UNIVERSITY OF COPENHAGEN
FACULTY OF SOCIAL SCIENCES

Pro-Kremlin Disinformation on Social Media

PhD Dissertation 2020

Yevgeniy Golovchenko

DEPARTMENT OF POLITICAL SCIENCE
FACULTY OF SOCIAL SCIENCES
UNIVERSITY OF COPENHAGEN
PHD DISSERTATION 2020 · ISBN 978-87-7209-378-9

YEVGENIY GOLOVCHENKO

Pro-Kremlin Disinformation on Social Media

SL grafik

PhD Dissertation

Pro-Kremlin Disinformation on Social Media

Author:
Yevgeniy Golovchenko

Supervisor:
Rebecca Adler-Nissen

This dissertation is submitted for the degree of Doctor of Philosophy

Department of Political Science
University of Copenhagen

June 15, 2020

ISBN 978-87-7209-378-9 (Printed book)
ISBN 978-87-7209-381-9 (E-book)
Printed by SL grafik, Frederiksberg, Denmark (slgrafik.dk)

Acknowledgements

This dissertation would not have been possible without a supporting community of scholars, friends, family and many other people who helped me along the way.

I would first and foremost like to thank my supervisor Rebecca Adler-Nissen. Rebecca has supported me from the beginning to the end, has offered me invaluable feedback on this dissertation. She taught me what it meant to be a scholar – through her precious advice, and by way of her own example as an innovative and supportive member of the scholarly community. Most importantly, my supervisor believed in me throughout my time as a PhD student. Words cannot do justice to the gratitude that I would like to express for this support.

Sune Lehmann Jørgensen also played an important role as my mentor. Sune gave me crucial feedback on all of the dissertation papers as well as more general guidance throughout the entire research process. In many ways, my conversations with Sune have been my gateway to computational social science. He is, without exaggeration, a driving force behind the discipline in Denmark and an inspiration to many scholars around the world. He has inspired me in my own career path and has accelerated my own interest in applying the insights from other disciplines within social science.

This dissertation is part of a larger research project Digital Disinformation, that is led by Rebecca as Principal Investigator and funded by the Carlsberg Foundation (project number CF16-0012). This dissertation exists thanks to the foundation's support.

While writing this thesis I learned that research is not just solitary work. Some of the most important insights only come through engaging with people from different scholarly communities and parts of the world. Thanks to generous opportunities provided by the Carlsberg Foundation and the Department of Political Science, I had the privilege to participate in conferences and to do research abroad. I would particularly like to thank Joshua A. Tucker for hosting me during my research stay at New York University. Not only did he help me settle in a new environment, he also made me a welcomed member of the community at the SMaPP lab (now CSMaP) and the Jordan Center for the Advanced Study of Russia. This resulted in an entire project on Russian interference in the US elections, and a research paper in this dissertation co-authored with Joshua A. Tucker, Cody Buntain, Megan Brown and Gregory Eady. I am very grateful to

the co-authors, as well as to the entire SMaPP community that supported this research. This includes Leon Yin, Nicole Baram, Sergey Sanovich, Richard Bonneau and many others.

I would also like to express my gratitude to Chiyuki Aoi at Tokyo University and Yasuyuki Todo at Waseda University for hosting me during my research stays in Japan. Both offered me enormous support and contributed to my intellectual growth. Here I had the privilege of developing research ideas in long and thought-provoking conversations with Yee-Kuang Heng, Vindu Mai Chotani, Paul Nadeu, Cesare M. Scartozzi, Shinichiro Asayama, Makoto Takahashi and Yuzuka Kashiwagi, who also made my stay at Waseda University possible.

I would like to thank members of the pre-defence committee, James Pamment and Asmus Leth Olsen, who put in many hours reading and commenting my drafts. I would also like to express my gratitude to Alicia Fjällhed for joining the pre-defence seminar and offering important feedback. Their feedback resulted in many crucial changes that drastically improved the quality of the dissertation frame.

I would also like to extend a special thanks to Isabelle Augenstein who played an important role in co-developing the code-book for paper A, Mareike Hartmann, the co-author for paper B and Alan Mislove for providing Twitter data for both of the papers. I would also like to thank Fanny Marie Brændholt Sørensen who assisted with a lot of the research and Sille Obelitz Søe, whose insightful work has heavily influenced my understanding of 'disinformation' as a theoretical concept. I am grateful to Michael Bang Petersen, Jason Reifler, Frederik Hjorth, JungHwan Yang, Mikhail Suslov, Frederico Botta, Amelia H. Arsenault, Jonas Gejl Pesersen, Yi Ma and Benjamin Carl Krag Egerod for being discussants at various seminars or conferences, where they gave thorough and helpful feedback on my dissertation papers. I would like to thank Margaret E. Roberts for taking the time to give important feedback on the part of my research that is related to censorship.

I would like to thank my sister and mother as well as Anina Lyck Uldum's family for their support. Cathrine Rasmussen helped me enormously by giving me feedback on a large part of my dissertation frame. Mike Young provided proofreading services for most of the dissertation frame. Lene Hansen gave me detailed feedback on multiple research papers, offered valuable advice on many issues (small and large) and showed me continuous support, for which I am immensely grateful. I have had the privilege to discuss and develop research ideas

with excellent scholars like Patrice Wangen, Kristin Anabel Eggeling, Anne Bach Nielsen, Hjalte Bonde Meilvang, Thorsten Brønholt, Deven Parekh, Sofiya Voytiv, Porey Lin, Christina la Cour, Bertel Teilfeldt Hansen, Johan Farkas, Michael Bosetta, Livia Rohrbach, Minda Holm, Christien Santos Rasmussen and many others. Vera Skvirskaja has greatly helped me understand the nuances of online propaganda. I am also heavily indebted to Valentina Shapovalova, who helped me by giving invaluable feedback to the dissertation frame and articles, as well as annotating parts of the data. She was also a source of many insights on pro-Kremlin disinformation.

I am indebted to the kind and helpful community of scholars as well as administrative staff at the Department of Political Science and the Copenhagen Center for Social Data Science. You are the reason why I enjoyed the last three and a half years. A special thanks goes to my former and current officemates, with whom I had the pleasure to discuss most aspects of my research and life as a PhD student: Rune Møller Stahl, Xuan Li, Jens Adriaan van der Ploeg, Simon Polichinel von der Maase and Marie Heldgaard Seestedt.

Lastly, I would like to thank my partner, Anina Lyck Uldum. Anina helped me with many aspects of this dissertation: From feedback on individual chapters and articles, to long conversations about new research ideas. Not only for this dissertation, but also for my master's thesis on the war in Ukraine, which, in many ways laid the foundation for this research. I am very grateful to have had this support throughout these years (and on three different continents). Perhaps most importantly, I would like to thank Anina for giving me a reason to come home and realize that academia, while important, is not the only thing that matters in life.

Preface

This article-based dissertation is a response to the following research question:

> *What is the nature and extent of pro-Kremlin disinformation and counter-disinformation on social media?*

It consists of a dissertation frame (chapters 1-6) and the following four research papers (chapters 7-10). The research papers can be read independently:

A. Measuring the Scope of Pro-Kremlin Disinformation on Twitter.

B. State, Media and Civil Society in the Information Warfare over Ukraine: Citizen Curators of Digital Disinformation.

C. Cross-Platform State Propaganda: Russian Trolls on Twitter and YouTube during the 2016 US Presidential Election.

D. Fighting Propaganda with Censorship: A Study of the Ukrainian Ban on Russian Social Media.

Paper B has been published in *International Affairs*. Paper C has been accepted in *International Journal of Press/Politics* and made available in *OnlineFirst*.

The purpose of the dissertation's 'frame' is to put the papers in context in multiple ways. It introduces the main themes and shows the relevance of the research. It also offers a broader overview and discussion of relevant literature and core theoretical concepts (such as 'disinformation'), than what is possible in the papers alone. Lastly, it presents the core findings of the papers in a distilled form, and discusses the results and implications.

Contents

Chapter 1
Introduction

This article-based dissertation examines the scope and nature of pro-Kremlin disinformation and counter-disinformation on social media during international conflict. It explores how different actors produce and counter (dis)information in the context of the deteriorating relations between Russia and the West as well as the war in Ukraine.

I will begin by offering background on one of the most significant events of the ongoing conflict and the dissertation's first case: Russia's annexation of Crimea in 2014. I will then introduce the debate on the nature and scope of pro-Kremlin disinformation and counter-disinformation, following the Western response to Russian 'information warfare'. I will proceed to delimiting the scope this dissertation, before introducing its empirical cases and papers. As I will show in the subsequent section, the dissertation's main contribution is its empirical findings related to a debate on pro-Kremlin disinformation that has largely lacked systematic evidence.

Russian interference in Ukraine and disinformation

On 27 February 2014, Russian soldiers without insignia appeared in the streets of Crimea in Ukraine.[1] As pro-Russian protests erupted in the region, the Russian military seized local municipality offices, the airport, the TV tower and other critical infrastructure. The unidentified troops surrounded Ukrainian military bases, raising tensions and stoking fears of a massacre.

On the international stage there was a state of confusion throughout the military spectacle. Should NATO confront Russia to stop the military operation? Should Ukrainian troops open fire on sight? Russian officials and the Russian state-controlled media claimed that Russia was not involved in the conflict and

[1] Parts of this section are based on the 'Background' section in Paper A.

therefore not responsible. According to this narrative, the unidentified armed men did not belong to the armed forces of the Russian Federation. Instead, they were described as a local self-defence force, or disenchanted police officers (Schreck, 2019). The 'little green men' were there to protect Crimeans from the Ukrainian 'Fascist Junta', the new pro-Western government in Kyiv.

After being captured by masked Russian soldiers, the Crimean parliament hastily proclaimed a 'referendum' on 16 March on whether Crimea should become part of the Russian Federation.[2] A few days later, Russia annexed Crimea – within just three weeks after the appearance of armed men without insignia. One month after the annexation, President Putin retracted his own disinformation narrative by admitting that the soldiers were indeed Russian (RT, 2014).

The Kremlin's theatrical play-out of the crisis has become a classic example of a massive disinformation campaign (Thornton, 2015). Commentators have argued that Russia used disinformation in Ukraine and the West to cover the events with a veil of confusion, and to mobilise local support for Russia during the most critical phase of the military operation. This has arguably helped the Russian armed forces to successfully capture an important part of Ukrainian territory with minimal military casualties (K.N.C., 2019; Snegovaya, 2015).

The Russian Annexation in 2014 served as a historical turning point and marked the beginning of a new international conflict. A conflict that came to involve liberal democracies around the world. The conflict's global aspect became even more apparent when Russian-backed separatists in Ukraine used a Russian land-to-air missile to shoot down Malaysia Airlines Flight 17 (MH17) in July 2014, killing 298 civilians on-board, a majority of whom were Dutch citizens (Joint Investigation Team (JIT), 2018; Escritt, 2015). Just like the Crimean annexation, the event was accompanied by a wave of pro-Kremlin disinformation (Oates, 2016) and led to sanctions against Russia. Following the war in Ukraine, the Western public did not just perceive post-Soviet Russia as a nuclear power, but also as a master of political deception. A regime that could use propaganda and disinformation to support military operations on the ground, influence foreign elections, and potentially destabilise entire countries (Frum, 2018; Wallance, 2018; Lockwood, Alisa, 2018; Lockie, 2017; Torossian, 2016; Pomerantsev, 2015).

[2]While Crimean authorities claimed that the referendum legitimately expresses the will of the people to join Russia, the UN passed a resolution stating that the referendum had "...no validity, cannot form the basis for any alteration of the status of Crimea" (Charbonneau and Donath, 2019).

It is important to note that Russia is not the only country to strategically exploit social media through manipulation or 'non-lethal warfare'. In fact, multiple government entities around the world use fake or partially fake accounts in attempting to influence online debates (Bradshaw and Howard, 2019). This includes countries like China (Han, 2015), South Korea (Keller et al., 2019), UK (MacAskill, 2015), Ukraine (Kottasova, 2015) and Israel (Daro, 2018). While authoritarian states have used social media to manipulate domestic audiences for some time (King, Pan, and Roberts, 2017; Keller et al., 2017; Morozov, 2012), more recent reports suggest that different state and non-state actors in China and other authoritarian states are following Russia's lead by also targeting *foreign* audiences online (Yang, 2019; Zhang, 2019). Twitter, for example, identified a network of state-backed fake accounts which were engaging in debates about protests in Hong Kong (Twitter, 2019). Many of their Tweets were in English, Indonesian, Portuguese, Spanish and other languages (Uren, 2019), suggesting that they were trying to manipulate a global audience, rather than just users in Hong Kong or China.

This pattern is not surprising, considering that online disinformation – in the form of fake accounts – can be easily produced at a relatively low cost. As several commentators have noticed, disinformation agents do not act in isolation, they are likely to improve their covert operations by learning from each other (Polyakova and Meserole, 2019). By understanding pro-Kremlin disinformation campaigns, scholars may better understand operations run by other actors, such as those run by the Chinese state or patriotic grassroots (Yang, 2019; Zhang, 2019). Perhaps even more importantly, understanding Russian disinformation may help decision-makers and the media prepare for future campaigns. While each country may develop its own approach to online manipulation, Russia's use of (dis)information in relation to Crimea (and later MH17) stands as an important example for how such campaigns may unfold.

Pro-Kremlin disinformation and the West

Both scholars and the broader public in the West have mostly understood the Crimean annexation, as well as pro-Kremlin disinformation in Ukraine, as 'information warfare' (Hoskins and Shchelin, 2018; Tanchak, 2017; Niekerk, 2015; Snegovaya, 2015; Darczewska, 2014). The Supreme Allied Commander Europe, General Philip Breedlove, went as far as describing the Russian operation in Crimea as "...the most amazing information warfare blitzkrieg we have ever seen

in the history of information warfare" (Vandiver, 2014). This relatively state-focused concept refers to the state's manipulative use of information for the purpose of achieving military and political goals (Thornton, 2015; Taylor, 2003; Myriam Dunn Cavelty and Mauer, 2008). When combined with military operations, Kremlin's strategic use of information is often referred to as 'hybrid warfare' (Lanoszka, 2016; Woo, 2015; Thiele, 2015; Reisinger and Gol'c, 2014a).

Both scholars and commentators perceived the 'information war' in Ukraine as a foreshadowing of Russia's use of disinformation to interfere in Western elections. This was shown to be the case, for example, when the (Russian state-affiliated) Internet Research Agency, popularly known as the 'Russian troll factory', launched a covert propaganda campaign through fake social media accounts – often posing as American users or local news outlets – in an attempt to influence the 2016 US presidential elections (Bastos and Farkas, 2019; Slutsky and Gavra, 2017; Xia et al., 2019; Zannettou et al., 2019a).

These events have been accompanied by an enormous public interest in Russian disinformation and influence campaigns. In the period 2014-2019, there were at least 24,437 English-language news articles from Northern Europe and North America alone that mentioned Russia together with 'fake news', 'misinformation' or 'disinformation', often intertwined with controversies related to Donald Trump. While the interest peaked in 2017 – following the November 2016 US presidential election – the interest in the topic continued to be relatively high throughout 2019, as shown in Figure 1.1.

The debate is predominantly shaped by theoretical assumptions and journalistic accounts. It is driven by at least three assumptions. First, media, decision makers and scholars often implicitly describe pro-Kremlin disinformation campaigns as entirely state-driven (Zannettou et al., 2019b; Unver, 2019; Frum, 2018; Thornton, 2015). This view is in line with state-focused theories on Russian information warfare, as well as the broader literature on political propaganda. Second, journalists and experts often assume that Russian disinformation has a vast reach in the West, overwhelming the online ecosystem with a 'disinformation flood' or even winning an information war against the West (Wallance, 2018; Lockwood, Alisa, 2018; Lockie, 2017; Torossian, 2016). Lastly, Russia's use of political disinformation is viewed as a strategy to support right-wing movements abroad or to weaken societies from within by supporting opposing sides on the ideological spectrum. The purpose of this dissertation is to challenge these views empirically by measuring the actual scope and nature of pro-Kremlin disinformation.