scalpel precision of manual research in favor of sheer volume. In particular, we have been focusing our collection upon tweets containing keywords about the October 2017 presidential election, with our time-range stretching back to April 2017, when the major parties formally announced their candidates. Our keyword list has been curated by the team of journalists and included the names of candidates and slogans in both Russian and Kyrgyz.

4.4.2 *Ethical Approval.*

Prior to starting, we underwent an extensive ethical review process from the university's ethical review board. We also underwent additional legal review in order to ensure compliance with both European and Kyrgyz laws regulating the use of personal identifiable information for both journalistic and machine-learning based research. Throughout this work, we have collected and analyzed only publicly-available data. We have collected data either through appropriate use of designated APIs or via manual reporting (i.e., reports of inauthentic behavior and accounts). Our work has involved no contact with the owners of the accounts aside from Piala, the whistleblower source.

5 CONCLUSION

Our research is still ongoing, but already it has yielded many important insights into how disinformation works from the inside. The combination of machine learning and investigative journalism have gained us something more than ground truth and important patterns of activity: it has gained a sense of the motivations and reasoning of those who engage in disinformation via fake accounts and fake news.

To close, we want to ring an alarm bell: the possibility of there being a next-generation threat arising not from bots – the traditional specter – but from the old-fashioned sockpuppet. As mentioned above, the sockpuppets to which Piala has focused our attention have a rather sinister task: to infiltrate audiences and manipulate them from within. We call this subset of sockpuppets "infiltrators", and we believe they represent an adversarial escalation of meatpuppets and Sybils [2, 14].

Sockpuppets were originally used to either a) praise, defend or otherwise support a person or organization, often for the purpose of manipulating an audience's opinion, or b) to circumvent a platform's restrictions on a person's behavior, particularly in cases of suspensions or bans [6]. Sockpuppets accomplish these tasks by typically posing as an independent third-party entity, wholly unaffiliated with its actual user [6]. Infiltrators go about things somewhat differently. We are seeing in Kyrgyzstan that, while they can and often do operate as traditional sockpuppets, infiltrators can also directly attack a target.

Based upon what we have seen with Real Time 312 and RE:АКЦИЯ, we may also reliably presume that infiltrators are pursuing the very narrow perceived self-interest of their controller, and not the interest of the audience. A philosophical way to think about why this is an important escalation in the sockpuppet concept is the following. Bots can be understood as *simulating* audiences, which is why their chief effect appears to be either to drive traffic or to drown out opposition, moderate or alternative voices [39]. In other words, bots do not really convince anyone of anything, they just yell in unison and extremely loudly. What infiltrators attempt to do is *assimilate* genuine audiences from within. According to Piala, the goal is nothing less than to persuade real people – and thus, ultimately, public opinion – of the message that this or that sockpuppet campaign is secretly attempting to promulgate [37]. All told, then, a significant element of deception and intention to commit harm is introduced by infiltrators.

At this juncture, were we to offer a schematic of infiltrator accounts based upon what our research, it would be the following:

- Infiltrator accounts are far less numerous than their bot counterparts, not only by necessity – as they depend upon real human beings creating and managing them, and hence their scalability becomes a function of the number and efficiency of real people – but also by design, for they are also careful to avoid the machine-learnable markers of forgery, such using as hashtags, re-posting content *verbatim*, and even direct linkages to each other within the online social network, such as "shares" of each other's content or establishing "friendships" with each other.
- In this latter respect, their preferred *modus operandi* is to either take over existing accounts – whether by cracking the password or, we suspect in some cases in Kyrgyzstan, paying "rent" to the real person – or engineering entirely new ones – typically by using the subscriber identity module (SIM) cards of elderly pensioners, their own relatives, or if circumstances permit, unregistered phone numbers. In this way, they appear within an online social network *ex nihilo*.
- They then proceed to actively develop personalities, replete with ethnicities and genders, even going so far as generating content that has nothing to do with their actual agenda but which gives them the appearance of full, regular identities, rather than the cycloptic, monomaniacal identities of bots.
- Finally, they proceed to target an existing constellation of accounts or community group, friending members who they identify as nodes within the social network linkages, and slowly entrenching themselves before finally directly connecting with each other – if they ever do.

Infiltrators will probably not altogether replace bots. However, as machine learning-based methods of bot detection slowly improve, the sun will gradually set upon the bot, and a new, dark dawn of the sockpuppet may rise.

ACKNOWLEDGMENTS

This work is supported in part by the Open Technology Fund's Information Controls Fellowship. The Civil Initiative for Internet Policies provides administrative, legal and logistical support. Our team of journalists is comprised of Rustam Khalimov, Arina Efremova, Rauf Akhmatov and Aziza Isakova.

REFERENCES

[1] 2019. CIA World Factbook: Kyrgyzstan Country Profile.
[2] Frank Ahrens. 2006. "'Puppets' Emerge as Internet's Effective, and Deceptive, Salesmen". *Washington Post* (2006).
[3] Nurlan Aliyev. 2019. "Protest Against Chinese Migrants in Kyrgyzstan: Sinophobia or Demands for Social Justice?". *CACI Analyst* (2019). https://www.cacianalyst.org/publications/analytical-articles/item/13568-protest-against-chinese-migrants-in-kyrgyzstan-sinophobia-or-demands-for-social-justice?.html
[4] DR Analytica. 2017. "2016 ICT Sector Overview for Kyrgyzstan". https://analytica.digital.report/wp-content/uploads/2017/07/Kyrgyzstan-The-2016-ICT-Sector-Overview.pdf
[5] Radio Azattyk, OCCRP, and Kloop.kg. 2019. "Plunder and Patronage in the Heart of Central Asia". (2019). https://www.occrp.org/en/plunder-and-patronage/
[6] Poland Bailey. 2016. *Haters: Harassment, Abuse, and Violence Online.* University of Nebraska Press.
[7] Ermek Baisalov. 2019. "How to Cope with Disinformation in Kyrgyzstan?". *CABAR.asia* (2019). https://cabar.asia/en/how-to-cope-with-disinformation-in-kyrgyzstan/
[8] Marco Bastos and Shawn T. Walker. 2018. "Facebook's data lockdown is a disaster for academic researchers". *The Conversation* (2018). http://theconversation.com/facebooks-data-lockdown-is-a-disaster-for-academic-researchers-94533
[9] BBC. 2017. "Kyrgyzstan Country Media Profile". (2017). https://www.bbc.com/news/world-asia-16187183
[10] Reporters Without Borders. 2020. "Absurd lawsuit against media outlets over corruption exposé in Kyrgyzstan". (2020).
[11] Andrew Chadwick, Cristian Vaccari, and Ben O'Loughlin. 2018. "Do Tabloids Poison the Well of Social Media? Explaining Democratically-Dysfunctional News Sharing". *News Media & Society* (2018).
[12] Yimin Chen, Niall J. Conroy, and Victoria L. Rubin. 2015. "Misleading Online Content: Recognizing Clickbait as 'False News'. *WMDD '15: Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection* (2015). https://doi.org/10.1145/2823465.2823467
[13] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec1. 2017. "Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions". *CSCW* (2017). http://dx.doi.org/10.1145/2998181.2998213
[14] John R. Douceur. 2002. "The Sybil Attack". In *Peer-to-Peer Systems (Lecture Notes in Computer Science).* https://doi.org/10.1007/3-540-45748-8_24
[15] Brigitte Dufour, Farid Tuhbatullin, and Asia-Pacific Human Rights Information Center. 2012. "Central Asia: Censorship and Control of the Internet and Other New Media". https://www.hurights.or.jp/archives/focus/section2/2012/03/central-asia-censorship-and-control-of-the-internet-and-other-new-media1.html
[16] Kamila Eshaliyeva. 2018. "Is anti-Chinese mood growing in Kyrgyzstan?". *openDemocracy.net* (2018). https://www.opendemocracy.net/en/odr/anti-chinese-mood-growing-kyrgyzstan/
[17] Massimo Flore, Alexandra Balahur, Aldo Podavini, and Marco Verile. 2019. "Understanding Citizens' Vulnerabilities to Disinformation and Data-Driven Propaganda". *JRC Technical Reports* (2019).
[18] Anonymous (for security reasons). 2019. "Kazakhstan Country Profile". In *Nations In Transit.* Freedom House.
[19] Franco Galdini. 2014. "The June 2010 'Events' Four Years On: Past, Present, Future".
[20] Zukhra Iakupbaeva. 2017. "Kyrgyzstan: A Haven for Reporters amid Love and Strife". *EurasiaNet.org* (2017). https://eurasianet.org/kyrgyzstan-haven-reporters-amid-love-and-strife
[21] Zukhra Iakupbaeva. 2018. "Minorities in Kyrgyzstan: changed by revolution". *openDemocracy.net* (2018). https://www.opendemocracy.net/en/odr/minorities-in-kyrgyzstan/
[22] Gulnara Ibraeva. 2019. "Kyrgyzstan Country Profile". In *Media Sustainability Index.* International Research and Exchanges Board. https://www.irex.org/sites/default/files/pdf/media-sustainability-index-europe-eurasia-2019-kyrgyzstan.pdf
[23] Garth Jowett and Victoria O'Donnell. 2005. "What Is Propaganda and How Does It Differ From Persuasion?". In *Propaganda and Persuasion.* Sage.
[24] Joshua Kucera. 2014. "U.S. Formally Closes Its Kyrgyzstan Air Base". *EurasiaNet.org* (2014). https://eurasianet.org/us-formally-closes-its-kyrgyzstan-air-base
[25] Bahtiyar Kurambayev, Laura Schwartz-Henderson, and Ken Winneg. 2018. "The Spiral Of Silence on Social Media: Cultures of Self-Censorship Online and Offline in Kyrgyzstan". *Internet Policy Observatory* (2018). https://globalnetpolicy.org/spiral-of-silence-kyrgyzstan/
[26] Peter Leonard. 2019. "Kyrgyzstan: Life in jail looming for ex-president". (2019).
[27] Alison MacKenzie1 and Ibrar Bhatt. 2018. "Lies, Bullshit and Fake News: Some Epistemological Concerns". *Postdigital Science and Education* (2018). https://doi.org/10.1007/s42438-018-0025-4
[28] Meerim Maturaimova. 2015. *"Memory of Territory as an Ethnic Narrative: Kyrgyz and Uzbek Narratives in Kyrgyzstan".* Master's thesis. Central European University.
[29] OSCE-ODIHR. 2018. "Final report on Kyrgyzstan presidential election". (2018). https://www.osce.org/odihr/elections/kyrgyzstan/374743
[30] G Pennycook and DG Rand. 2018. "Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking". *Journal of Personality* (2018). https://doi.org/10.1111/jopy.12476
[31] G Pennycook and DG Rand. 2019. "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning". *Cognition* (2019). https://doi.org/10.1016/j.cognition.2018.06.011
[32] Katharine Quinn-Judge and Paul Stronski. 2016. "Kyrgyzstan at Twenty-Five: Treading Water". Carnegie Endowment for Peace. https://carnegieendowment.org/2016/07/21/kyrgyzstan-at-twenty-five-treading-water-pub-64152
[33] Meet Rajdev and Kyumin Le. 2015. "Fake and Spam Messages: Detecting Misinformation During Natural Disasters on Social Media". *2015 IEEE / WIC / ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (2015). https://doi.org/10.1109/WI-IAT.2015.102
[34] Christopher Rickleton. 2020. "China's virus storm over Central Asia: Panic is being sown by trouble-making social media users.". *EurasiaNet.org* (2020). https://eurasianet.org/chinas-virus-storm-over-central-asia
[35] Christopher Rickleton and Nurjamal Djanibekova. 2019. "Internet provides new space for Kyrgyzstan's north-south divide: Online disinformation is manipulating regional divisions.". (2019).
[36] Christopher Schwartz. 2015. Field notes.
[37] Christopher Schwartz. 2019. Piala interview notes.
[38] Christopher Schwartz and Alisher Khamidov. 2016. "Kyrgyzstan: Corrupt, Anarchic – and Stable?". *The Diplomat* (2016).
[39] Christopher Schwartz and Rebekah Overdorf. 2019. "Subtle Censorship Via Adversarial Fakeness in Kyrgyzstan". In *19th Privacy Enhancing Technologies Symposium.* https://arxiv.org/abs/1906.08021
[40] Inga Sikorskaya. [n.d.]. "A brief history of conflict in Kyrgyzstan". *Peace Insight* ([n. d.]). https://www.peaceinsight.org/blog/2015/09/a-brief-history-of-conflict-in-kyrgyzstan/
[41] Sputnik.kg. 2019. "Страница в поддержку Кашкара Джунушалиева появилась в Instagram". (2019). https://ru.sputnik.kg/society/20190817/1045465013/kashkar-dzhunushaliev-instagram-podderzhka.html
[42] Edson Tandoc, Zheng Wei Lim, and Rich Ling. 2017. Defining 'Fake News': A typology of scholarly definitions. *Digital Journalism* (2017). https://doi.org/10.1080/21670811.2017.1360143
[43] Committee to Protect Journalists. 2007. "Alisher Saipov Profile". (2007).