Bohdan Andrusyak, BSc

Principle-Guided Propaganda Analysis - Case Study on Russian Military Intervention in Ukraine

Master's Thesis

to achieve the university degree of

Master of Science

Master's degree programme: Software Engineering and Management

submitted to

Graz University of Technology

Supervisor

Dr. Roman Kern

Institute for Interactive Systems and Data Science
Head: Univ.-Prof. Dipl-Inf. Dr. Stefanie Lindstaedt

Graz, May 2019

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date

Signature

Abstract

Propaganda is one of the biggest problems in the modern world because it provokes conflicts which can lead to a great loss of human life. The annexation of Crimea and following conflict in Eastern Ukraine is a prime example of it. This conflict lead to thousands of lost lives and millions of displaced people. The lack of research on the topic of unsupervised propaganda detection led us to devise methods for analysing propaganda that does not rely on fact checking or makes use of a dedicated ground truth. Instead, we base our measures on a set of guiding principles that constitutes the intention of an propagandist authors. For each of these principles we propose techniques from the fields of Natural Language Processing and Machine Learning. We have chosen the Russian military intervention in Ukraine as our focus, and the Russian News and Information Agency as our data source. We found the representation of Ukraine to be remarkably different to other countries, hinting that the principles of propaganda might be applicable in this case. Our quantitative analysis paves the way to more in-depth qualitative analysis.

Contents

Abstract **iii**

1. Introduction **1**

2. Background **3**
- 2.1. Propaganda 3
- 2.2. Information Warfare 5
- 2.3. Russian Hybrid Warfare 8
  - 2.3.1. Annexation of Crimea 8
  - 2.3.2. Russian strategy 10
  - 2.3.3. Understanding Russian Propaganda 12
  - 2.3.4. Principles of Propaganda 14
- 2.4. Machine Learning 16
  - 2.4.1. Random Forest 16
  - 2.4.2. Neural Network 17
  - 2.4.3. Evaluation techniques 17
- 2.5. Nature Language Processing 19
  - 2.5.1. Statistical Language Modeling 20
  - 2.5.2. Neural Language Modeling 21
  - 2.5.3. Word2Vec 22

3. Related Work **25**
- 3.1. Extremist Propaganda Detection 25
- 3.2. Russian Propaganda Detection 27

4. Methodology **29**
- 4.1. Finding Propagandistic Articles 29
  - 4.1.1. Detecting Shift in Narrative 29
  - 4.1.2. Extracting propagandistic articles 30

4.2. Analyzing Principles of Propaganda 32
4.2.1. Principle of Long Lasting Impact 32
4.2.2. Principle of Desired Information 32
4.2.3. Principle of Emotional Agitation 34
4.2.4. Clarity Principle 34
4.2.5. Principal of Supposed Obviousness 35
4.3. Influence on Other Countries 36

5. Dataset **37**
5.1. Data Collection 37
5.2. Data Preprocessing 39
5.3. Exploratory Data Analysis 40

6. Results **45**
6.1. Shift in Narrative 45
6.2. Extracting suspected articles 47
6.3. Investigating Principle of Long Lasting Impact 49
6.4. Investigating Principle of Desired Information 50
6.5. Investigating Principle of Emotional Agitation 53
6.6. Investigating Clarity Principle 53
6.7. Investigating Principle of Supposed Obviousness 56
6.8. Measure Influence on Other Countries 60

7. Discussion **61**

8. Conclusion **65**

Bibliography **67**

A. RIA News Statistic **77**

B. Results **81**

List of Figures

2.1. Continuous Bag-Of-Word (CBOW) architecture of word2vec model 22
2.2. Skip-gram architecture of word2vec model 23

4.1. The process of detecting shift in media narrative 30
4.2. The process of detecting suspected articles 31
4.3. The approach for measuring the long-lasting impact 33

5.1. Homepage of RIA News 38
5.2. Amount of articles published by RIA News per year 40
5.3. Amount of articles published by RIA News each year that mention Ukraine in absolute numbers and its relative share based on all articles published in the respective year. 41
5.4. Amount of articles per day, mentioning Ukraine published by RIA News in 2014. 42
5.5. Amount of articles per day, mentioning Ukraine and word ”fascism” together, published by RIA News in 2014. 43

6.1. Similarity between vector-words ”State”, ”Country” and ”fascism” 46
6.2. Changes in similarity between country names and word ”fascism”. 47
6.3. Distribution of articles per topic 48
6.4. Results of classifier bias 51
6.5. Average amount of views per article 52
6.6. Average percentage of negative words per article 54
6.7. Result of Flesh reading-ease score and LIX 55
6.8. Average amount of words per article 56
6.9. Average amount of unique stems per article 57
6.10. Value of parameter β in Heaps’ law for the articles 58
6.11. Percentage of articles that contain the term ”Zapad” in the text . . 59
6.12. Measurements for the similarity between country name and fascism 60

1. Introduction

So long as there are men, there will be wars.

Albert Einstein

In 2014 world faced unprecedented event – the forceful change of borders of the independent country without a declaration of war, when Russian Federation annexed Ukrainian peninsula Crimea. Occupation of Crimea and following Russian military innervation in Eastern Ukraine caused thousands of lost lives and millions of displaced people (Office for the UN High Commissioner for Human Rights (OHCHR) 2018). This conflict is not like any conflict before and was named ”Hybrid War” (Rojansky & Kofman 2015). Where ”hybrid” part of the term denotes a combination of previously defined types of warfare: conventional, irregular, political and information (Rojansky & Kofman 2015). Before the annexation of Crimea, all Russian federal television and radio channels, newspapers and a multitude of online resources have been employed in disinformation campaign regarding the situation in Ukraine (Darczewska 2014). This large scale campaign led to the division of Ukrainian society into two groups: Ukrainian speaking and Russian speaking. Consequent antagonising of those groups and creating a fictional thread to the Russian speaking population, resulted in the fact that the victims of aggression (citizens of Crimea) did not resist (Darczewska 2014). Furthermore, people who had undergone necessary psychological and informational treatment took part in the separatist movement which led to appearing of terrorist organisation: ”Donetsk People’s Republic” and ”Luhansk People’s Republic” (Verkhovna Rada of Ukraine 2015).

The fact that as of May 2019 Crimea is still under control of Moscow and Eastern parts of Ukraine are submerged into the war with pro-Russian terrorists is the prime example of successful implementation of information warfare tools. Furthermore, taking to consideration constant thread of Russian aggression towards

Baltic states (Thornton & Karagiannis 2016), Kazakhstan and Poland (Rojansky & Kofman 2015) leads to the increasing demand for means of counteraction. It is important to understand that the most media sources in Russia are under control of government (*Russia country profile* 2012) and, as Fedor (2015) showed it, all forms of media were working together to create a negative image of Ukraine. Thus, it is impossible to cover in one work all means of propaganda in different media. For this reason, this thesis is focused on detecting propaganda in textual forms, mainly in online newspapers. Furthermore, as propaganda played an important role in the annexation of Crimea (Darczewska 2014, Fedor 2015), the main focus of the research is on the year 2014. Based on this, the goal of this thesis is to answer the following research questions:

- Based on what measurements newspaper article can be considered propagandistic?
- What methods of Natural Language Processing and Machine Learning can be used for analysing propaganda?

2. Background

2.1. Propaganda

> A lie that is half-truth is the darkest of all lies.
>
> Alfred Tennyson

Throughout history, propaganda has been a powerful weapon. Propaganda was instrumental to the Thirteen Colonies gaining independence. At the close of the First World War, many Germans concluded that British propaganda had contributed significantly to their defeat. At Nuremberg, the Allies charged the Nazis with poisoning the minds of the German people with propaganda, sparking the Second World War and inciting genocide (Macdonald 2014). All of this lead to a great amount of research in different aspects of propaganda, which resulted in the creation of a wide variety of definitions.

Smith & Lasswell (1946) introduced one of the first postwar definitions of propaganda is one means by which large numbers of people are induced to act together. This definition clearly shows that the propaganda aims at masses. In the following years, French philosopher Jacques Ellul defined propaganda as a form of information that panders to our insecurities and anxieties (Ellul 1973). Importance of this definition is in the fact that the author recognises that propaganda aims at creating negative feelings and anxiety. Both of this definitions lacked the precise description of the goal of propaganda, that is why Jowett & O'Donnell (1986) in their definition stated that propaganda is the deliberate, systematic attempt to shape perceptions, manipulate cognitions, and direct behaviour to achieve a response that furthers the desired intent of the propagandist. In the 1996 previous definition was enriched by Nelson (1996), who indicated that propaganda is a form of purposeful persuasion that attempts to influence the emotions, attitudes, opinions, and actions of specified target audiences for ideological, political or commercial purposes through the controlled transmission of one-sided messages (which may or may not be factual) via mass and direct media channels. Even though from previous definitions, it is hard to understand the relation between propaganda and truthfulness, but this question was answered by Cunningham (2002): ”Propaganda is indifferent to truth and truthfulness, knowledge and understanding; it is a form of strategic communication that uses any means to accomplish its ends”.

Aforementioned definitions were summarized by Bachrach et al. (2009): ”Propaganda appears in a variety of forms. It is strategic and intentional as it aims to influence attitudes, opinions and behaviours. Propaganda can be beneficial or harmful. It may use truth, half-truths or lies. Propaganda taps into our deepest values, fears, hopes and dreams, to achieve the maximal result.” Due to the fact, that given definition describes the goals and methods of propaganda, we are using it as the main definition of propaganda in this research.

The most effective propaganda combines entertainment and education. The entertainment elements attract the audience, while the educational aspect decreases the perception that the message is propaganda, it persuades. Unlike education, which seeks to present an objective view, propaganda is biased through the particular use of facts (Macdonald 2014).

Depending on the source, propaganda can be white, black or grey. In white propaganda, the source is known and is usually official. In black propaganda, it is concealed, and a false source, and in grey propaganda the source is obscured. Black propaganda is more difficult to create than white propaganda because white propaganda can contain mistakes and still be effective since it is known to be from a foreign source, whereas black propaganda cannot contain certain types of errors (Macdonald 2014). For example, during the Second World War, the British operated a radio station that pretended to be operated by loyal Germans broadcasting from inside occupied Europe. While Germans might expect official British messages to contain inaccurate information about Germany that any German would know, the same error from a station pretending to be German would cripple the station’s credibility (Macdonald 2014).

The influence of propaganda had in both World Wars made it a crucial part of all military strategies for all future conflicts. With the rise of informational systems in the second part of the 20s century, propaganda became one of the tools of newly emerged type of war – informational warfare.

2.2. Information Warfare

> All warfare is based on deception. Hence, when we are able to attack, we must seem unable; when using our forces, we must appear inactive; when we are near, we must make the enemy believe we are far away; when far away, we must make him believe we are near.
>
> Sun tzu, The Art of War

The term ”information warfare” started appearing in the last decade of the 20s century, since information and information technologies became crucial for state security and defence. For a better understanding of what ”information warfare”, first, we are defining the term ”warfare”. Warfare is the set of all lethal and nonlethal activities undertaken to subdue the hostile will of an adversary or enemy (Szafranski 1997). According to Colonel Szafranski warfare is not same as war, because warfare does neither require a declaration of war, nor does it require the existence of a widely recognised as ”a state of war”, and warfare is not necessarily aimed at killing the enemy but on suppressing the enemy (Szafranski 1997). Informational warfare is part of conventional warfare which is aimed at subduing enemy by disrupting or destroying an enemy's knowledge and beliefs.

Libicki & University (1995) outlined seven different forms of information warfare, which include:

1. **Command-and-Control Warfare**. The goal of this strategy is to disrupt knowledge and information transfer of enemy by decapitating the enemy's command structure from the body of command forces. Which can be achieved by physically liquidating the enemy's commanders or by interfering with enemy's communication systems.

2. **Intelligence-Based Warfare**. Intelligence-based warfare has to application offensive and defensive. The idea of the offensive idea is to maximise utilization of distributed information systems for enemy detection (e.g. a target is detected through a fusion of sensor readings). On the contrary, defensive one is aimed at preserving invisibility and untraceability of own forces in cyberspace and battlefield.
3. **Electronic Warfare**. The operations of electronic warfare are aimed at destroying the enemy's information systems, mainly communications. This goal is achieved by utilising radio-electronic to degrade the physical basis for transferring information and cryptography to intercept messages in cyberspace.
4. **Psychological Warfare**. The psychological warfare encompasses the use of information against the human mind. The authors divide psychological warfare into four categories:

    - operations against the national will
    - operations against opposing commanders
    - operations against troops
    - cultural conflict

    The key factor of counter-will is to breaks enemy's will to fight either by velvet glove ("accept us as friendly") or the iron fist ("or else"). Authors recognise global broadcasting media as a powerful tool which can change people perception of an ongoing conflict. The operations aimed at enemy commanders have the goal of confusing and disorienting them which will lead to non-optimal decisions on their side. Cultural warfare is changing values and beliefs of the opposing nation to make them less hostile and more open to cooperation.
5. **Hacker Warfare**. The operations of hacker warfare are taking place mostly in cyberspace. The intent of an attack can range from total paralysis to intermittent shutdown, theft of information, theft of services, illicit systems' monitoring and intelligence collection, the injection of malicious message traffic, and access to data for blackmail.
6. **Economic Information Warfare**. This type of warfare is based on using methods that are used in economic warfare such as blockade. The prominent strategy of economic information warfare is information blockade. The effectiveness of an information blockade presumes an era in which the well-being of societies will be as affected by information flows of material supplies. The

authors theorise that nations would strangle others' access to external data which would lead to crippling the economies of those nations. The use of psychological methods against the other side's forces offers variations on two traditional themes: fear of death and potential resentment between the trench and the home front. The authors predicted that this type of warfare would be most powerful when information technology can broadcast threats or resentment-provoking information to individual opposing troops.

7. **Cyberwarfare**. According to authors, cyberwarfare of the seven forms of information warfare, the least tractable because by far the most fictitious, differing only in degree from information warfare as a whole. Cyberwarfare is a broad category that includes information terrorism, semantic attacks and simula-warfare. Information terrorism is a subset of computer hacking, aimed not at disrupting systems but at exploiting them to attack individuals. The semantic attack is also using tools of hacker warfare to make attacks on information systems in which affected the system will be perceived as operating correctly, but it will generate results that are different from reality. The idea of simula-warfare is to simulate conflicts on computers to develop better plans for real-life conflict.

Libicki & University (1995) concluded that the most applicable methods of information warfare are all tools of Command-and-Control Warfare, Electronic Warfare, Intelligence-Based Warfare and psychological operations against commanders and forces. Authors concluded that psychological operations against the national will and culture are arguably forms of warfare, they came to this conclusion because most of the biggest global broadcasting services are independent and do not take sides in conflicts. However, it is not the case of Russian, where the biggest mass media company are owned by the government or pro-government organisations (*Russia country profile* 2012), which makes these forms of informational warfare much more potent. Moreover, research by Libicki & University (1995) was done before September 11 attacks, because in 2014 "The 9/11 Commission Report" reorganised the power of propaganda and recommended the use of various types of psychological warfare to improve the image of the USA in the world (Macdonald 2014). Thus making all form of psychological warfare as applicable as other previously mentioned forms.

Moreover, the rapid growth of networking and the Internet, which allowed people to communicate and transfer information in rates never seen before, have an enormous effect on informational warfare. Traditional information warfare topics such as offensive and defensive operations, espionage, ethics and legalities, propaganda and intelligence remain as crucial as before although there are some new additions. Such as cybersecurity, critical infrastructure protection, cyberterrorism, technology convergence, individual warfare and space (Williams 2010). It caused changes in the goals of information warfare, where traditional information warfare was aimed at physical protection of the nation, and modern have a stronger focus on defending critical infrastructures to protect nation, society and people (Williams 2010).

2.3. Russian Hybrid Warfare

2.3.1. Annexation of Crimea

> I will say this clearly: There are no Russian troops in Ukraine
>
> — Vladimir Putin

The annexation of Crimea shocked the whole civilised world, grabbing the attention of the scientific community. The ongoing occupation is well documented and described in many works and was defined as ”hybrid war”. The ”hybrid” aspect of the term denotes a combination of previously defined types of warfare, whether conventional, irregular, political or information (Rojansky & Kofman 2015). For a better understanding of how that warfare where combined, a summary of events that lead to the occupation is needed. According to Rojansky & Kofman (2015), since February 2014 Russia has conducted two distinct phases of operations in Ukraine, beginning with the occupation and annexation of Crimea, and continuing with the invasion of Eastern Ukraine’s Donbas industrial region. Crimea began as a covert military operation, combining ambiguity, disinformation, and the element of surprise at the operational level with more traditional aids such as electronic warfare. The annexation was completed by a traditional military invasion and occupation of the peninsula, using Russia’s airborne, naval infantry, and motor rifle brigades. This operation was unique, because Russia’s Sevastopol naval base, the status of forces arrangements in Crimea, and additional agreements on the transit of troops in Ukraine enabled deployments and tactics that would not otherwise have been possible. Furthermore, in 1997 Ukraine and Russian Federation signed a ”Treaty on friendship, cooperation and partnership between Ukraine and the Russian Federation” (Ukraine and Russian Federation 1997), where Article 2 states: ”High Contracting Parties shall honour each other’s territorial integrity and shall acknowledge the inviolability of the borders existing between them.” Which means that officially Moscow positioned itself as an ally of Ukraine but in reality was launching hostile operation on Ukrainian territory.

Bebler (2015) in his work shows that the psychological aspect of this war was well planned by invading Russian forces. Because immediately after the takeover on February 28, 2014, Russian security personnel shut off all Ukrainian television channels, imposed a tight blockade on the land border with the mainland Ukrainian territory, closed the Simpheropol airport’s flights from Ukraine and thus prevented the diffusion on Crimea of Ukrainian printed media (which are still issued mostly in the Russian language). The population of Crimea was thus subjected to one-sided information and often outright disinformation by the Russian state-controlled mass media. The intense propaganda campaign, almost like that during the “Cold War” depicted the interim Ukrainian authorities in Kyiv as “fascists” or “neo-Nazi” who presumably threatened the Russian and Russian-speaking population with “genocide”. Public harassment and intimidation of Crimean Tatars by the so-called “people’s self-defence” forces and by unidentified men in military fatigues, as well as physical and verbal threats to Ukrainian opponents of secession were reported. Fifteen pro-Ukrainian journalists and activists were abducted, detained and ill-treated. The Russian authorities barred the return of Mustapha Dzhemilev, a leader of the Crimean Tatars and a member of the Ukrainian Parliament. One known Tatar protester was reportedly abducted, apparently tortured, and found dead.

To legitimise the annexation of invading forces organised a referendum with two options:

- Are you in favour of Crimea joining the Russian Federation as a subject of the Russian Federation?
- Are you in favour of re-establishing the 1992 constitution of the Republic of Crimea and Crimea’s status as a part of Ukraine?

The maintenance of the territorial and status quo was not given as an option in that referendum, and no international observers were present (Peters 2015). Besides that, the referendum was held under the unpredictable conditions of Russian military occupation. The presence in public places of armed local Russian irregulars, of Russian Cossacks and even Serbian “Chetniks”, as well as of masked “little green men" undoubtedly belonging to the Russian Armed Forces, certainly had an intimidating effect on the opponents of Crimea's secession (Bebler 2015). Results of the referendum showed that 81.36 per cent of the registered voters took part in Crimea's referendum and 96.77 per cent of them voted for its separation from Ukraine and for reuniting with Russia.

Nevertheless, even if the Crimean referendum was organised according to all international legal standards, the "unification" of Crimea with Russian, cannot be justified by it. Holding a free and fair referendum is only a necessary, but not a sufficient condition for a territorial realignment to be accepted as lawful by international law. The operation could therefore not constitute a legal basis for the new territorial status quo (Peters 2015). Thus, till the day of writing this thesis, Crimea is still illegally occupied by the Russian Federation, and citizen of the peninsula are denied their fundamental human rights (Office of the United Nations High Commissioner for Human Rights 2017).

2.3.2. Russian strategy

Due to the possibility Moscow will repeat that same strategy against other post-Soviet countries, researchers and military agencies, particularly from the NATO members countries, put great effort to summarise and describe strategy that was used during annexation of Crimea. Christopher S. Chivvis, Associate Director of the International Security and Defense Policy Center and a senior political scientist at the RAND Corporation[1], in his testimony on Russian hybrid warfare (Chivvis 2017) before the Committee on Armed Services United States House of Representatives gave a rundown on characteristics, objectives and toolkit of Russian hybrid warfare. Below we describe critical aspects of the testimony.

Russian hybrid warfare can be characterised by three factors: economisation of force, persistency and population-centricity. It economises the use of force because Russia would stand little chance of winning a protracted conventional conflict

[1]RAND Corporation ("Research ANd Development") is an American nonprofit global policy think tank created in 1948 by Douglas Aircraft Company to offer research and analysis to the United States Armed Forces. It is financed by the U.S. government and private endowment, corporations, universities and private individuals. The company has grown to assist other governments, international organisations, private companies and foundations, with a host of defence and non-defence issues. https://www.rand.org/

with NATO, Moscow seeks instead to pursue its interests without the overt use of military power if possible. Russia may still use its conventional and even nuclear threats as part of a hybrid strategy, but in general, it prefers to minimise the actual employment of traditional military force. Russian hybrid warfare is persistent because it breaks down the traditional binary delineation between war and peace. The reality of hybrid war is an ever-changing intensity of the conflict. The third factor - population-centricity was evident during whole Crimean operation, as Moscow has realised the importance of an approach that seeks to influence the population of target countries through information operations, proxy groups, and other influence operations.

The main goals of hybrid warfare are:

- Capturing territory without resorting to overt or conventional military force
- Creating a pretext for overt, conventional military action
- Using hybrid measures to influence the politics and policies of countries in the West and elsewhere

These goals are achieved with a variety of tools and mechanisms. One of the most prominent methods is information operations for shaping political narratives in many countries. This is achieved with the help of traditional media, such as Russia Today[2] and Sputnik Media[3], and social media, such as Vkontakte[4], YouTube[5], Twitter[6] (Aro 2016). To maximise pro-Russian presence in social media a big amount of people employed forming so-called ”troll factories” (Čižik 2017). Close to information operations are cyber operations that are aimed at hacking into Western information systems to collect valuable information. The information is then used to influence elections and other political outcomes outside Russia's borders.

For the operations that do not take space in informational and cyber domains, Kremlin employees proxies. Proxies are often groups that have a deep sympathy with Russia's objectives. One of the Kremlin's typical proxies is the Night Wolves, a biker club and ultranationalist, anti-American gang, whose leader is a personal friend of President Putin. Other, more dangerous proxies are terrorist organisations such as ”Donetsk People's Republic” and ”Luhansk People's Republic” (Verkhovna

[2] https://www.rt.com/
[3] https://sputniknews.com/
[4] https://vk.com/
[5] https://www.youtube.com/
[6] https://twitter.com/

Rada of Ukraine 2015), which are fighting in Eastern Ukraine. Moreover, Moscow can use political and economic pressure to achieve its goals.

It is important to note that Russia's use of hybrid strategies is not new; it takes roots from the Soviet era (Chivvis 2017). After the fall of the Soviet Union, Russia inherited infrastructure, methods and practices, and continued developing them into what is now modern Russian hybrid warfare.

2.3.3. Understanding Russian Propaganda

> ...it was not in anyone's mind, frankly, that the radicals and neo-Nazis would prevail on the Ukrainian political arena...
>
> Sergey Lavrov, Minister of Foreign Affairs of the Russian Federation

As mentioned before, modern Russian warfare has its origin in the Soviet Union. Propaganda is one of the methods of informational warfare which has a long history of being used by the Soviet Union. For example, Soviets funded "euro-communist" political parties, encouraged antinuclear protest movements, and wanted to manipulate the European media (Chivvis 2017).

Origins of modern Russian propaganda is well described by Jolanta Darczewska, the deputy director of Centre for Eastern Studies[7], in work "The Anatomy of Russian Information Warfare the Crimean Operation a Case Study" (Darczewska 2014). In this work author shows, that present-day Russian propaganda is derived directly from spetspropaganda (special propaganda) theory, which was first taught as a separate subject in 1942 at the Military Institute of Foreign Languages. Spetspropaganda was removed from the curriculum in the 1990s to be reintroduced in 2000 after the institute had been reorganised. The institute is now known as the Military

[7] Centre for Eastern Studies in Polish research centre established in 1990. Centre's main task is to compile information on significant events and political, social and economic processes taking place in Poland's international surrounding and make it available to Poland's state authorities, prepare analyses, expert opinions and forecast studies. The Centre employs around 50 analysts. The Centre's primary activity is entirely funded from public resources. https://www.osw.waw.pl/en

Information and Foreign Languages Department of the Military University of the Ministry of Defence of the Russian Federation.

The reason behind the transformation of Kremlin's propaganda institute is an increase in interest in information warfare issues initiated by the work on the Information Security Doctrine of the Russian Federation announced in 2000 (Ministry of Foreign Affairs of the Russian Federation 2000). This Doctrine was one of the first policy documents issued by President Vladimir Putin's Security Council. The Doctrine's textual representation of threat delimits the boundaries of national identity and legitimises the exercise of control over the mass media. Employing identity politics in the service of a "strong state", the Kremlin has intensified its regulation of the communications infrastructure to ensure the uniformity and stability of the nation's worldview (Carman 2002). This Doctrine also leads to the changes in regulations of broadcast networks in Russia, which resulted in the creation of creating a new state-owned Russian Television and Radio Broadcasting Network. The goal of the new organisation was to be in charge of maintaining transmission facilities for both state and private broadcasters (Carman 2002). Due to overwhelming state control of media, according to Freedom House[8], since 2003 Russian freedom of the press is one of the worst in the world, on the level with such countries as Somalia, Rwanda and Swaziland (Freedom House 2017).

The narratives of Russian propaganda also bare similarity to the Soviet one. As it was mentioned before, after invading Crime, Russian media started an informational campaign depicting the new Ukrainian government as "fascist" and "neo-Nazi". To understand why this strategy was used, it is important to understand that the Soviet Union's victory over Nazi Germany is a central tenet of Russian national identity (Mijnssen 2010). The victory in World War II, or as it called in Russia – Great Patriotic War, serves as a morality tale of suffering and redemption and a foundation myth (Wood 2011). That is why the Russian government dictates and promotes the "correct" interpretation of history. Challenges to its view in the countries of the former USSR are condemned as anti-Russian and "fascist" and detrimental to Russian national interests (Mijnssen 2010). Thus, everybody who is deemed to be a "fascist" is automatically enemy to the whole Russian nation. This strategy is far from being new, as it was described by Soviet Colonel-General Volkonogov in work "Psychological Warfare", where the victory over the Nazism is unification factor for

[8]Freedom House is an independent watchdog organisation dedicated to the expansion of freedom and democracy around the world. https://freedomhouse.org

nations of Soviet Union to opposing ”imperialistic West” (Volkogonov 1983).

2.3.4. Principles of Propaganda

Analysing the Russian media rhetoric during the Russian invasion in Crimea, Darczewska (2014) outlines five key principles of successful propaganda, and Humen (2017) provides detailed examples of media materials used during annexation of Crimea for each principal. The outlined principles are:

1. principle of massive and long-lasting impact
2. principle of desired information
3. principle of emotional agitation
4. clarity principle
5. principle of supposed obviousness

The principle of massive and long-lasting impact implies the usage of particular labels attached to Ukrainian activists: ‘orange plague’ (from the Orange Revolution in Ukraine 2004), ‘Banderivtsy’ (from the name of Stepan Bandera, the leader of XX century Ukrainian nationalist movement) and other political stigmas (Humen 2017). Moreover, besides using already established labels, it aims at creating a new one, in case of Crimean operation the goal was to attach the label of ”fascist” to the new Ukrainian government and Ukrainian atavists. The Russian media at that time was filled with headlines like: ”Neo-Nazism: Made In Ukraine”[9], ”70 years later: fascism is marching again in Ukraine...”[10] and ”Lavrov: Kiev condones the neo-Nazi movements of Ukraine”[11].

The principle of desired information aims at certain templates of people‘s collective memory which predefined particular beliefs and are vulnerable to emotional provocations. Due to the importance of Soviet victory over Nazism in Russian collective memory, any mentioning of ”fascism” attracts the attention of Russian people. Moreover, Russians and Russian-speaking people expect that their rights should be protected, so they believed the manipulated information that the Russian language had been banned (Darczewska 2014). Articles with titles like: ”The Russian Federation spoke about ignoring the interests of Russian-speaking residents

[9] https://ria.ru/20141025/1030502474.html
[10] https://ria.ru/20141030/1030939584.html
[11] https://ria.ru/20141020/1029133792.html

in Ukraine"[12] and "Lavrov: Russia paid little attention to the Russian-speaking population of Ukraine"[13], created a pretext for invading Crimea. As a result, the majority of Russians believed in this artificial threat and supported the annexation of Crimea (Balzer 2015).

The principle of emotional agitation focuses on bringing the recipients of the message to a condition in which they will act without much thought, even irrationally (Darczewska 2014). The prime example would be the story of "crucified boy"[14]. The great news episode made by Channel One Russia presents an interview of the allegedly Ukrainian refugee in Russia, who describes the "demonstrative execution" of a three-year-old child by the Ukrainian soldiers. The woman provides very shocking details of the child being nailed to the announcement desk "like a Jesus" on the main square of the city of Slovyansk. Of course, this story was a fake (Stop Fake 2014), but because states control Russian media, the story was not removed or shown to be a fake.

The clarity principle implies the usage of easy-to-understand language carrying catchy and straightforward, but strongly politicised load (Humen 2017). Repetition is also one of the tools used for delivering a clear message to the audience. For example, Humen (2017) noticed that in the article "Novorossiia – Born in Flame"[15] word "fascism" and its derivations are mentioned for fifteen times. According to Humen (2017), the author of the article, utilizes such a rhetoric in order to create the simplistic image of Novorossiia[16] as a part of great Russian history. Importantly, this image is constructed as opposed to fascism.

The principle of supposed obviousness indicates the audience‘s unquestioned consumption of the information, which carries a politically loaded implication (Humen 2017). According to Darczewska (2014), this principle causes the propaganda thesis to be associated with the already created political myth. These myths are predefined by the principle of massive and long-lasting impact they construct the model of the society‘s worldview which inclines people to particular judgments (Humen 2017). Humen (2017) argues that Russian media construct an alternative reality in

[12]https://ria.ru/20140328/1001462211.html

[13]https://ria.ru/20141119/1034069099.html

[14]http://www.1tv.ru/news/world/262978

[15]http://izvestia.ru/news/570647

[16]Novorossia is the name often used by Russian media when referring to Ukrainian territories occupied by terrorist organizations "Donetsk People's Republic" and "Luhansk People's Republic"

which the Ukrainian government is fascist and Ukrainian activists are neo-Nazis and due to this news about ”Ukrainian fascist” are consumed without any proofs provided.

2.4. Machine Learning

Machine Learning is the scientific study of algorithms, and statistical models that enable a machine to change its structure, program, or data based on inputs or in response to external information to achieve future performance improves (Nilsson 1996, Bishop 2006). Nilsson (1996) identified the reason why it is important for a machine to have the ability to learn; those reasons include:

- Some tasks cannot be defined well except by example, meaning that it is possible to specify input/output pairs but not a concise relationship between inputs and desired outputs.
- Human designers often produce machines that work as well as desired in the environments in which they are used. Certain characteristics of the working environment might completely be known at design time.
- Environments change over time.
- New knowledge about tasks is constantly being discovered by humans.

Machine learning approach to solving a problem can be described as the process of tuning parameters of an adaptive model based on data to produce the desired output. There are two major ways of how the machine can learn: supervised and unsupervised learning. The idea of supervised learning is that for each data point desired output is known. In the case of unsupervised learning, the desired output, and model learn hidden relations in data. A typical task of supervised learning is classification, of unsupervised – clustering. There are a great variety of different machine learning models, in the course of this work we use models based on Random Forest and Neural Networks.

2.4.1. Random Forest

Often to improve results of machine learning, different models are trained, and their results are aggregated, such approach is called Ensemble Learning (Liaw et al. 2002).

Random Forest, introduced by Breiman (2001), is a prime example of this approach. The idea of the proposed method is to combine results of binary tree-structured models, which are constructed by splits of subsets of data X into two descendant subsets, beginning with data X itself (Breiman et al. 1984). Moreover, each tree is constructed with a much smaller subset of original data, and this technique is called bootstrapping. In addition, split in trees used in a random forest, are based on a random subset of predictors, which adds one more layer of randomness. This strategy makes random forest robust against overfitting (Breiman 2001).

2.4.2. Neural Network

Neural Networks or Artificial Neural Networks are machine learning models inspired by the structure of the human brain. The human brain as computational systems is inspiring because it is a highly complex, nonlinear, and parallel information-processing system. It can organise its structural constituents to perform certain computations Haykin (2009).

Just like the human brain consists of neurons, the neural network is built with perceptrons – a mathematical model which are inspired by biological neuron (Rosenblatt 1958). The goal of perception is to correctly classify set of external inputs $x_1, x_2, ..., x_n$ into two classes, y_1 or y_2, based on weights w_i and non linear activation function f. There are different nonlinear activation functions. However, classical uses are heaviside step functions. Perceptron is limited to the classification of linearly separable patterns. To overcome this problem, multiple perceptrons can be stacked into layers. Then the layers are connected between each other in the way, that every node is fully connected to the next layer. Such architecture forms one of the many types of neural networks and allows to approximate non-linear functions.

2.4.3. Evaluation techniques

To decide if the model is performing well some evaluation techniques are needed. There are different approaches to evaluation for different machine learning application. For this thesis evaluation of classification of labelled data is the most relevant. Classification models can be evaluated with the confusion matrix, the technique that was first introduced by Miller & Nicely (1955). In the original work, authors evaluate performance over multiple classes, and this work is aimed at binary classification.

A confusion matrix is a specific layout that is used for visualisation of the model's performance. Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class or vice versa (Powers 2011). In the case of binary classification, there are only two classes – positive and negative, which means that confusion matrix will have only four fields, because of this they have specific names:

- True Positive (TP). The predicted label is positive, and the true label is positive.
- True Negative (TN). The predicted label is negative, and the true label is negative.
- False Positive (FP). The predicted label is positive, but the true label is negative.
- False Negative (FN). The predicted label is negative, but the true label is positive.

From the confusion matrix, we can derive measurements of accuracy, precision, recall and F1-score. Accuracy shows the proportion of correct results to all results:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision is the number of correctly classified objects, out of all objects which were classified as positive:

$$Precision = \frac{TP}{TP + FP}$$

The recall represents the number correctly classified objects, out of all objects which should have been classified as positive:

$$Recall = \frac{TP}{TP + FN}$$

F1-Score is a harmonic mean between precision and recall:

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

2.5. Nature Language Processing

Natural Language Processing is sub-filed of computer science and artificial intelligence dedicated to analysing text and speech. In this work definition proposed by Liddy (2001) will be used:

> Natural Language Processing is a theoretically motivated range of computational techniques for analysing and representing naturally occurring texts at one or more levels of linguistic analysis to achieve human-like language processing for a range of tasks or applications.

In addition to the definition, Liddy (2001) also provides further explanation for key elements. For example, the author explains necessity imprecise notion of ”range of computational techniques” – because there are multiple methods or techniques from which to choose to accomplish a particular type of language analysis. The phrase ”naturally occurring text” implies that text can be oral or written of any mode, genre and language. The only requirement is that they be in a language used by humans to communicate with one another. The text is analysed also should not be specifically constructed for the analysis, but rather that the text is gathered from actual usage. The notion of ”levels of linguistic analysis” refers to the fact that there are multiple types of language processing known to be at work when humans produce or comprehend language. It is thought that humans normally utilise all of these levels since each level conveys different types of meaning. However various Natural Language Processing systems utilise different levels or combinations of levels of linguistic analysis, and it differs from application to application.

Natural Language Processing has a variety of applications which include:

- Text Categorization. Text categorisation or text classification is the activity of labelling texts with thematic categories from predefined set of labels. It is applied in many contexts, document indexing based controlled vocabulary, to document filtering, automated metadata generation, word sense disambiguation, the population of hierarchical catalogues of Web resources, and in general any application requiring document organisation or particular and adaptive document dispatching (Sebastiani 2002).
- Named Entity Recognition. The task of identifying information units like names, including person, organisation and location names, and numeric expressions including time, date, money and per cent expressions, is called

Named Entity Recognition Nadeau & Sekine (2007). Initially, this task was formed on the Sixth Message Understanding Conference (MUC-6) which was focusing on Information Extraction (Grishman & Sundheim 1996).

- Sentiment Analysis. The goal of Sentiment Analysis is to analyse people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organisations, individuals, issues, events, topics, and their attributes (Liu 2012). Opinions are central to almost all human activities because they have key influence on people's behaviours. Sentiment Analysis has much application because businesses and organisations always want to find consumer opinions about their products and services (Liu 2012).

There is much more application, but these three applications are the most relevant to the work done in this thesis.

As the definition states, Natural Language Processing uses a lot of different techniques and algorithms for achieving results for each given application. One of the most used technique is Language Modeling, which is divided into two main approaches Statistical Language Modeling and Neural Language Modeling.

2.5.1. Statistical Language Modeling

The task of Statistical Language Modeling is developing probabilistic models, which can assign a probability to sentences in a language. Besides assigning a probability to each sequence of words, the language models also assign a probability for the likelihood of a given the word (or a sequence of words) to follow a sequence of words (Goldberg & Hirst 2017).

According to Goldberg & Hirst (2017), formally, the task of language modelling is to assign a probability to any sequence of words $w_{1:n}$, in other words – to estimate $P(w_{1:n})$. This can be rewritten as:

$$P(w_{1:n}) = P(w_1)P(w_2|w_1)P(w_3|w_{1:2})P(w_4|w_{1:3})...P(w_n|w_{1:n-1})$$

Since the last term in the equation still requires conditioning on $n - 1$ words, statistical language models make use of the markov-assumption, stating that the future is independent of the past given the present. Meaning that a k-th order markov-assumption assumes that next word in a sequence depends only on the last k words (Goldberg & Hirst 2017).

The implementation of statistical language models ranges from very simple –based on trigrams, to much more complicated, which use caching, clustering, higher-order n-grams, skipping models, and sentence-mixture models (Goodman 2001). Statistical Language Modeling becomes very useful in a large variety of areas including speech recognition, optical character recognition, handwriting recognition, machine translation, and spelling correction (Goodman 2001).

2.5.2. Neural Language Modeling

The recent increase in the usage of neural networks in the Natural Language Processing led to the changes in language modelling. New models appeared that used neural networks, and it became, what is known now as Neural Language Modeling. Neural network approaches are achieving better results than classical statistical methods because they allow conditioning on increasingly large context sizes with only a linear increase in the number of parameters, they alleviate the need for manually designing backoff orders, and they support generalisation across different contexts (Goldberg & Hirst 2017). It is achieved by using word embedding – a real-valued vector that represents each word in a project vector space. Such representation of words is learned based on their usage, which results in a similar representation of similar words (Kim et al. 2015). The word embedding representation allows models to scale better with the size of the vocabulary.

Bengio et al. (2006) described the neural network approach with three following properties:

- Associate each word in the vocabulary with a distributed word feature vector.
- Express the joint probability function of word sequences in terms of the feature vectors of these words in the sequence.
- Learn simultaneously the word feature vector and the parameters of the probability function.

It represents the model where both the representation and probabilistic model are learned together directly from raw text data.

There are a great variety of different neural language models which have different advantages and disadvantages. In the course of this thesis model developed by Mikolov et al. (2013) was used.

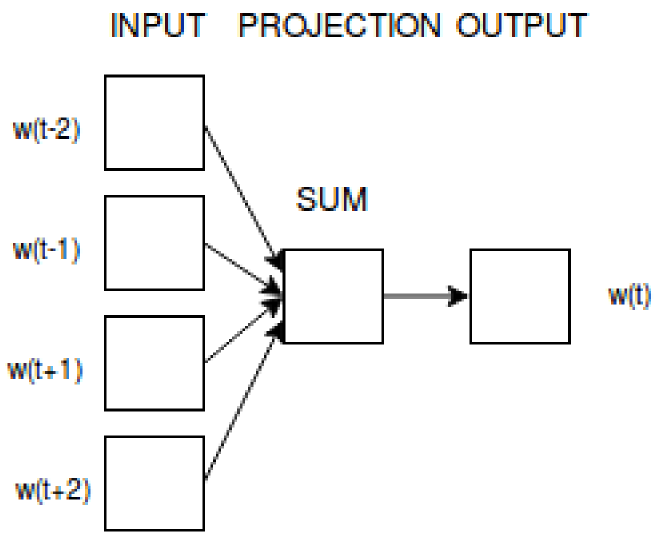

Figure 2.1.: Continuous Bag-Of-Word (CBOW) architecture of the model for continuous word representation introduced by Mikolov et al. (2013). CBOW predicts the current word based on the context.

2.5.3. Word2Vec

Mikolov et al. (2013) introduced two new model architectures that can be used for learning high-quality word vectors from huge data sets with billions of words, and with millions of words in the vocabulary. Moreover, obtained word vectors have a modest dimensionality of $50 - 100$.

The first proposed architecture is similar to the classical neural network approach, but the non-linear hidden layer is removed, and the projection layer is shared for all words. Thus all words get projected into the same position because their vectors are averaged. Mikolov et al. (2013) described this architecture as a bag-of-words model as the order of words on history does not influence the projection. Authors obtained the best performance by building a log-linear classifier with four future and four history words at the input, where the training criterion is to classify the current (middle) word correctly. Since, unlike the standard bag-of-words model, this architecture uses a continuously distributed representation of the context, it was denoted as Continuous Bag-Of-Words or CBOW for short. The CBOW architecture is illustrated in Figure 2.1 The second architecture introduced by Mikolov et al.

(2013) is similar in the way to CBOW, but instead of predicting the current word based on the context, it tries to maximise classification of a word based on another word in the same sentence. Authors use each as an input to a with continuous projection layer, and predict words within a certain range before and after the current word. The architecture is illustrated on the Figure 2.2

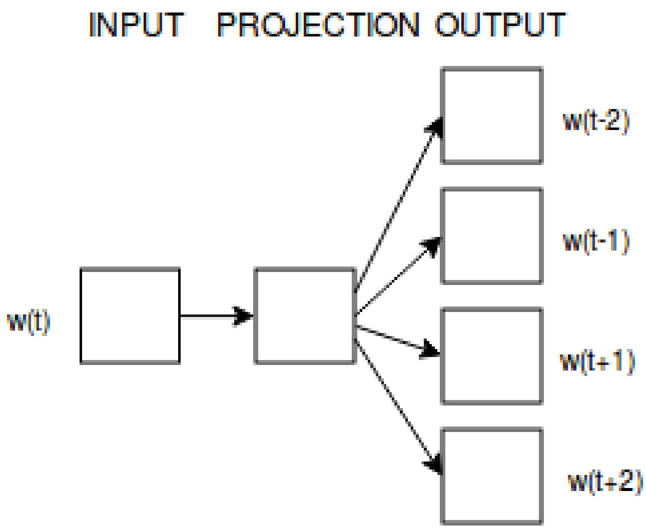

Figure 2.2.: Skip-gram architecture of model for continuous word representation introduced by Mikolov et al. (2013). Skip-gram predicts surrounding words given the current word.

Mikolov et al. (2013) showed that introduced models, when trained on a large amount of data, produce vectors which can be used to answer to answer very subtle semantic relationships between words, such as a city and the country it belongs to, for example, France relates to Paris as Germany relates to Berlin. These models became known in the literature as word2vec and were used for different applications in Natural Language Processing including sentiment detection (Liu 2017, Rexha et al. 2016, Zhang et al. 2015), text classification (Lilleberg et al. 2015, Johnson & Zhang 2015) and named entity recognition (Sienčnik 2015).

3. Related Work

With growth Internet and social media, the problem of propaganda became very important, because now propagandists can target a huge amount of people with relative ease. That is why most of the research focuses on the detection of propaganda on social media. Before the attention of scientific community was attracted to Russian propaganda, the research was focusing on detection extremist propaganda, which was sparked by terrorist attacks carried by ISIS[1], Al-Qaeda[2] and other radical jihadist organizations. However, after the annexation of Crimea and alleged intervention of Russia into US elections, lead to an increasing amount of research. The following section describes the work done in both fields.

3.1. Extremist Propaganda Detection

The terrorist attacks that were carried by people who were not members of a terrorist organisation, but were indoctrinated by extremist propaganda, raised concern in the society. Social media platforms recognised that extremists were using them as a medium for delivering their radical believes (Awan 2017). However, the task of preventing spared of radical ideas is not the easy one, because moderators have to find perpetrators of this between millions and millions of accounts[3]. Of course, it is impossible to manually inspect every account; that is why machine learning methods have been used to automatize the process.

The research in the direction of detection of extremist propaganda started with collecting data based on keywords associated with ISIS (Berger & Morgan 2015),

[1] Islamic State of Iraq and Syria – jihadist militant group and terrorist organization
[2] jihadist militant multi-national terrorist organization founded by Osama bin Laden
[3] According to statista.com in the fourth quarter of 2018, Twitter had 321 million monthly active users. https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/

which lead to the development of tools for the automatic detection of extremist content. Ashcroft et al. (2015) developed one of such tools, in their work authors extract stylometric, time-based and sentiment based features. Based on those features authors trained three different classifiers: Support Vector Machine, Naive Bayes and AdaBoost, all of which has very high evaluation results. A similar approach was employed by Kaati et al. (2015), but for detecting, if the user is a supporter of radical ideas.

However, it was argued by Ferrara (2017) that such keyword-based approach are limited and, to improve it, the author used dynamic activity-connectivity maps, based on network and temporal activity patterns and investigated the dynamics of social influence within ISIS supporters. For their work Ferrara (2017) used the dataset of more than 25 thousands accounts, that were manually labelled as ISIS supporters. This work was continued by Badawy & Ferrara (2018). Authors focused on content written in the Arabic language. Findings of this work yield essential insights about how social media is used by radical militant groups to target the Arab-speaking world and reveal important patterns in their propaganda rhetoric.

Holm (2017) dedicated their work to monitoring jihadist propaganda in other media – online newspapers and magazines published by terrorist organisations. The idea of the work was to detect the authorship and ideology of the given text. The author managed to show that by pairing an ideological-content detector with an authorship detector, higher recall can be obtained. Johnston & Weiss (2017) also focused their work on the data from different sources which included news articles, Wikipedia entries, paste site pages, and terrorist magazines. For feature extraction authors used Doc2Vec, which is a more generalised version of Word2Vec. The resulting method showed to robust and able to learn to classify extremist multilingual text of varying length effectively.

It is important to note, that most of the work done in the field of detection extremist propaganda fall into the category of supervised learning because there are available datasets that were labelled by experts. Moreover, all media produced by terrorist organisations can be considered as propaganda.

3.2. Russian Propaganda Detection

The world's attention to Russian propaganda was attracted after claims that Moscow interfered in 2016 US elections. Kremlin chose social media as one of the ways of influencing US citizen, that lead to research in the field of detecting so-called trolls and bots. One of such research is done by Badawy et al. (2018). The authors used a network-based machine learning method to determine the political ideology of most of the users and ran state-of-the-art bot detection analysis on users who engaged with Russian Trolls. They determined that bots were engaged in both liberal and conservative domains. Stewart et al. (2018) aimed their research on the strategy that was very effective in the Ukrainian crisis – polarisation of society. In this paper, authors have located Russian affiliated troll accounts in the re-tweet network of a politically polarised conversation surrounding race and shootings in the United States. What is more, Russian propaganda is not just interfering with the politics of other countries, but also promotes dangerous ideas such as the anti-vaccination movement, which was detected by Broniatowski et al. (2018).

The worrying statistic was discovered by Stukal et al. (2017). The author showed that in the period from February 2014 to December 2015, an especially considerable period in Russian politics, among accounts, actively tweeting about Russian politics the proportion of Tweets produced by bots usually exceeded 50%. Even though there are many proofs that Russian conducts propaganda campaigns in different countries, to the best of our knowledge, there was no research done with a focus on Russian national media.

4. Methodology

As it was mention in Chapter 1, this work is focusing on the analysing of propaganda in online newspapers. Since propaganda campaigns can have different goals and methods, the aim of this research is on investigating the phenomenon that was observed by several researchers (Darczewska 2014, Bebler 2015, Humen 2017) – an antagonising group of people by depicting them as fascists.

In the following section, we describe the methodology of this work. It can be summarised into a two-step process: first – find articles that potentially cause a shift in perception of the targeted country towards fascism; second – check if those articles correspond to the principles of propaganda described in Section 2.3.4. Bellow, when referring to articles that cause a shift in the perception of the targeted country, the term "suspected articles" is used.

4.1. Finding Propagandistic Articles

4.1.1. Detecting Shift in Narrative

We frame the problem of determining if possible propaganda is present in dataset as a detection of the shift in the media narrative. To do this data from different periods is required. With this data, we train word2vec models for each period. Thus, with the models produce vector representation for each word used in the dataset, having this it is possible to calculate the cosine similarity between pairs of words. Further in the text, when referring to similarity, it is meant the cosine similarity. Cosine similarity between two vectors A and B is calculated with the following formula:

$$similarity(A, B) = \frac{A \cdot B}{||A|| \times ||B||} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i} \times \sqrt{\sum_{i=1}^{n} B_i}}$$

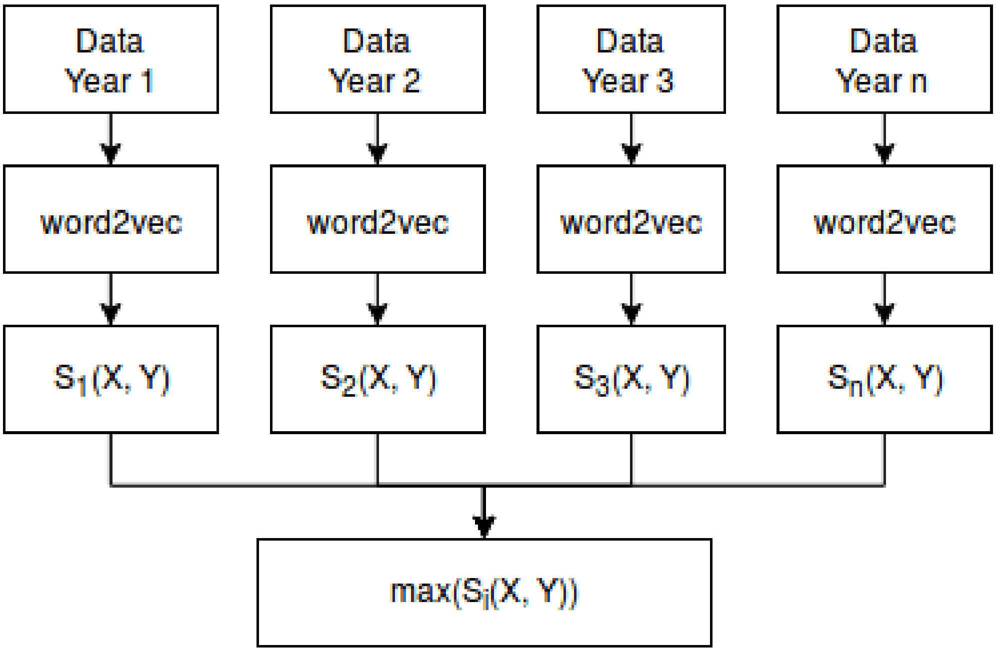

Figure 4.1.: The process of detecting a shift in the media narrative. Data, which consist of any textual sources, should be divided into years, and for every year word2vec model is trained. S_i is a similarity measurement between two word-vectors. Where X is targeted country (e.g. Ukraine) and Y negative term (e.g. fascism). The result is a year with the maximum similarity between targeted country and negative term

The resulting similarity ranges from -1 to 1. With this measurement, it is possible to detect the changes in the representation of the target group. For example, if it is observable that the similarity between word-vectors of words "Ukraine" and "fascism" is increasing over time, it is an indicator of alteration in the depiction of Ukraine. Therefore, it is possible to find the year in which similarity was the highest. This process is outlined in the Figure 4.1

4.1.2. Extracting propagandistic articles

When the data, in which similarity between targeted country and fascism, or any other negative term, is the highest, is detected, it is needed to find which articles are causing such changes. The data should be divided into different groups based on different categories. For example, based on topic, date of publishing, author, etc.

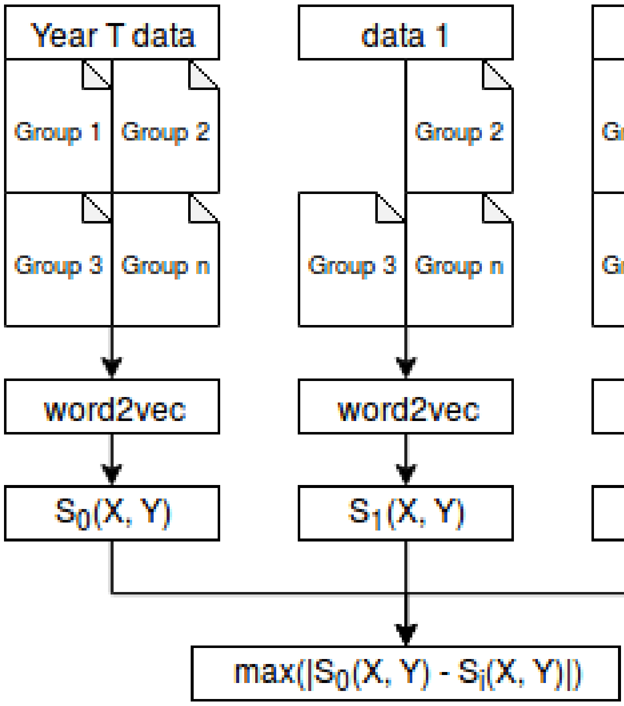

Figure 4.2.: The process of finding articles that are responsible for the shift in the media narrative. Year T – is the year when similarity S_0 between X and Y is the highest, and X is target country, Y is a negative term. Data is divided into n groups, based on different criteria. Then each, group is removed from data, and new word2vec model is trained. Based on the new model similarity S_i between X and Y are calculated and compared to S_0.

One of the most important groups are articles that mention the target country and negative term together. Then, for each group, the new word2vec model must be trained without articles from the suspected group. Having newly trained word2vec model, the similarity between the target country and the negative term should be calculated again and compared with the original value. The goal of this process is to find a group which, when removed, minimises similarity between target country and negative term. Articles from such group are responsible for the shift in narrative. This process is outlined in the Figure 4.2.

4.2. Analyzing Principles of Propaganda

However, the fact, that articles from suspected groups caused a shift in the media narrative is not enough to determine if suspected articles are propagandistic or not. It is needed to see if suspected articles correspond to the principles of propaganda introduced by Darczewska (2014) and described in the Section 2.3.4. Every principle requires a different type of evaluation.

4.2.1. Principle of Long Lasting Impact

Principle of long-lasting impact is measured by the bias of the text classification algorithm. To measure the bias, we train a classifier. In this work, we use Random Forest classifier, but any other classifier can be employed for this task. The goal of the classifier is to be able to determine if the article contains mentioning a negative term. The negative term and all derivative forms are removed from the articles, and to see if the classifier is biased against the targeted country the country names are removed too. Then articles were converted into vectors, by converting every word with word2vec model and taking a pointwise maximum of obtained vectors. After evaluation, bias can be measured by classifying articles which mention target country and do not have any mentioning of negative term and seeing how much of those articles were classified as one that contains fascism (false positive). The obtained number should be compared to other countries, and if it is higher in the case of the target country, bias is present. The method is summarised in the Figure 4.3.

4.2.2. Principle of Desired Information

The principle of the desired information aims to attract as many people as possible. Therefore, good measurement is comparing the number of views, or other user interaction, depending on what features are available in the dataset, to the number of views on other articles. Thus, it is expected that propagandistic articles attract much more traction than other articles.

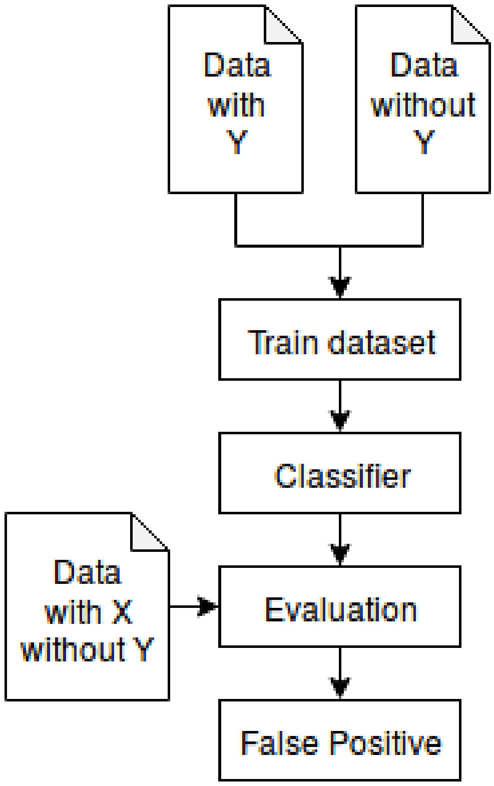

Figure 4.3.: The approach for measuring the long-lasting impact of propagandistic articles based on classifier bias. For training dataset equal amount of texts with and without negative term, Y should be sampled. All country names and derivative forms of Y must be removed. Binary classifier, which classifies if a text contains term Y, should be trained. Then texts with country name X and without negative term Y should be classified, and the amount of false positives is the value of bias.

4.2.3. Principle of Emotional Agitation

The goal of emotional agitation is to cause negative emotions in targets mind and connect it to the desired subject, which will encourage recipients to act without a critical assessment of the acquired information. To see if this occurs in suspected articles, we perform sentiment analysis based on the lexicon look-up. The idea of this approach is to have a dictionary of words that are labelled positive or negative, and count the percentage of those words in the text. It is expected that suspected articles have a higher amount of negative words.

4.2.4. Clarity Principle

Clarity principle is intended to ensure that the propagandistic measure is easy to understand. It can be achieved by simplifying the text. To see if the texts of suspected articles are simpler and easier to read, the readability score can be calculated. In this work, we use Flesch reading-ease test (Flesch 1979) and LIX readability test (Björnsson 1983).

Originally Flesh reading-ease test developed for measuring how difficult is a passage in English is to understand. The following formula calculates Flesh Reading-ease Score (FRS):

$$FRS = 206.835 - 1.015\left(\frac{\textit{amount of words}}{\textit{amount of sentences}}\right) - 84.5\left(\frac{\textit{amount of syllables}}{\textit{amount of words}}\right)$$

The constants in this formula were developed for English, because of this, those constants have to be adapted to the Russian language. In Russian, the average length of a sentence is less (due to less use of functional words, such as articles or auxiliary verbs), and the words are on average longer, Oborneva (2005) proposed following adaptations:

$$FRS_{rus} = 206.835 - 1.3\left(\frac{\textit{amount of words}}{\textit{amount of sentences}}\right) - 60.1\left(\frac{\textit{amount of syllables}}{\textit{amount of words}}\right)$$

Higher resulting scores indicate that text is easier to read. For the English, scores were interpreted by paring them with school grade, where the score of 100 would represent 5th grade, meaning that text is very easy to read and easily understood by an average 11-year-old student, and text with a score lower than 30 is for college graduates – very difficult to read (Flesch 1979). However, this interpretation is not always correct for other languages. Therefore, in this work, when the readability score of suspected articles is compared to the readability score of other articles to determine if they are easier to read.

Due to the fact that Flesh reading-ease is very language depended, it was decide to use one more measurement, which is more language independent – LIX readability test (Björnsson 1983). It is calculated as follows:

$$LIX = \frac{\textit{amount of words}}{\textit{amount of sentences}} + \frac{\textit{amount of long words} \cdot 100}{\textit{amount of words}}$$

Where "long words" are words with more than 6 letters. In contrast to FRS, higher LIX score represents texts that are hard to read.

Moreover, as it was shown in Section 2.3.4, to simplify text the repetition of words is used. However, it is expected that with the increasing size of text amount of repeated words is also increasing. Thus, the suspected articles can be checked if they obey Heaps' Law, which describes the number of distinct words in a document (or set of documents) as a function of the document length (Egghe 2007). It is formulated as:

$$V_R(n) = Kn^{\beta}$$

Where V_R is the number of distinct words in an instant text of size n. K and $beta$ are free parameters determined empirically. For English, K is between 10 and 100, and β is between 0.4 and 0.6. However, Gelbukh & Sidorov (2001) showed that the parameters are changing for different languages and topics. Due to this fact, the goal would be comparing how much different is β of suspected articles from β of other articles.

4.2.5. Principal of Supposed Obviousness

Methods mentioned above can be easily adapted for different languages and settings. However, it is hard to achieve this with the principle of supposed obviousness. The reason is – the goal of this principle is to cause propaganda thesis to be associated with the already established political myth. Such a myth is different from country to country. To find such a myth, one should research the history and geopolitics of countries involved in the conflict. In the case of Russia, the term ”Zapad” (”Запад”) can be used. This term, which means west, is a collective name for USA, NATO and West Europe. In the Soviet Union “Zapad” was a representation of “imperialism” and “exploitation of workers” (Volkogonov 1983), and in Russia “Zapad” become the centre of “liberalism” and “moral failure” (Vázquez-Liñán 2017). Therefore, to check if the principal of supposed obviousness holds in the suspected articles, the occurrence of the term ”Zapad” can be calculated.

4.3. Influence on Other Countries

To see if suspected articles would have had the same impact on other countries, we replaced the targeted country name with a different country. After that, word2vec models are retrained with newly modified suspected articles. If the similarity between selected country and the negative term has increased, it indicates that suspected articles were written in the way to achieve the shift in how the target country is perceived.

5. Dataset

Considering the fact, that in this thesis propaganda is viewed as part of informational warfare, for the analysis the stated owned media source was selected – RIA Novosti (RIA News or RIA)[1]. RIA News stands for Russian News and Information Agency. This institution has a long history dating back to 1941. The agency was founded under the resolution of USSR Council of People's Commissars and the Communist Party Central Committee, "On the Establishment and Tasks of the Soviet Information Bureau", and was named Soviet Information Bureau (Sovinformburo). Its main task was to oversee work to cover international, military events and the events of the country's domestic life in periodicals and on the radio (RIA Novosti 2013). After the fall of Soviet Union, the Soviet Information Bureau was reorganised into RIA Novosti and was placed under the control of the Press and Information Ministry. On 9 December 2013[2] the agency was reorganised again, and it became part of state-owned international agency Rossiya Segodnya (Russia Today).

Moreover, RIA News is a suitable source of data, because, as it was shown in the Chapter 4, for the approach used in this work data from different periods is required. Official website of RIA News provides articles since 2001 till nowadays.

5.1. Data Collection

To obtain articles from RIA News web crawler was developed. This application was developed with Python language and with using functionality provided by package Scrapy[3]. Due to the fact, that website https://ria.ru also has links to external resources, we designed the web crawler as a two-step process: first, collect all links to

[1] https://ria.ru/
[2] Annexation of Crimea started in the February 2014
[3] https://scrapy.org/

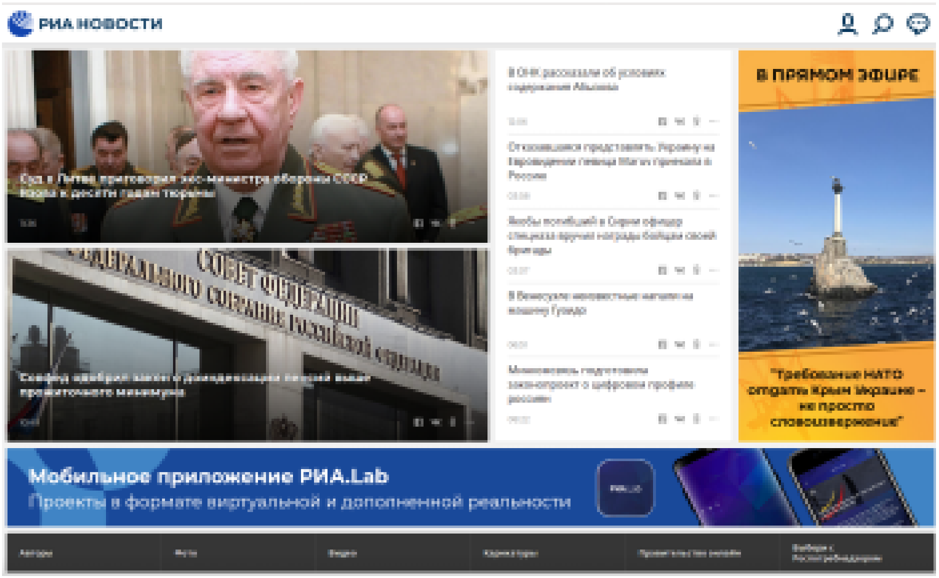

Figure 5.1.: Homepage of RIA News (https://ria.ru). It can be seen, that this web page has advertisement banners. That is why, web crawler was designed to collect links to news articles first, and then crawl those news articles.

news articles; second, crawl articles from collected links. Example of the homepage of RIA News is presented in the Figure 5.1.

We collected links to news articles by using RIA News search function. The idea of this approach is to send the following HTTP POST request:

https://ria.ru/services/search/getmore/?offset=i

with the following parameters:

'date_from': 'yyyy-mm-dd'

'date_to': '-yy-yy'

'interval': 'period'

'limit': '20'

'query': ''

'search_area': 'all'

Where 'date_from' and 'date_to' are beginning and end dates of the period of interest; 'interval' indicates the time frame for the search query, value 'period' means that the server needs to return articles between 'date_from' and 'date_to'; 'limit' – sets the number of articles returned by one request, it was noticed that if 'limit' cannot be set to the value of more than 20; 'query' is the search term, here dot is used in order to collect all articles, because it is impossible to send request with empty 'query'; 'search_area' indicates which areas of the website should be searched. The result of the request is list of links.

Due to the fact that one request return only 20 links, it is needed to iterate over all results with parameter i, which starts from 0 and increases by 20 till the moment when server return empty result. Furthermore, we noticed that setting the time frame to more than two months, the request does not return all articles, thus request were send with the one-month time frame.

Next step was to collect news articles from obtained links. We decided to download whole HTML pages, because they may contain metadata which can be useful in further research.

5.2. Data Preprocessing

After all data was downloaded, we had to extract texts from the HTML pages. Text extraction was done with Python language and packaged Beautiful Soup[4]. All articles from RIA News have the same HTML structure. Therefore it was possible to find HTML tags which store key information of the article. For the goals of this thesis, we extract headlines, article bodies and view counts. Those parts can be found with followings tags: 'headline', 'articleBody' and 'b-statistic__item m-views'. Then the extracted text was stored in text files, each article in a separate file.

To prepare data for training of word2vec models, we grouped articles per year and tokenised them on the sentence level. In the obtained sentences, we lemmatised every word. After that, we wrote each sentence into the file as a new line. In order to be able to find articles based on words, we made an index of words to the corresponding URL address. We published this index and grouped text files(Bohdan 2019).

[4]https://www.crummy.com/software/BeautifulSoup/bs4/doc/, version 4.4.0

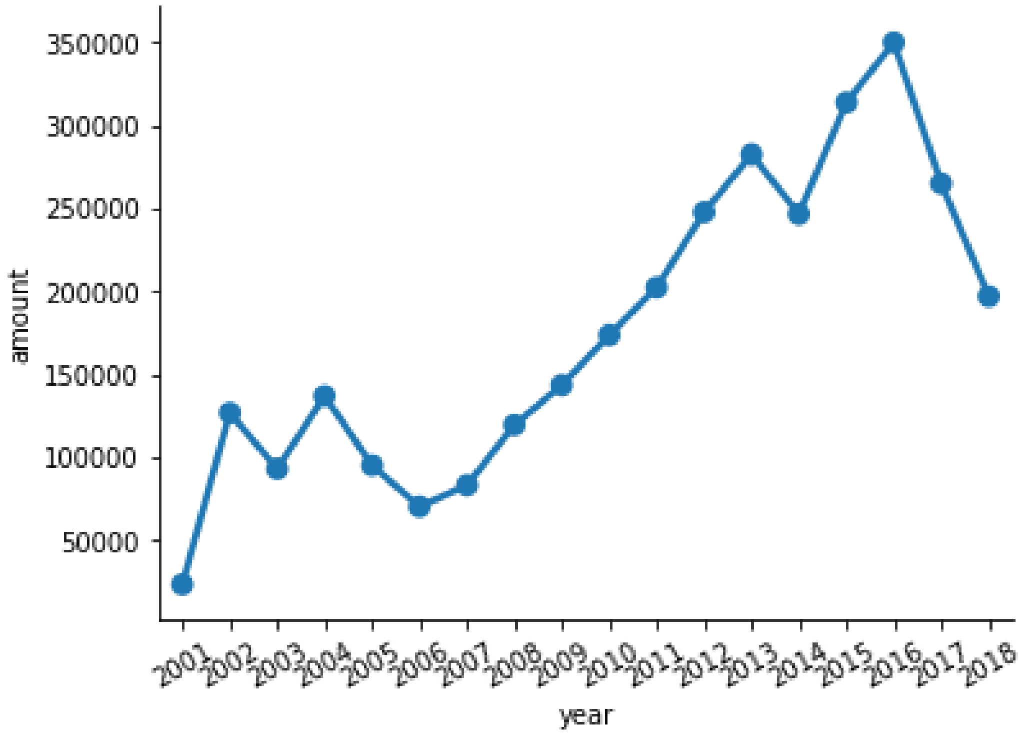

Figure 5.2.: Amount of articles published by RIA News per year. In total there are 3173079 between January 2001 and August 2018

5.3. Exploratory Data Analysis

For this work, we collected articles between the January 2001 and August 2018. In total it is 3173079 articles. Amount of articles published per year is shown on the Figure 5.2 and detailed statistic can be found in the Appendix A.1. As this work is focused on the representation of Ukraine, it is useful to investigate how much articles mention Ukraine. There are in total 265766 such articles. However, the distribution over the year is not equal, which is shown on the Figure 5.3. It can be observed that there is an enormous increase in the number of articles that mention Ukraine, from 10328 to 49123. It means that, in 2014, 19.9% of all articles, published by RIA News, were mentioning Ukraine. It can be seen from Appendix A.2 and A.3, in 2014 Ukraine was the most reported country after Russia itself. Considering this tremendous increase and the fact that the year 2014 is the year where Crimea was

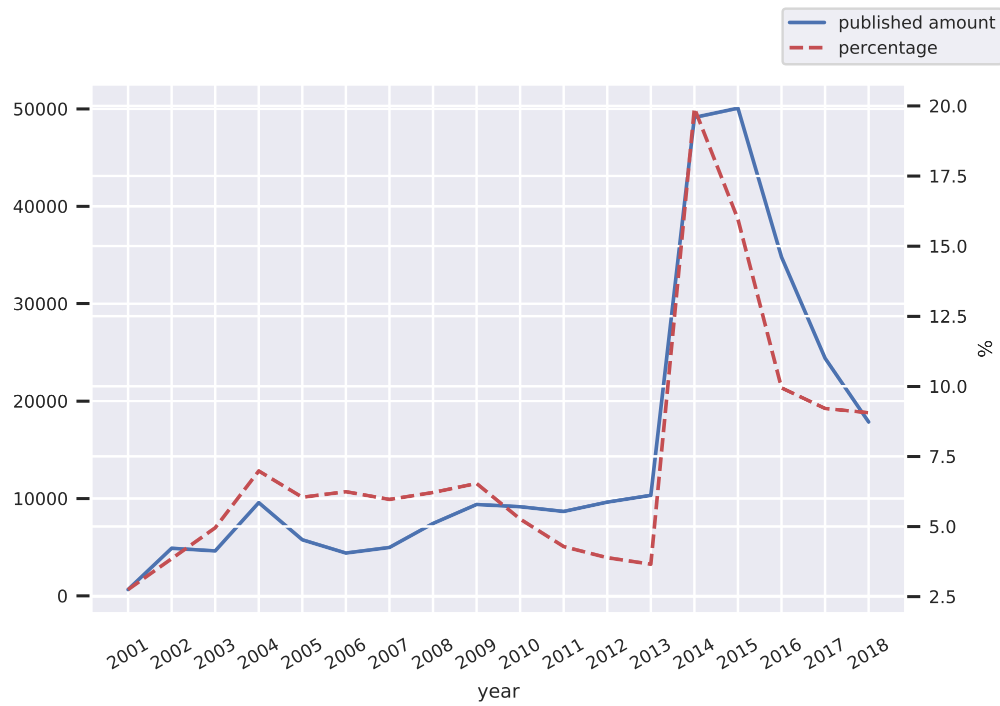

Figure 5.3.: Amount of articles published by RIA News each year that mention Ukraine in absolute numbers and its relative share based on all articles published in the respective year.

occupied, the year 2014 is the main focus of this work. The daily amount of articles mentioning Ukraine in 2014 is presented on the Figure 5.4, and detailed numbers are in the Appendix A.4. It can be observed, that the amount of articles published per day, is increasing rapidly from January, and reaching a maximum of 335 articles in one day on the 18th of July. Moreover, from the March most of the days have more than 100 articles mentioning Ukraine. As, it was mentioned in Section 2.3.3, there was an influx of articles that describe Ukraine as a fascist country. Overall, in 2014 Ria News published 372 articles that mention Ukraine and the word "fascism" in one article. The distribution over days can be seen on the Figure 5.5 and exact numbers are in the Appendix A.5. It can be observed that in 2014 every week there was at least 1 such article per week, reaching its peak in May, when 12 articles that mention Ukraine and word "fascism" were published in one day.

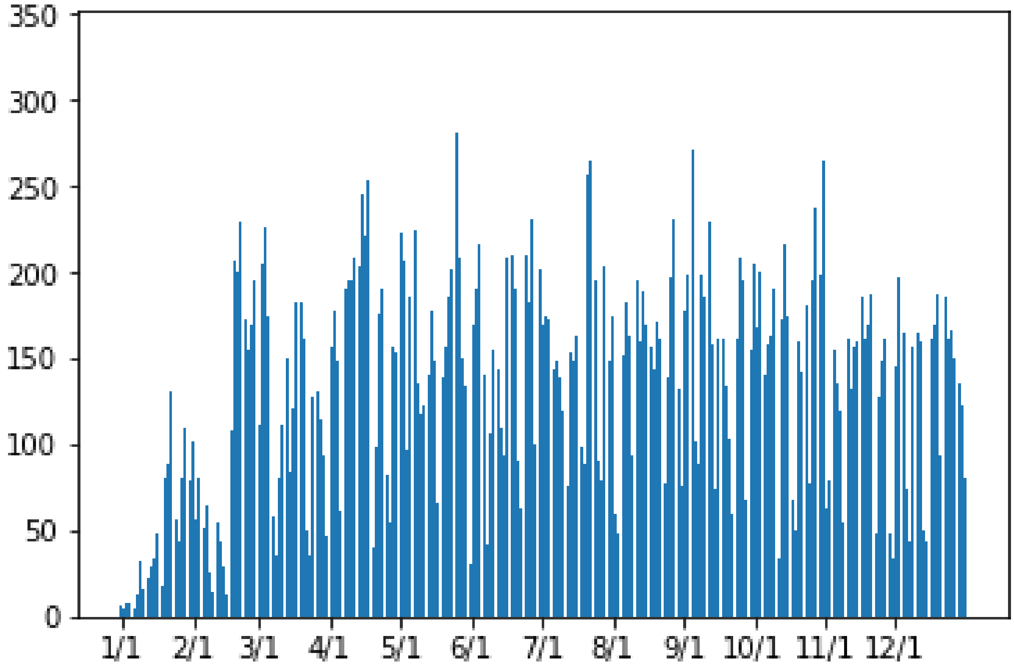

Figure 5.4.: Amount of articles per day, mentioning Ukraine published by RIA News in 2014.

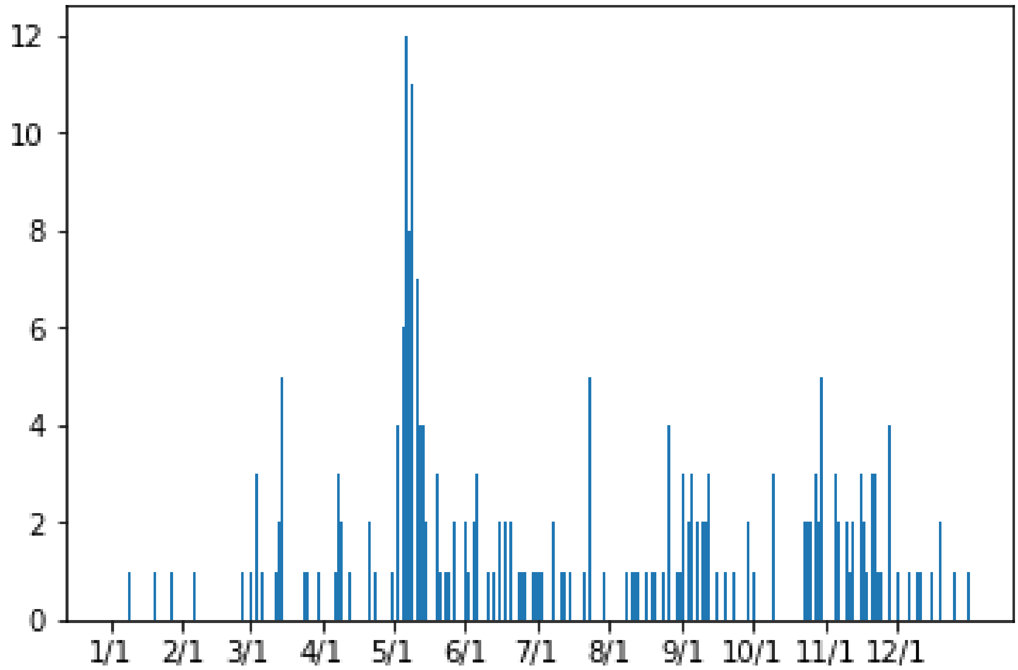

Figure 5.5.: Amount of articles per day, mentioning Ukraine and word ”fascism” together, published by RIA News in 2014.

6. Results

6.1. Shift in Narrative

As described in the Section 4.1.1, we use word2vec models to detect shifts in the media narrative. In the course of this work we use the implementation provided by Python package Gensim Řehůřek & Sojka (2010).

The articles collected from RIA News were divided into groups based on the year of publishing. After that the data was prepared for training word2vec models, preparation steps include:

- Dividing texts into sentences
- Removing punctuation
- Converting upper case to lower case
- Lemmatization of every word

Then for each year, separate word2vec models were trained. All models used CBOW architecture, described in the Section 2.5.3, and same sets of parameters. Following parameters were used:

- 'size': 100. The dimensionality of the feature vectors.
- 'window': 10. The maximum distance between the current and predicted word within a sentence.
- 'alpha': 0.025. The initial learning rate.
- 'min_count': 2. All words with a total frequency lower than this will be ignored.

Based on the reasons described in the Section 2.3.3 and the fact that there were 372 articles that mention Ukraine and fascism together, it was decided to investigate how similarity between Ukraine and word "fascism" was changing over the years. However, we noticed that similarity between words "fascism" and "country" or

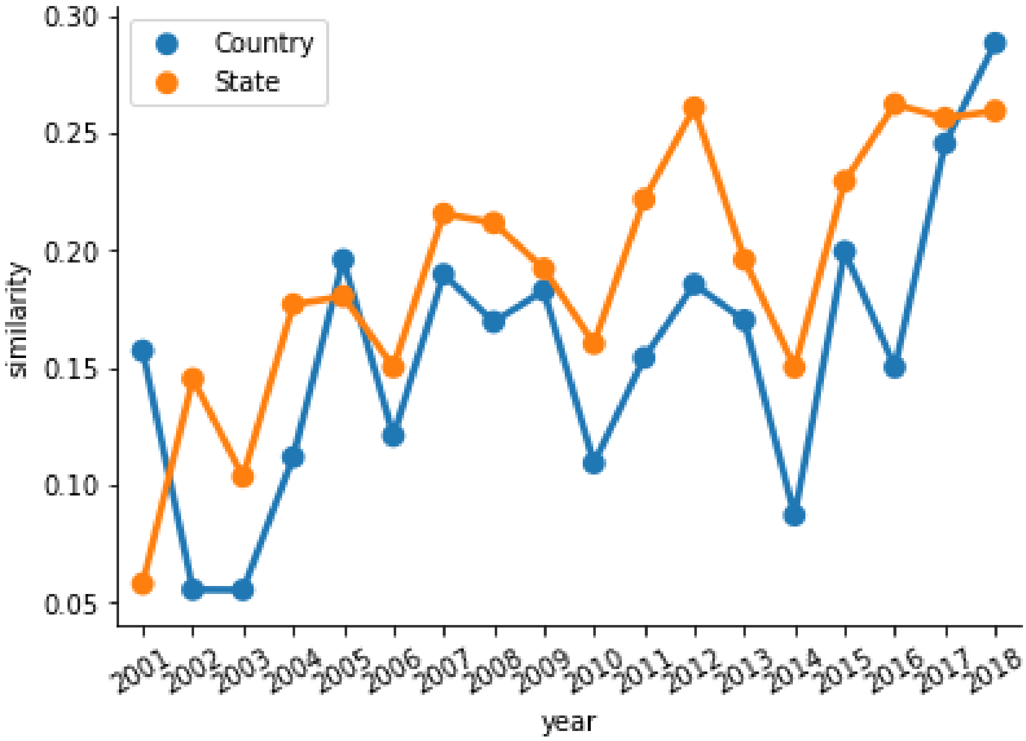

Figure 6.1.: Similarity between vector-words ”State”, ”Country” and ”fascism”

”state” was also changing over the years, which is demonstrated on the Figure 6.1. Due to this, when calculating the similarity between country and fascism, first the mean vector of word-vectors ”country” and ”state” were subtracted from word-vector of a country name.

To see if the changes in similarity for Ukraine are different from other countries, we compare the values concerning Ukraine to two sets of countries. First, ”West” countries, this are the states which historically where opposing Soviet Union and, later, Russia, which includes 12 countries that founded NATO[1] and Germany. Second set consists of Post-Soviet countries[2]. The Figure 6.2 shows changes in similarity

[1]NATO was founded on the 4th of April by USA, France, United Kingdom, Italy, Canada, Norway, Belgium, Netherlands, Denmark, Portugal, Iceland and Luxembourg

[2]Post-Soviet countries are Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan

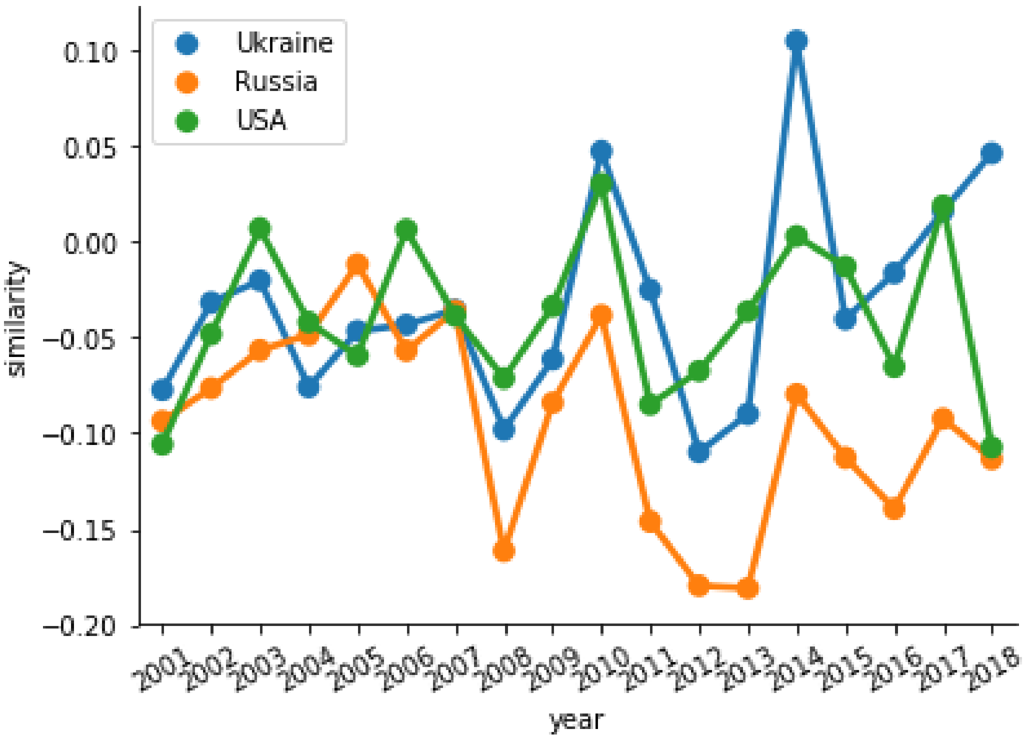

Figure 6.2.: Changes in similarity between country names and word "fascism".

with fascism for Ukraine, Russia and USA, measurements for the rest of the countries can be found in Appendix B.1 and B.2. It is visible, that in 2014 there was a peak in the similarity between Ukraine and fascism, with a value of 0.10505. What is more, Ukraine had the highest similarity among all countries used for comparison in that year. That leads to the next step – extracting articles that cause such behaviour.

6.2. Extracting suspected articles

To find articles that may have caused the increase in similarity between Ukraine and fascism, we divide all articles from 2014 that mention Ukraine in groups, than we train new word2vec models without data from selected groups. It was decided to divide data based on the topic provided in the URL of the page; for example,

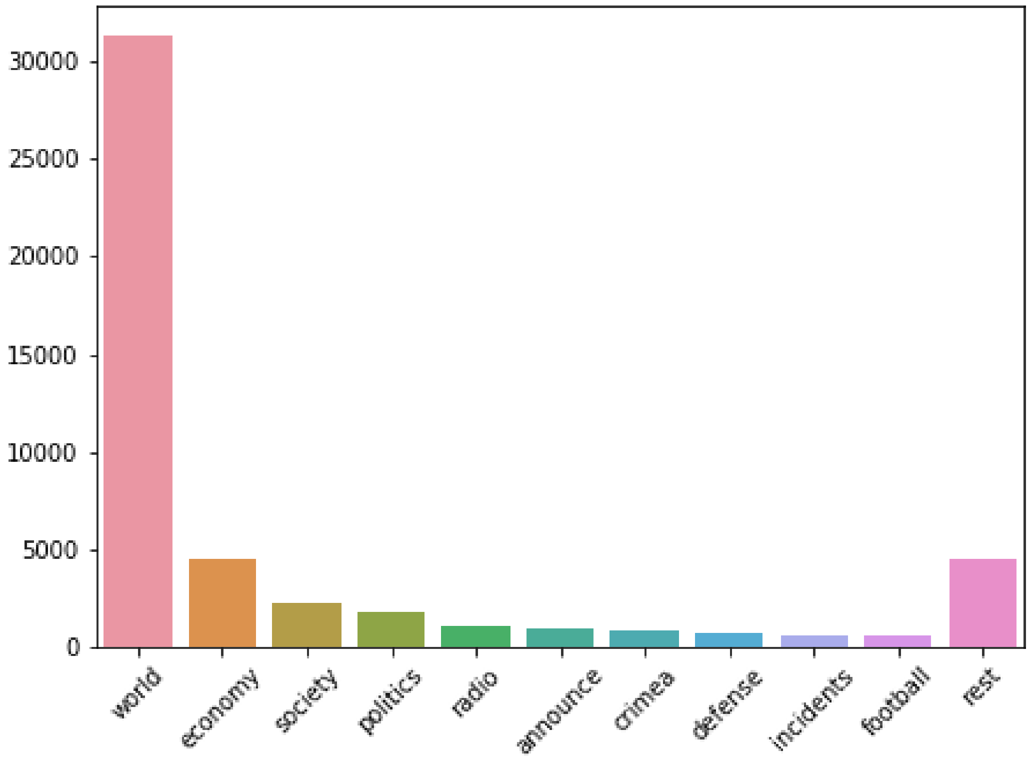

Figure 6.3.: Distribution of articles per topic. All articles were published in 2014 and have mentioning of Ukraine. In total there are 136 topics, but top 10 categories contain more than 90% of data, thus all other categories are joined into the group "rest"

https://ria.ru/politics/20011127/24002 has topic "politics". Overall, in 2014 articles about Ukraine were divided into 136 topics. However, as it can be seen on the Figure 6.3, amount of articles per topic are not equally distributed, where top 10 topics have 90.7% of all articles and topic "world" has 63%. In addition to those groups, we form one more group, which consist of articles that mention Ukraine and fascism together, and label it "fascism". Therefore, in total there are 12 groups: "world", "economy", "politics", "radio", "announce", "crimea", "defence", "incidents", "rest" (remaining 126 categories) and articles that mention Ukraine and Fascism together, labelled as "fascism".

As described in the Section 4.1.2, the goal is to find a group removing of which from training dataset minimise the similarity between Ukraine and fascism. The

Removed Group	Amount of Articles	Similarity(Ukraine, fascism)	difference
fascism	372	**0.015**	0.09
announce	967	0.082	0.023
economy	4554	0.089	0.016
world	31347	0.095	0.001
crimea	791	0.105	0.0
politics	1731	0.105	0.0
defense	745	0.108	-0.003
incidents	634	0.113	-0.008
society	2213	0.119	-0.013
rest	4553	0.126	-0.021
radio	1026	0.132	-0.027
football	562	0.134	-0.029

Table 6.1.: Measurement of similarity between Ukraine and fascism after removing selected group from dataset and retraining word2vec models. Groups are formed based on the topic provided in the URL of article plus additional group ”fascism”, which contains articles that mention Ukraine and fascism together

results of this process are presented in the Table 6.1. It can be observed, that removing articles, which mention Ukraine and fascism together, substantially decreases similarity, from 0.105 to 0.015. To check, if such decrease in similarity is unique to the word ”fascism”, we found all words that have about same absolute and relative frequency, and about the same similarity to Ukraine as word ”fascism”. In the articles that are published in 2014, there are two such words ”zamorozok” (slight frost) and ”razoruzhenie” (disarmament). From Table 6.2, it can be seen that the three words have initially nearly the same similarity, but only for ”fascism” the similarity was reduced to almost zero. Therefore, **articles that were published in 2014 and mention Ukraine and fascism together are suspected of being propagandistic**.

6.3. Investigating Principle of Long Lasting Impact

After determining articles that are suspected of propaganda, it is needed to analyse if those articles correspond to principles of propaganda. First one is the principle

Word	Occurrence	with Ukraine	Initial Similarity	Modified Similarity
fascism	5953	372	0.105	**0.015**
razoruzhenie	6884	367	0.152	0.145
zamorozok	5237	375	0.145	0.191

Table 6.2.: The table shows impact of removing articles on similarity measurement.

of long-lasting impact and, as it was described in the Section 4.2.1, to check if suspected article corresponds to this principle, it needs to train text classifier, which determines if the article contains mentioning of fascism. For this task we use Random Forest Classifier implemented in scikit-learn[3] (Pedregosa et al. 2011).

Since all the work in this thesis is focusing on the timeframe of Crimean occupation, we use only articles from the year 2014 for this analysis. These articles were divided into two groups: with mentioning of fascism and without. Then, for the training dataset, an equal amount of articles from both groups were sampled. After that, we remove the word "fascism" and country names from articles. Afterwards, all articles were converted to vector representation with word2vec model, which were trained on all articles published by RIA News. Finally, the classifier was evaluated with 10 fold validation method.

Resulting classifier had a mean accuracy of 86%. The bias is represented by the percentage of articles, that have no mentioning of fascism, but was classified as one that has. The Figure 6.4 demonstrates results for the subset of selected countries and the Appendix B.3 has the measurement of all countries used for comparison. It can be seen, that the bias is the highest in the case of Ukraine, reaching 46.1%.

6.4. Investigating Principle of Desired Information

Since RIA News have view counter on every article, it was possible to see which articles had more attention from readers. To see, if articles that mention Ukraine and fascism together have more views, articles were grouped by countries and averaged the number of views per article was calculated. However, we noticed that mentioning of Ukraine in the article greatly boost the number of views. Therefore,

[3]https://scikit-learn.org/stable/index.html

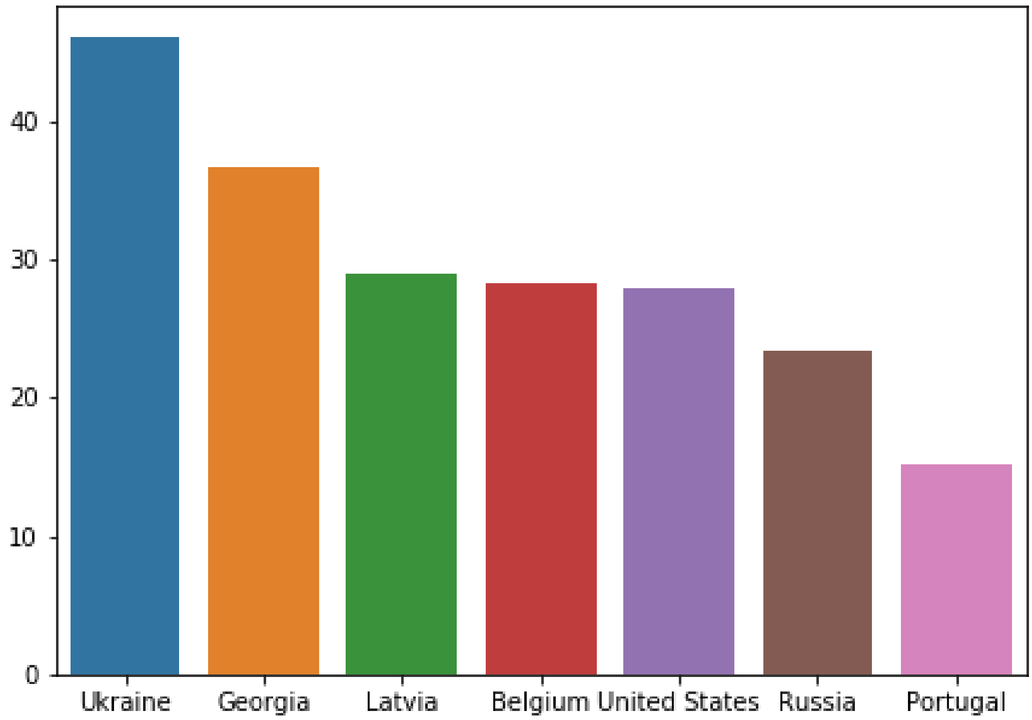

Figure 6.4.: Results of classifier bias, when determining if the article contains mentioning of fascism. Bias is the percentage of articles that did not have any mentioning of fascism, but was classified as one that has mentioning of fascism.

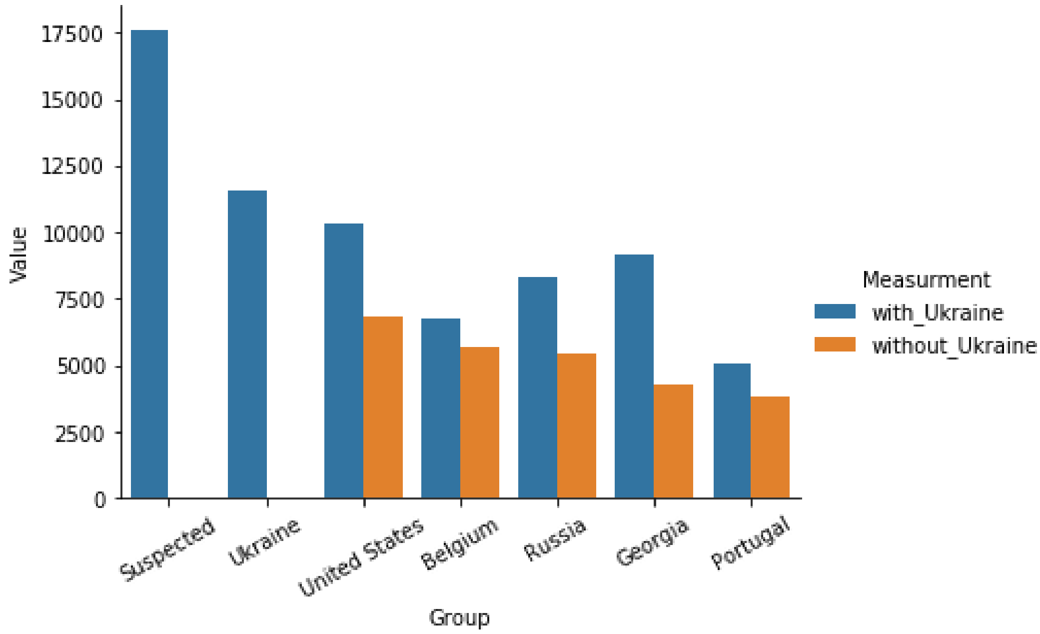

Figure 6.5.: Average amount of views per article. "Suspected" are articles that mention Ukraine and fascism together. Measurements were done with and without articles that mention Ukraine.

when counting the average amount of views for other counties, articles that mention Ukraine were excluded. In the case of Ukraine, articles that mention fascism were excluded, because those articles are in the suspected group. It can be seen in the Appendix B.5, that excluding articles for Ukraine drastically decrease the average amount of views per article, for example, articles that mention Moldova and Ukraine have on average 6063.7 views, without Ukraine – 2937.0. Moreover, articles that mention Ukraine have on average 11561.3 views per article, which is higher than any other country used for comparison, and articles that mention Ukraine and fascism together have 17635.9 average views per article. The great difference in average views can be observed on the Figure 6.5.

6.5. Investigating Principle of Emotional Agitation

As it was stated mentioned before, the goal of the principle of emotional agitation is to invoke negative emotions in the recipient of information. To investigate if suspected articles provoke negative emotions, we calculate the percentage of negative words per article. The list of negative words was obtained from Russian sentiment lexicon – RuSentiLex, which was formed by Loukachevitch & Levchik (2016). The lexicon has over 12 thousands words, which are labelled as ”positive”, ”neutral” or ”negative”. In the lexicon word ”fascism” and all derivative words are labelled as negative. Therefore, because it is known that all suspected articles have word ”fascism” in them, the word and all derivative forms were removed from the list. After that percentage of negative words per article – negativity, was calculated and averaged based on country.

Since suspected articles contain country names, they were ignored when measuring negativity. The Figure 6.6 demonstrates results for the subset of selected countries and the Appendix B.4 has the measurement of all countries used for comparison. It can be observed that all articles about Ukraine and not just ones that mention fascism on average have a higher percentage of negative words per article.

6.6. Investigating Clarity Principle

The next step is to analyse if the suspected articles were simplified in any matter. As it was described in the Section 4.2.4, for this task measurement of Flesh reading-ease score (FRS) and LIX score are used. The results can be found in the Appendix B.6 and for selected country visualisation is provided on the Figure 6.7. It can be observed that derivation of values between the suspected group and other is insignificant. For measuring repetition in the articles, all words were reduced to the stem. We do the stemming with the of Python package – Natural Language Toolkit (NLTK)[4]. Then, we calculate the percentage of unique stems per article. The Appendix B.7 shows the measurement for the average amount of words per article and the average percentage of unique stems, the Figures 6.8 and 6.10 visualise these measurements for selected countries. It can be observed that suspected articles have the highest amount of words, but those words are not unique. To see if such behaviour is

[4]https://www.nltk.org/

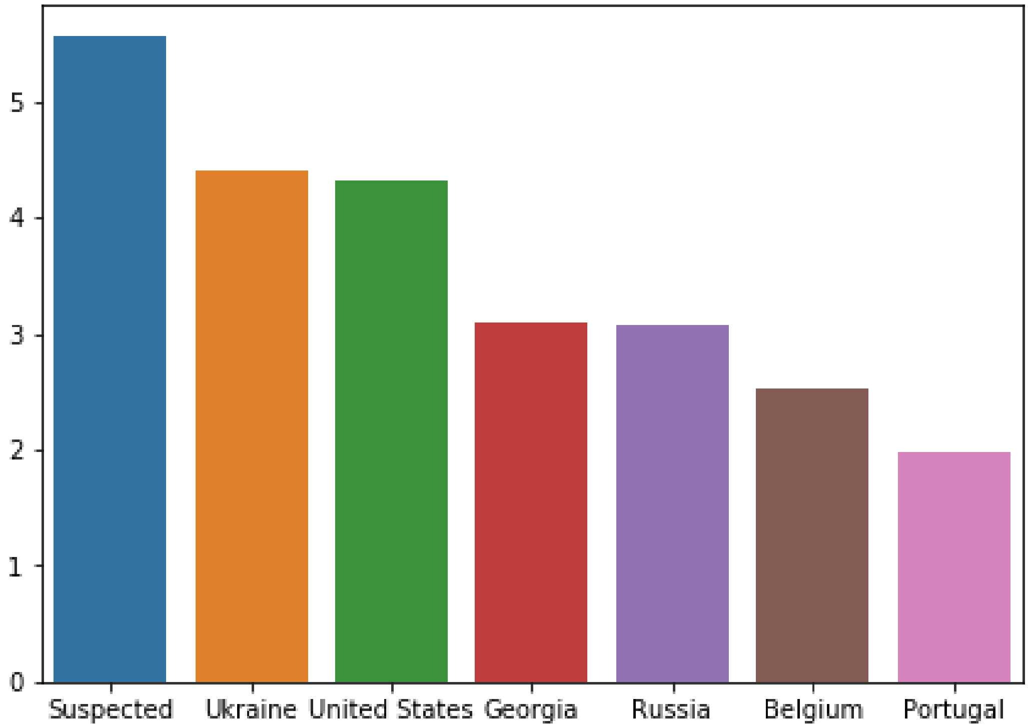

Figure 6.6.: Average percentage of negative words per article, without word ”fascism” and all derivative forms. ”Suspected” are articles that mention Ukraine and fascism together.

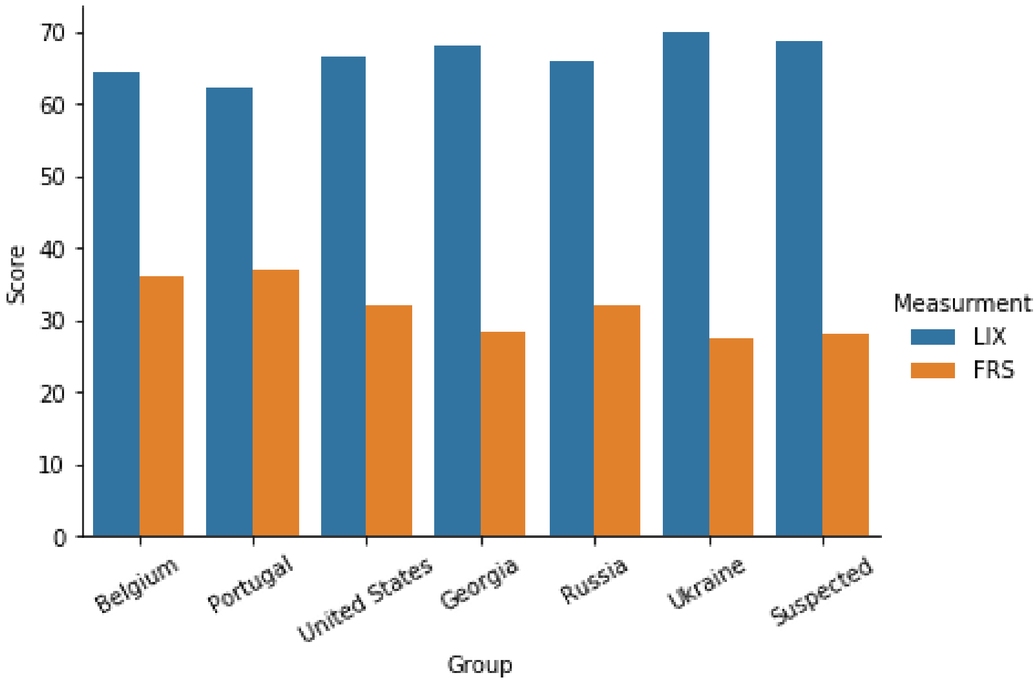

Figure 6.7.: Result of Flesh reading-ease score (FRS) and LIX for article grouped by country, and "Suspected" are articles that mention Ukraine and fascism together. Higher values of FRS represent higher simplicity. Lower values of LIX – higher simplicity.

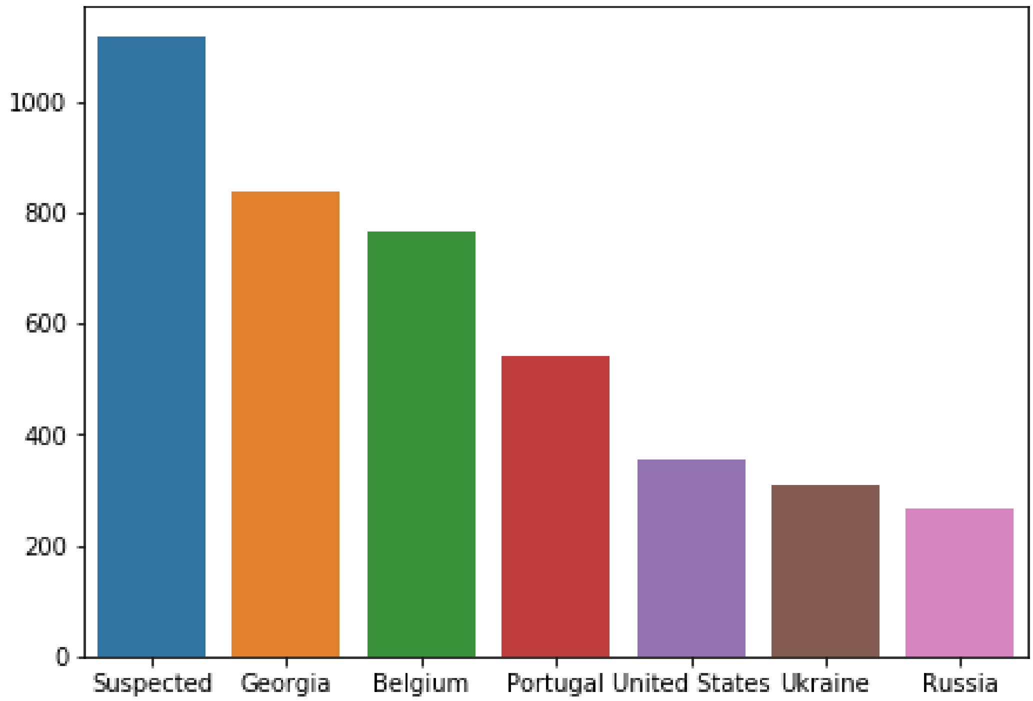

Figure 6.8.: Average amount of words per article grouped by country, and "Suspected" are articles that mention Ukraine and fascism together

expected from the texts, Heaps' β, with fixing parameter $k = 10$, for selected texts were calculated. It can be seen from Figure 6.10 that, the difference between values of β for different groups is negligible.

6.7. Investigating Principle of Supposed Obviousness

As mention before in the Section 4.2.5, term "Zapad" ("Запад") is translated as west. To distinguish, if the word "Zapad" refers to the collective name of western countries or the cardinal direction of the compass, we check the spelling. "Zapad" as the name is always written with a capital letter and as the direction – with the little letter, but it can be written with capital if the word is at the beginning of a sentence.

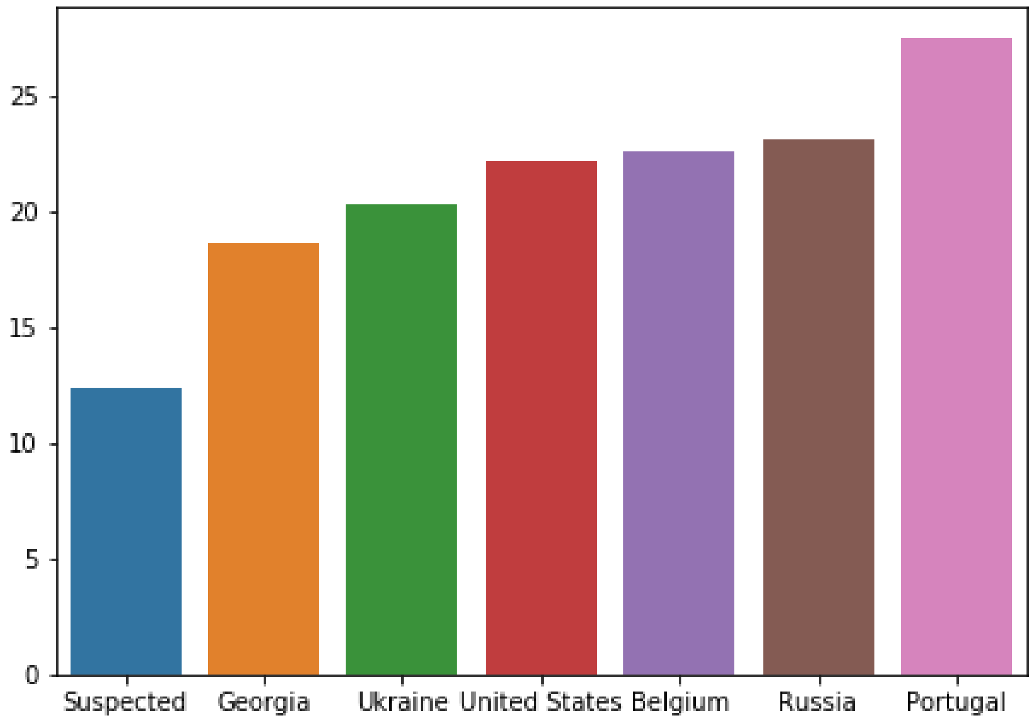

Figure 6.9.: Average amount of unique stems per article grouped by country, and "Suspected" are articles that mention Ukraine and fascism together

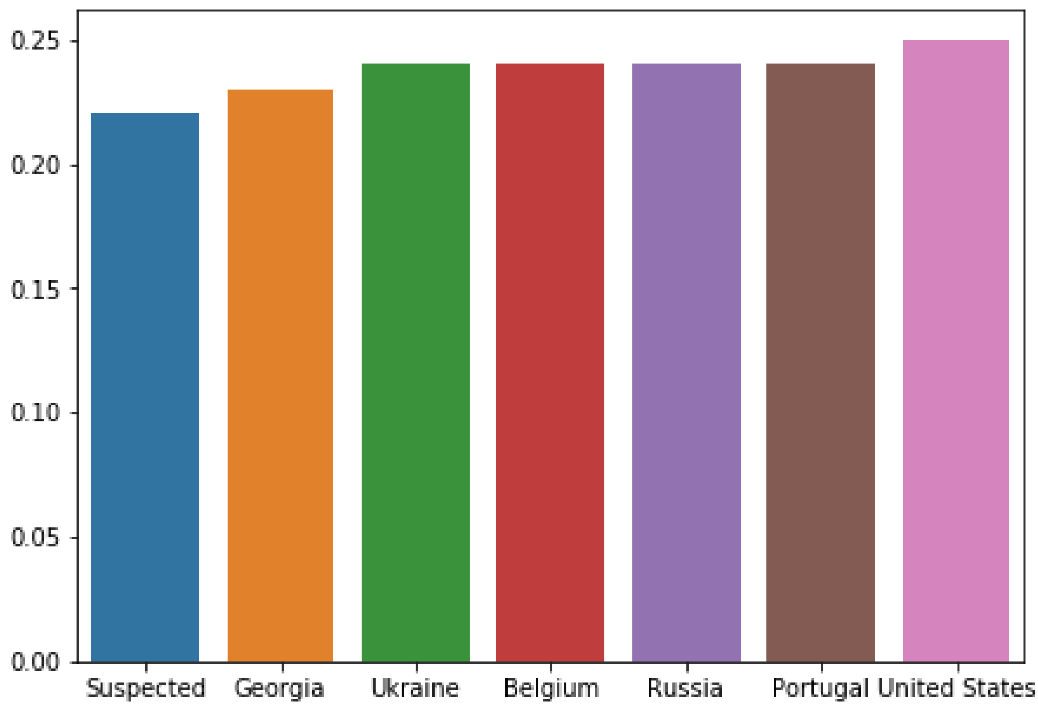

Figure 6.10.: Value of parameter β in Heaps' law for text from the group with fixing parameter $k = 10$. Articles are grouped by country and "Suspected" are articles that mention Ukraine and fascism together.

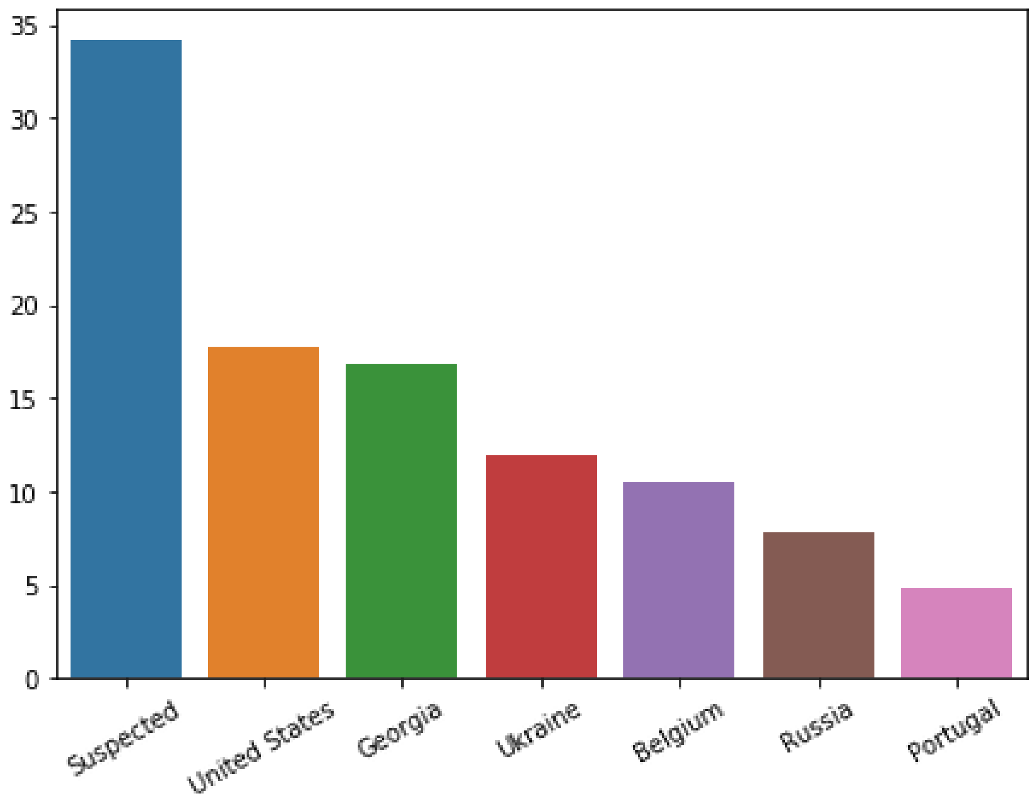

Figure 6.11.: Percentage of articles that contain the term ”Zapad” in the text. Term ”Zapad” has a negative context and is used for Western countries. Articles are grouped by country, and ”Suspected” are articles that mention Ukraine and fascism together.

Thus, to check if the article mentions the negative myth of ”Zapad”, we calculate the occurrence of word ”Zapad”, with a capital letter and not at the beginning of the sentence. The results for selected countries are shown on the Figure 6.11 and measurements for the rest can be found in the Appendix B.8. It can be seen, that more than 30% of suspected articles mention ”Zapad”.

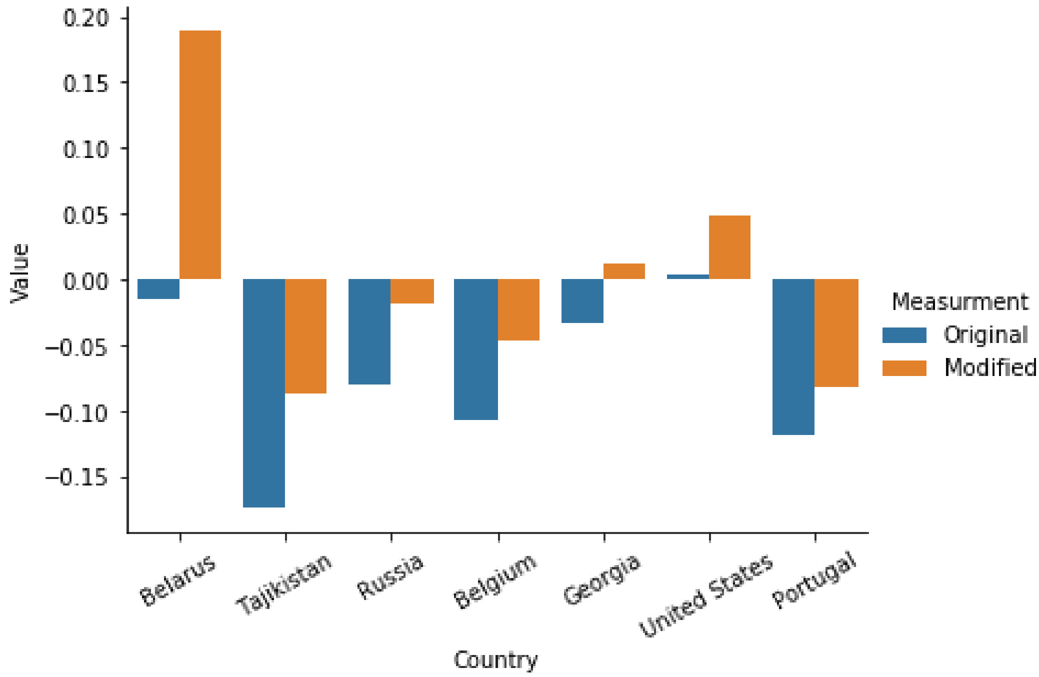

Figure 6.12.: Measurements for the similarity between country name minus mean of word-vectors ”state” and country”, and the word ”fascism”. Original are values obtained from word2vec models trained on articles from 2014. Modified – are from word2vec model trained with the injection of the modified version of suspected articles, where Ukraine was replaced with selected country name.

6.8. Measure Influence on Other Countries

The last step of the analysis is to see if the suspected articles would have had the same impact on other countries. To achieve this goal, every article from the suspected group word ”Ukraine” is replaced with country name from the set used for comparison. After that, word2vec models are retrained with these new ”fake” articles. Then similarities between country name minus mean of word-vectors ”state” and ”country”, and ”fascism” from the original model and new model are compared. Measurements for all countries used for comparison can be found in the Appendix B.9 and some values are visualized in the Figure 6.12. It can be seen that for most of the countries adding modified articles to the training set increased similarity to fascism. In the case of Belarus, the similarity raised tremendously.

7. Discussion

Performed analysis of RIA News articles confirmed the observation of Bebler (2015) and Humen (2017) – during the period of the Crimean annexation, Russian news media were portraying Ukraine as a fascist country. This fact can be observed in the shift of media narrative about Ukraine, measured in the similarity between Ukraine and fascism. The sudden jump in this measurement, from -0.08986 in 2013 to 0.10505 in 2014, indicates that the change of similarity did not happen naturally.

Then, the amount of articles that mention Ukraine and fascism together, which was reaching more than 10 such articles per day, seems to be artificial and not corresponding to real life events. Moreover, removing those articles from training dataset decreased similarity between Ukraine and fascism from 0.10505 to 0.01502. The new value seems to be more corresponding to the value from 2013. Thus, we selected these 372 articles for further analysis.

Further analysis showed that machine models trained on the articles from 2014, would be biased to classify articles about Ukraine as an article that mentions fascism, even if the article has none of it. Comparing it to the bias against other countries, showed that for Ukraine bias is the highest. Which indicates that suspected article corresponds to the principal of long-lasting impact, as even machines, not talking about humans, become biased against Ukraine after reading articles published by RIA News.

Then, we showed that the articles about Ukraine are most viewed in comparison to other countries. However, if the word "fascism" mentions the number increased even more. It confirms that information was presented in the way that attracts the attention of the reader, which is the exact goal of the principle of desired information.

Afterwards, it was shown that the articles about Ukraine have the highest amount of negative words per article, and this is without counting the word "fascism" and all derivative forms. Moreover, the articles that mention Ukraine and fascism together have an even a higher percentage of negativity. Taking into consideration the role which World War II and fascism as the depiction of true enemy plays in Russian collective memory, it is reasonable to expect that such articles will cause negative emotions connected to Ukraine. Such a strategy is a perfect example of a principal of emotional agitation.

The analysis has shown that, in terms of readability, suspected articles are not different from other articles. However, it is essential to note that newspapers articles are aimed at the mass population, therefore they are written in a way that average adult person can understand them without any problems. Nonetheless, it was noticed that suspected articles are different aspects – they have much more words per article, but the number of unique words are smaller, meaning the repetition is used. When comparing to other articles based on Heaps' law, it seems that such reception is expected from the Russian language.

Furthermore, to add oil to the fire, more than a third part of suspected articles mention ”Zapad”, already established negative myth. Combining it with the usage of ”fascism”, the reader should feel that it is obvious that nothing good can come from Ukraine.

The last, step of the analysis revealed that replacing Ukraine with other countries can increase the similarity to fascism. Meaning that such articles can be created for any country, especially knowing that Kremlin pushes the ”correct” version of history, in which anyone can be fascist.

Considering all of this, the likelihood of those articles being specially designed as a tool of propaganda is very high. However, the truth can be discovered only with a flow of time, when these events will be just history and all secret documents will be published.

Coming back to the research questions stated in the Chapter 1:

- Based on what measurements newspaper article can be considered propagandistic?
- What methods of Natural Language Processing and Machine Learning can be used for detection of propaganda?

Propaganda can have different types, forms and goals, that is why there can not be one measurement to detect all of it. In this work propaganda which is aimed at building a negative image of the target group was analysed. For this case, we use the measurement of similarity between the targeted group and the negative term. For this task, word2vec model can be used, because this model can capture the relation between two terms. However, using only this measurement is not enough, that is why in the course of these work methods for measuring correspondence to principles of propaganda described by Darczewska (2014). Principles and methods are as follows:

- Long-Lasting Impact – classifier bias
- Desired Information – measuring the interaction of a user with propagandistic material
- Emotional Agitation – sentiment analysis with a focus on negative emotions
- Clarity – readability scores and repetition detection
- Supposed Obviousness – measuring co-occurrence with already established myth.

All the methods, except the one used principle of the principal of supposed obviousness, can be adapted for different languages and settings. The problem of detecting supposed obviousness is that one would have to know the history and geopolitics of countries involved in the conflict.

8. Conclusion

> Si vis pacem, para bellum
>
> Vegetius, De Re Militari

Propaganda is one of the biggest problems in the modern world because it provokes conflicts which can lead to a great loss of human life. The annexation of Crimea and following conflict in Eastern Ukraine lead to thousands of lost lives and millions of displaced people. The lack of research on the topic of unsupervised propaganda detection led us to introduce the first iteration of measurements which can be used for quantifying qualities of propaganda.

In conclusion, the advantage of proposed methods is the fact that they can be easily adapted to different languages and settings. However, they require the understanding of the geopolitics of involved countries. We plan to continue working on this research and apply these methods, to the different occurrence of propaganda such instances of propaganda as jihadism, white supremacy and other cases of extremism.

Bibliography

Aro, J. (2016), ‘The Cyberspace War: Propaganda and Trolling as Warfare Tools’, *European View* **15**(1), 121–132.

Ashcroft, M., Fisher, A., Kaati, L., Omer, E. & Prucha, N. (2015), Detecting jihadist messages on twitter, *in* ‘2015 European Intelligence and Security Informatics Conference’, IEEE, pp. 161–164.

Awan, I. (2017), ‘Cyber-extremism: Isis and the power of social media’, *Society* **54**(2), 138–149.

Bachrach, S., Phillips, E., Luckert, S. & Museum, U. S. H. M. (2009), *State of Deception: The Power of Nazi Propaganda*, United States Holocaust Memorial Museum.
URL: *https://books.google.at/books?id=hzJ7PwAACAAJ*

Badawy, A. & Ferrara, E. (2018), ‘The rise of Jihadist propaganda on social networks’, *Journal of Computational Social Science* **1**(2), 453–470.
URL: *https://doi.org/10.1007/s42001-018-0015-z*

Badawy, A., Ferrara, E. & Lerman, K. (2018), Analyzing the digital traces of political manipulation: The 2016 russian interference twitter campaign, *in* ‘2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)’, IEEE, pp. 258–265.

Balzer, H. (2015), ‘The Ukraine Invasion and Public Opinion’, *Georgetown Journal of International Affairs* **16**(1), 79–93.
URL: *http://www.jstor.org/stable/43773670*

Bebler, A. (2015), ‘Crimea and the Russian-Ukrainian Conflict’, *Romanian Journal of European Affairs* **15**(1), 35–54.

Bengio, Y., Ducharme, R., Vincent, P. & Jauvin, C. (2006), *Neural Probabilistic Language Models*, Vol. 3, pp. 137–186.

Berger, J. M. & Morgan, J. (2015), 'The isis twitter census: Defining and describing the population of isis supporters on twitter', *The Brookings Project on US Relations with the Islamic World* **3**(20), 4–1.

Bishop, C. M. (2006), *Pattern recognition and machine learning*, springer.

Björnsson, C.-H. (1983), 'Readability of newspapers in 11 languages', *Reading Research Quarterly* pp. 480–497.

Bohdan, A. (2019), 'Ria news corpus 2001-2018'.
URL: *https://doi.org/10.5281/zenodo.2647439*

Breiman, L. (2001), 'Random forests', *Machine learning* **45**(1), 5–32.

Breiman, L., Friedman, J., Olshen, R. & Stone, C. (1984), *Classification and Regression Trees (CART)*, Vol. 40.

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C. & Dredze, M. (2018), 'Weaponized health communication: Twitter bots and russian trolls amplify the vaccine debate', *American journal of public health* **108**(10), 1378–1384.

Carman, D. (2002), 'Translation and Analysis of the Doctrine of Information Security of the Russian Federation: Mass Media and the Politics of Identity', *Pacific Rim Law & Policy Journal Association* **II**(2), 339–362.

Chivvis, C. (2017), 'Understanding Russian "Hybrid Warfare": And What Can Be Done About It: Addendum', *Understanding Russian "Hybrid Warfare": And What Can Be Done About It: Addendum* .

Cunningham, S. (2002), *The Idea of Propaganda: A Reconstruction*, Praeger.
URL: *https://books.google.at/books?id=2kCFgv6FzuUC*

Darczewska, J. (2014), *The Anatomy of Russian Information Warfare. The Crimean Operation. A Case Study*, Vol. May 2014.
URL: *http://www.osw.waw.pl/en/publikacje/point-view/2014-05-22/anatomy-russian-information-warfare-crimean-operation-a-case-study*

Egghe, L. (2007), 'Untangling herdan's law and heaps' law: Mathematical and informetric arguments', *Journal of the American Society for Information Science and Technology* **58**(5), 702–709.

Ellul, J. (1973), *Propaganda: the formation of men's attitudes*, Vintage Books.

Fedor, J. (2015), 'Introduction: Russian Media and the War in Ukraine', *Journal of Soviet and Post-Soviet Politics and Society* **1**(1), 1–12.

Ferrara, E. (2017), 'Contagion dynamics of extremist propaganda in social networks', *Information Sciences* **418-419**, 1–12.
URL: *http://www.sciencedirect.com/science/article/pii/S0020025517305030*

Flesch, R. (1979), 'How to write plain english: Let's start with the formula', *University of Canterbury* .

Freedom House (2017), 'Press Freedom's Dark Horizon', *Freedom of the Press Report 2017* (April), 30.
URL: *https://freedomhouse.org/sites/default/files/FOTP_2017_booklet_FINAL_April28.pdf*

Gelbukh, A. & Sidorov, G. (2001), Zipf and heaps laws' coefficients depend on language, *in* 'International Conference on Intelligent Text Processing and Computational Linguistics', Springer, pp. 332–335.

Goldberg, Y. & Hirst, G. (2017), *Neural Network Methods in Natural Language Processing*, Morgan & Claypool Publishers.

Goodman, J. T. (2001), 'A Bit of Progress in Language Modeling', *Computer Speech & Language* **15**, 403–434.

Grishman, R. & Sundheim, B. (1996), Message understanding conference-6: A brief history, *in* 'Proceedings of the 16th Conference on Computational Linguistics - Volume 1', COLING '96, Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 466–471.
URL: *https://doi.org/10.3115/992628.992709*

Haykin, S. (2009), *Neural Networks and Learning Machines.*
URL: *https://arxiv.org/pdf/1312.6199v4.pdf*

Holm, R. R. (2017), Natural language processing of online propaganda as a means of passively monitoring an adversarial ideology, Technical report, Naval Postgraduate School Monterey United States.

Humen, Y. (2017), 'Five Principles of Media Propaganda and Russian Geopolitical Expansionism', *Dialog media studios* .

Johnson, R. & Zhang, T. (2015), Semi-supervised convolutional neural networks for text categorization via region embedding, *in* 'Advances in neural information processing systems', pp. 919–927.

Johnston, A. H. & Weiss, G. M. (2017), Identifying sunni extremist propaganda with deep learning, *in* '2017 IEEE Symposium Series on Computational Intelligence (SSCI)', IEEE, pp. 1–6.

Jowett, G. & O'Donnell, V. (1986), *Propaganda and Persuasion*, SAGE Publications.
URL: *https://books.google.at/books?id=N0eRQgAACAAJ*

Kaati, L., Omer, E., Prucha, N. & Shrestha, A. (2015), Detecting multipliers of jihadism on twitter, *in* '2015 IEEE International Conference on Data Mining Workshop (ICDMW)', IEEE, pp. 954–960.

Kim, Y., Jernite, Y., Sontag, D. & M. Rush, A. (2015), 'Character-aware neural language models'.

Liaw, A., Wiener, M. et al. (2002), 'Classification and regression by randomforest', *R news* **2**(3), 18–22.

Libicki, M. C. & University, N. D. (1995), *What is information warfare? / by Martin C. Libicki*, National Defense University, Institute for National Strategic Studies [Washington, D.C.] .
URL: *http://purl.access.gpo.gov/GPO/LPS14211*

Liddy, E. D. (2001), 'Natural Language Processing. In Encyclopedia of Library and Information Science', *Encyclopedia of Library and Information Science* .
URL: *https://surface.syr.edu/cnlp*

Lilleberg, J., Zhu, Y. & Zhang, Y. (2015), Support vector machines and word2vec for text classification with semantic features, *in* '2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)', IEEE, pp. 136–140.

Liu, B. (2012), 'Sentiment analysis and opinion mining', *Synthesis lectures on human language technologies* **5**(1), 1–167.

Liu, H. (2017), 'Sentiment analysis of citations using word2vec', *arXiv preprint arXiv:1704.00177* .

Loukachevitch, N. V. & Levchik, A. (2016), Creating a general russian sentiment lexicon., *in* 'LREC'.

Macdonald, S. (2014), *Propaganda and Information Warfare in the Twenty-First Century*.

Mijnssen, I. (2010), 'The Victory Myth and Russia's Identity', *Russian Analytical Digest (RAD)* pp. 6–9.

Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013), 'Efficient estimation of word representations in vector space', *Proceedings of Workshop at ICLR* **2013**.

Miller, G. A. & Nicely, P. E. (1955), 'An analysis of perceptual confusions among some english consonants', *The Journal of the Acoustical Society of America* **27**(2), 338–352.

Ministry of Foreign Affairs of the Russian Federation (2000), 'Information Security Doctrine of the Russian Federation'.

Nadeau, D. & Sekine, S. (2007), 'A survey of named entity recognition and classification', *Lingvisticae Investigationes* **30**.

Nelson, R. (1996), *A Chronology and Glossary of Propaganda in the United States*, Greenwood Press.
URL: *https://books.google.at/books?id=ySwYAAAAIAAJ*

Nilsson, N. J. (1996), 'Introduction to machine learning: An early draft of a proposed textbook'.

Oborneva, I. (2005), 'Automate the assessment of the quality of text perception [in Rus.]', *Bulletin of the Moscow City Pedagogical University. Series: Informatics and Informatization of Education* (5), 86–91.

Office for the UN High Commissioner for Human Rights (OHCHR) (2018), 'Report on the human rights situation in Ukraine, 16 May to 15 August 2018', (August).
URL: *https://www.ohchr.org/Documents/Countries/UA/ReportUkraineMay-August2018_EN.pdf*

Office of the United Nations High Commissioner for Human Rights (2017), 'Situation of human rights in the temporarily occupied Autonomous Republic of Crimea and the city of Sevastopol (Ukraine)', pp. 1–33.
URL: *https://www.ohchr.org/Documents/Countries/UA/Crimea2014_2017_EN.pdf*

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011), 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research* **12**, 2825–2830.

Peters, A. (2015), *The Crimean Vote of March 2014 as an Abuse of the Institution of the Territorial Referendum*, pp. 278–303.

Powers, D. M. (2011), 'Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation'.

Řehůřek, R. & Sojka, P. (2010), Software Framework for Topic Modelling with Large Corpora, *in* 'Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks', ELRA, Valletta, Malta, pp. 45–50. `http://is.muni.cz/publication/884893/en`.

Rexha, A., Kröll, M., Dragoni, M. & Kern, R. (2016), Polarity Classification for Target Phrases in Tweets: A Word2Vec Approach, *in* H. Sack, G. Rizzo, N. Steinmetz, D. Mladenić, S. Auer & C. Lange, eds, 'The Semantic Web', Springer International Publishing, Cham, pp. 217–223.

RIA Novosti (2013), 'RIA NOVOSTI'.
URL: *https://web.archive.org/web/20130928023413/http://rianovosti.com/docs/about/novosti.html*

Rojansky, M. & Kofman, M. (2015), 'Kennan Cable No.7: A Closer look at Russia's "Hybrid War"', (7).
URL: *https://www.wilsoncenter.org/publication/kennan-cable-no7-closer-look-russias-hybrid-war*

Rosenblatt, F. (1958), 'The perceptron: a probabilistic model for information storage and organization in the brain.', *Psychological review* **65**(6), 386.

Russia country profile (2012).
URL: *http://news.bbc.co.uk/2/hi/europe/country_profiles/1102275.stm*

Sebastiani, F. (2002), 'P1-Sebastiani', *ACM Computing Surveys* **34**(1), 1–47.
URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.6513*

Siencnik, S. K. (2015), Adapting word2vec to named entity recognition, *in* 'Proceedings of the 20th nordic conference of computational linguistics, nodalida 2015, may 11-13, 2015, vilnius, lithuania', number 109, Linköping University Electronic Press, pp. 239–243.

Smith, B. L. & Lasswell, H. (1946), 'Propaganda, Communication, and Public Opinion', *Social Forces* **25**(2), 229–230.
URL: *https://dx.doi.org/10.2307/2571574*

Stewart, L. G., Arif, A. & Starbird, K. (2018), Examining trolls and polarization with a retweet network, *in* 'Proc. ACM WSDM, Workshop on Misinformation and Misbehavior Mining on the Web'.

Stop Fake (2014), 'Fake: Crucifixion in Slovyansk'.
URL: *https://www.stopfake.org/en/lies-crucifixion-on-channel-one/*

Stukal, D., Sanovich, S., Bonneau, R. & Tucker, J. A. (2017), 'Detecting bots on russian political twitter', *Big data* **5**(4), 310–324.

Szafranski, R. (1997), 'A Theory of Information Warfare; Preparing for 2020', p. 12.

Thornton, R. & Karagiannis, M. (2016), 'The Russian Threat to the Baltic States: The Problems of Shaping Local Defense Mechanisms', *The Journal of Slavic Military Studies* **29**(3), 331–351.
URL: *https://doi.org/10.1080/13518046.2016.1200359*

Ukraine and Russian Federation (1997), 'Treaty on friendship, cooperation and partnership between Ukraine and the Russian Federation', (52240).
URL: *https://treaties.un.org/doc/Publication/UNTS/No Volume/52240/Part/I-52240-08000002803e6fae.pdf*

Vázquez-Liñán, M. (2017), 'Historical memory and political propaganda in the Russian Federation', *Communist and Post-Communist Studies* **50**(2), 77–86.
URL: *http://www.sciencedirect.com/science/article/pii/S0967067X17300156*

Verkhovna Rada of Ukraine (2015), 'On the Appeal of the Verkhovna Rada of Ukraine to the United Nations, the European Parliament, the Parliamentary Assembly of the Council of Europe, the NATO Parliamentary Assembly, the OSCE Parliamentary Assembly, the GUAM Parliamentary Assembly, the national parliaments of the countries of the world on the recognition of the Russian Federation as an aggressor state'.
URL: *https://zakon.rada.gov.ua/laws/card/en/129-19*

Volkogonov, D. (1983), 'Psychological warfare', *Moscow: Voyenizdat [in Rus.]* .

Williams, P. a. H. (2010), ‘Information warfare: Time for a redefinition’, *Australian Information Warfare and Security Conference* (December), 37–44.
URL: *http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1036&context=isw*

Wood, E. A. (2011), ‘Performing memory: Vladimir Putin and the celebration of World War II in Russia’, *Soviet and Post Soviet Review* **38**(2), 172–200.

Zhang, D., Xu, H., Su, Z. & Xu, Y. (2015), ‘Chinese comments sentiment classification based on word2vec and svmperf’, *Expert Systems with Applications* **42**(4), 1857–1863.

Čižik, T. (2017), *Information Warfare - New Security Challenge for Europe*, Centre for European and North Atlantic Affairs (CENAA).

Appendix

Appendix A.

RIA News Statistic

year	total amount	about Ukraine
2001	23771	654
2002	127046	4893
2003	93308	4622
2004	137216	9578
2005	95383	5764
2006	70543	4403
2007	83411	4976
2008	119732	7439
2009	143672	9389
2010	173895	9158
2011	202427	8673
2012	247691	9630
2013	282449	10328
2014	246622	49123
2015	313816	50075
2016	349877	34793
2017	265033	24408
2018	197187	17860

Table A.1.: Total amount of articles per year published by RIA News and amount of articles that mention Ukraine

year	Belgium	Canada	Denmark	France	Iceland	Italy	Luxembourg	Netherlands	Norway	Portugal	United Kingdom	United States	Germany
2001	321	229	124	758	26	507	27	112	218	192	690	4471	820
2002	1145	1523	1053	3934	202	2639	387	731	811	856	3331	19359	5224
2003	773	944	419	3937	81	1997	167	910	441	391	3829	15796	4057
2004	1185	1231	567	4769	109	3164	462	1323	828	993	3716	14507	4468
2005	758	831	387	3411	98	2017	504	742	542	334	3005	10767	3321
2006	688	805	481	3042	121	2136	219	550	532	372	2775	9605	3109
2007	766	905	450	3736	97	2342	247	686	501	529	3962	12030	3696
2008	1031	1570	705	5920	372	3423	432	1187	746	660	4480	16349	5256
2009	1603	2098	1221	6365	276	4030	384	1389	1119	815	4846	17792	6438
2010	2022	2410	1136	7396	1025	3907	420	1634	1330	1327	5208	20387	7195
2011	2791	3795	2395	9781	438	7323	578	1486	2484	1628	6930	22736	8251
2012	1903	3582	1570	9739	260	7804	683	2071	1885	2131	7528	23888	10339
2013	2166	3725	1610	9753	575	7414	695	2539	2275	1564	7095	27271	9677
2014	2066	5274	1309	9157	264	6393	558	2980	3182	1398	6275	31544	10652
2015	3368	5162	1862	21281	548	8456	1304	3447	2873	1677	9184	44298	21048
2016	3897	5097	1752	16265	880	7970	794	3639	2923	1886	11302	52352	14454
2017	2119	3645	1254	11361	375	5394	468	2150	1722	1737	7245	47323	11946
2018	1537	3559	1629	8777	305	4142	342	1904	1701	949	7755	34314	8599

Table A.2.: Amount of articles per year publish by RIA News. Presented countries are considered to be ”West” countries, meaning 12 countries founders of NATO and Germany

year	Armenia	Azerbaijan	Belarus	Estonia	Georgia	Kazakhstan	Kyrgyzstan	Latvia	Lithuania	Moldova	Russia	Tajikistan	Turkmenistan	Ukraine	Uzbekistan
2001	194	190	26	129	494	340	210	182	137	26	9196	425	49	654	275
2002	1314	1662	425	1255	3318	2874	2052	1228	1357	180	46814	1968	396	4893	1413
2003	1774	1616	231	716	2135	2608	1641	886	887	160	34668	1486	255	4622	1034
2004	3349	2724	282	1548	8276	3472	1789	1814	1720	375	44287	2167	318	9578	1723
2005	829	1059	179	884	2956	1880	2163	1020	1014	310	31932	886	154	5764	1214
2006	791	767	150	567	2975	1265	820	700	385	259	25612	611	204	4403	609
2007	831	899	157	1552	3021	1832	846	704	568	141	30973	534	215	4976	643
2008	1429	1302	233	1136	6964	2204	887	1250	967	173	45301	752	223	7439	911
2009	1860	1670	384	1346	4054	2829	1222	1888	1118	441	51074	954	332	9389	1231
2010	1936	1506	450	1125	2656	3779	2868	1352	1059	284	62245	1236	291	9158	1241
2011	2007	1455	540	1312	2186	3728	1334	1515	1637	248	73023	1251	215	8673	1047
2012	1733	1891	546	1227	2236	3833	1337	1896	1731	257	89972	1120	196	9630	1368
2013	1967	2013	443	1318	2818	4199	1433	2186	1806	190	99241	1196	145	10328	1639
2014	1850	1682	490	1482	2535	4089	1257	2365	1660	380	101398	1122	182	49123	1388
2015	3232	2304	786	2222	3319	5874	2539	3103	2030	554	130863	2004	330	50075	1754
2016	5812	5125	1090	2375	3388	6314	3058	2882	2557	547	143959	1786	327	34793	2597
2017	2409	2234	735	1501	3116	4189	1757	1994	1706	387	119422	986	212	24408	1443
2018	2141	1303	602	1151	1796	2527	810	1753	1393	305	92263	630	175	17860	913

Table A.3.: Amount of articles per year publish by RIA News. Presented countries are Post-Soviet countries

day	January	February	March	April	May	June	July	August	September	October	November	December
1	6	102	111	165	94	30	202	175	177	205	63	145
2	4	56	112	156	223	169	169	60	199	168	79	197
3	7	81	205	178	207	190	175	48	271	201	111	192
4	8	64	226	149	96	217	173	186	225	55	99	189
5	9	70	174	61	185	160	77	153	272	21	155	165
6	9	52	200	21	229	170	61	152	101	141	135	74
7	4	65	159	163	221	141	144	183	89	159	119	43
8	13	26	58	191	224	42	148	163	199	163	54	157
9	32	14	36	196	136	106	139	94	185	190	52	213
10	16	31	80	195	117	155	119	50	213	191	149	135
11	8	25	112	208	122	166	166	144	195	76	161	164
12	7	55	148	75	232	128	83	196	230	33	133	160
13	22	43	124	117	178	144	75	160	159	172	157	50
14	28	29	150	203	141	109	153	189	74	217	160	44
15	34	13	84	245	177	94	149	169	162	174	77	158
16	48	11	121	221	149	209	163	70	197	163	81	156
17	50	42	183	254	66	250	282	66	183	208	186	162
18	12	108	179	166	56	212	335	156	162	68	161	170
19	17	206	153	61	181	210	98	144	134	50	169	188
20	81	201	182	40	138	190	89	171	103	160	188	94
21	88	229	162	98	156	91	257	162	60	142	129	56
22	130	159	50	176	185	62	265	207	136	158	50	166
23	95	97	35	191	202	174	215	96	161	168	49	185
24	116	173	128	224	64	192	224	78	161	181	128	161
25	56	155	127	202	187	210	195	138	209	77	149	167
26	44	170	118	82	281	183	90	197	196	196	161	150
27	81	195	130	54	208	231	79	231	67	238	204	70
28	109	197	115	156	150	100	204	226	56	176	124	58
29	73	–	94	153	134	40	220	235	149	192	48	136
30	65	–	47	171	168	149	221	133	155	198	33	123
31	79	–	162	–	52	–	148	75	–	265	–	80

Table A.4.: Amount of per day in 2014, published by RIA News and mention Ukraine

day	January	February	March	April	May	June	July	August	September	October	November	December
1	0	0	0	0	1	2	1	0	3	1	1	1
2	0	0	1	0	0	1	1	0	2	1	1	2
3	0	1	0	0	4	1	1	1	2	2	0	1
4	0	1	3	3	1	4	1	1	2	0	0	0
5	0	0	4	0	6	2	0	0	3	0	3	0
6	0	1	0	0	6	3	0	0	0	0	2	1
7	0	0	1	1	12	0	0	0	2	0	2	0
8	0	0	0	3	8	0	2	1	1	1	0	0
9	1	0	0	2	11	4	0	0	1	3	0	0
10	0	1	0	3	2	1	2	0	2	3	2	1
11	0	0	3	2	1	1	2	1	2	0	1	1
12	0	0	0	0	7	0	1	1	3	0	2	0
13	0	0	1	1	4	1	1	1	0	0	1	0
14	0	0	2	0	4	0	0	0	1	1	4	1
15	0	0	5	0	2	2	1	0	1	2	0	0
16	1	0	0	0	0	0	0	0	1	0	3	1
17	1	0	1	3	1	2	0	1	0	0	2	0
18	0	0	0	0	0	2	0	0	0	0	1	0
19	0	0	0	0	0	0	0	1	1	0	2	2
20	1	0	0	0	3	2	0	1	0	1	2	0
21	0	0	0	2	1	0	1	1	0	2	3	0
22	1	1	0	2	1	3	1	0	0	0	3	0
23	0	0	0	0	5	3	1	0	1	2	1	0
24	0	0	2	1	1	1	5	1	0	2	1	0
25	0	0	1	0	1	1	0	0	0	2	3	1
26	0	1	1	0	0	1	0	4	2	0	0	1
27	1	1	0	0	2	0	0	3	0	2	0	0
28	4	2	0	0	2	0	3	0	0	3	4	0
29	0	–	0	5	2	1	0	0	2	2	0	0
30	0	–	2	0	0	1	1	1	0	5	0	0
31	0	–	1	–	0	–	0	1	–	0	–	1

Table A.5.: Amount of per day in 2014, published by RIA News and mention Ukraine together with word ”fascism”

Appendix B.

Results

year	Belgium	Canada	Denmark	France	Iceland	Italy	Luxembourg	Netherlands	Norway	Portugal	United Kingdom	United States	Germany
2001	-0.00790276	-0.0985466	-0.0541557	0.101023	-0.0815865	0.0231081	-0.0903535	-0.0662788	-0.181053	-0.0619983	-0.0518189	-0.105726	0.0534751
2002	-0.0626738	-0.059342	-0.0892731	0.0509444	-0.0467383	0.0228027	-0.0435795	-0.0208968	-0.059131	-0.0757869	-0.0493826	-0.0481296	0.105865
2003	0.0173316	0.0142613	-0.0254083	0.0530598	-0.0509294	0.00456258	-0.0246405	0.00134847	-0.0500771	0.0158057	-0.0199206	0.00710533	0.054161
2004	-0.127356	-0.148809	-0.0734493	0.0156537	-0.144138	-0.0430687	-0.100726	-0.139661	-0.109835	-0.019024	-0.0984065	-0.041886	-0.0233473
2005	-0.150643	-0.192389	-0.112019	-0.0125537	-0.157423	-0.111581	-0.100793	-0.11397	-0.151129	-0.14563	-0.0792648	-0.0594637	0.0591751
2006	-0.109843	-0.126772	-0.130488	-0.0218953	-0.236918	-0.0527573	-0.23476	-0.113408	-0.223574	-0.124044	-0.134427	0.0064753	-0.00098292
2007	-0.15814	-0.162772	-0.131052	-0.0446521	-0.2084	-0.0964745	-0.178574	-0.160542	-0.135686	-0.234687	-0.111432	-0.0385805	-0.0296443
2008	-0.161109	-0.114992	-0.162906	-0.0768428	-0.129955	-0.110914	-0.177669	-0.137339	-0.178889	-0.172672	-0.159941	-0.0712159	-0.0595826
2009	-0.205083	-0.0984087	-0.11432	-0.0535005	-0.170863	-0.0733914	-0.188335	-0.154881	-0.128924	-0.119022	-0.0979476	-0.0335378	0.00364338
2010	-0.115041	-0.0499854	-0.0865358	0.0349313	-0.0532963	-0.0690109	-0.125714	-0.0598155	0.0140356	-0.0269668	-0.0123589	0.0303527	0.0343359
2011	-0.202528	-0.110437	-0.165181	-0.0429053	-0.210859	-0.119672	-0.208032	-0.229033	-0.10166	-0.17875	-0.134146	-0.0849544	-0.108327
2012	-0.215868	-0.115675	-0.125424	-0.0812895	-0.266874	-0.255751	-0.176893	-0.204078	-0.0970067	-0.185902	-0.186317	-0.0673724	-0.0784427
2013	-0.210314	-0.108293	-0.127008	-0.0925125	-0.187041	-0.176787	-0.109889	-0.112165	-0.132907	-0.122249	-0.133804	-0.0362173	-0.0814578
2014	-0.106563	-0.0619918	-0.080054	-0.0670537	-0.0944956	-0.0672727	-0.121857	-0.115721	-0.062008	-0.117624	-0.0923417	0.00306998	-0.0249574
2015	-0.214613	-0.129321	-0.124652	-0.0164617	-0.223918	-0.159694	-0.257045	-0.194367	-0.111963	-0.187151	-0.106943	-0.0129133	-0.0277989
2016	-0.122684	-0.200237	-0.162938	-0.0720617	-0.0934225	-0.145934	-0.212698	-0.159979	-0.205665	-0.162385	-0.136122	-0.0651614	-0.000329901
2017	-0.152532	-0.175688	-0.205048	-0.0145054	-0.222739	-0.130963	-0.230226	-0.108258	-0.180167	-0.135265	-0.141078	0.0189379	-0.0136932
2018	-0.22368	-0.153211	-0.209271	-0.0988099	-0.213661	-0.154181	-0.254604	-0.197726	-0.142449	-0.180132	-0.130804	-0.107274	-0.0869691

Table B.1.: Similarity between vectors of country name minus mean value of word-vectors ”state” and ”country”. Presented values are for ”West” countries, meaning 12 counties founders of NATO and Germany.

	Armenia	Azerbaijan	Belarus	Estonia	Georgia	Kazakhstan	Kyrgyzstan	Latvia	Lithuania	Moldova	Russia	Tajikistan	Turkmenistan	Ukraine	Uzbekistan
2001	-0.170323	-0.142157	-0.101415	-0.0668925	-0.154794	-0.117511	-0.135622	0.0338898	-0.128122	-0.11326	-0.0935689	-0.0816883	-0.0959133	-0.0773244	-0.160816
2002	-0.109494	-0.0953643	-0.097479	0.0326361	-0.0617996	-0.144745	-0.102269	0.0715527	-0.0775159	-0.127928	-0.0763933	-0.0660646	-0.116923	-0.0316547	-0.0658331
2003	-0.0236678	0.0173846	0.00828974	0.0479148	-0.092162	-0.0108425	0.00622179	0.0790301	-0.0855238	-0.0105992	-0.05658	0.0417461	0.00854921	-0.020099	0.0248339
2004	-0.0934048	-0.0960497	0.0273298	0.0511891	-0.180778	-0.0580465	-0.100068	0.0495756	-0.0929503	0.00635343	-0.0484655	-0.0887094	-0.134118	-0.0758176	-0.0894795
2005	-0.126634	-0.10063	-0.0202744	0.107502	-0.0887437	-0.126872	-0.0932561	0.0832608	-0.0472717	-0.0657809	-0.0116136	-0.143411	-0.219858	-0.0465538	-0.0571265
2006	-0.162149	-0.187737	-0.111774	0.025563	-0.0670068	-0.167585	-0.188712	-0.017721	-0.111865	-0.122328	-0.0567648	-0.192769	-0.18503	-0.0432627	-0.161572
2007	-0.203078	-0.155786	-0.133348	0.184094	-0.0923634	-0.197151	-0.181928	-0.0516812	-0.0731504	-0.112192	-0.0359601	-0.174677	-0.235919	-0.0355061	-0.131952
2008	-0.147379	-0.183662	-0.123912	0.0366534	0.0499194	-0.198729	-0.18454	-0.0924169	-0.152801	-0.151369	-0.161391	-0.198314	-0.290078	-0.0980458	-0.19428
2009	-0.0993007	-0.074534	-0.0426836	-0.0369677	0.0218773	-0.0917446	-0.165685	-0.0739408	-0.117942	-0.0305018	-0.0839379	-0.185559	-0.142121	-0.0613728	-0.152565
2010	-0.00380962	-0.0166598	-0.0471768	0.0331355	0.0857522	-0.0316696	-0.0611673	0.0341529	0.0216808	-0.0436562	-0.0383162	-0.0850272	-0.0662964	0.0474398	-0.0700338
2011	-0.100691	-0.0993298	0.00938255	-0.0323403	-0.0252093	-0.123369	-0.129332	-0.0527624	-0.0389608	-0.0407069	-0.146059	-0.0598394	-0.132562	-0.0250819	-0.0817141
2012	-0.161739	-0.143197	-0.0388929	-0.0647075	-0.0604818	-0.155796	-0.158933	-0.012989	-0.114119	-0.121113	-0.179578	-0.125715	-0.208093	-0.109804	-0.166838
2013	-0.166435	-0.0840196	-0.0442281	-0.106084	-0.12033	-0.147164	-0.151531	-0.0430037	-0.101686	-0.0913727	-0.180922	-0.151199	-0.164403	-0.0898678	-0.148109
2014	-0.166456	-0.10386	-0.0154207	-0.0124576	-0.0332464	-0.106969	-0.201237	0.0544098	-0.0228395	0.0540821	-0.0795584	-0.174167	-0.126625	0.10505	-0.148703
2015	-0.225698	-0.216934	-0.0660499	-0.0674874	-0.10623	-0.242322	-0.209299	-0.0469708	-0.0846168	-0.13718	-0.112742	-0.146092	-0.257969	-0.0402997	-0.204564
2016	-0.177801	-0.155545	-0.13542	-0.155788	-0.133524	-0.193283	-0.209117	-0.047795	-0.118268	-0.123492	-0.139366	-0.165987	-0.235185	-0.0162263	-0.186983
2017	-0.235247	-0.237927	-0.129904	-0.121352	-0.0874336	-0.23799	-0.193535	-0.0961011	-0.0954959	-0.12723	-0.0923043	-0.211599	-0.269582	0.0161504	-0.290321
2018	-0.097699	-0.125713	-0.0997817	-0.0542087	-0.0759291	-0.219342	-0.228117	-0.0306388	-0.00744485	-0.12828	-0.113327	-0.19057	-0.262966	0.0463974	-0.211466

Table B.2.: Similarity between vectors of country name minus mean value of word-vectors ”state” and ”country”. Presented values are for Post-Soviet countries

Country	Bias
Ukraine	46.1
Moldova	45.5
Belarus	39.0
United Kingdom	37.0
Georgia	36.7
Kyrgyzstan	34.1
Tajikistan	32.8
Lithuania	32.6
Azerbaijan	32.4
Luxembourg	32.1
Armenia	31.5
Estonia	29.9
Latvia	28.9
Belgium	28.2
United States	27.9
Germany	27.1
Denmark	27.0
France	26.9
Kazakhstan	25.1
Turkmenistan	24.3
Uzbekistan	23.8
Russia	23.3
Netherlands	23.0
Italy	22.7
Norway	20.6
Iceland	20.6
Canada	19.7
Portugal	15.2

Table B.3.: Results of classifier bias, when determining if the article contains mentioning of fascism. Bias is percentage of articles that did not have any mentioning of fascism, but were classified as one that have mentioning of fascism.

Country	Negativity
Suspected	5.569344
Ukraine	4.408623
United States	4.329780
United Kingdom	3.494819
Luxembourg	3.267849
Tajikistan	3.141795
Georgia	3.097488
France	3.079415
Uzbekistan	3.076405
Russia	3.074768
Germany	3.068594
Moldova	3.008679
Canada	2.991568
Lithuania	2.970952
Norway	2.970849
Estonia	2.936913
Netherlands	2.846991
Kyrgyzstan	2.785828
Latvia	2.719390
Italy	2.602697
Azerbaijan	2.529195
Belgium	2.522775
Belarus	2.488924
Kazakhstan	2.412033
Iceland	2.374537
Turkmenistan	2.284974
Armenia	2.247289
Denmark	2.067966
Portugal	1.970618

Table B.4.: Average percentage of negative words per article, without word ”fascism” and all derivative forms. ”Suspected” are articles that mention Ukraine and fascism together.

Group	Average Views per Article	without Ukraine
Suspected	17635.88	–
Ukraine	11561.35	–
Norway	11437.43	9226.50
Lithuania	10346.26	8093.68
Netherlands	9721.76	7568.91
Canada	10688.83	7498.23
Denmark	8797.62	7003.22
United States	10340.80	6847.33
Iceland	16701.15	6343.21
Germany	10429.59	6181.42
Luxembourg	6825.06	6033.45
France	9261.67	5710.15
Belgium	6749.47	5704.74
Azerbaijan	8239.28	5611.25
Latvia	8027.45	5446.96
Russia	8290.83	5422.85
Estonia	8364.75	5373.13
United Kingdom	7720.92	5328.45
Italy	7455.63	5319.66
Kazakhstan	6654.69	5039.78
Georgia	9153.72	4260.15
Portugal	5071.76	3826.31
Turkmenistan	4506.54	3804.74
Armenia	6109.72	3790.62
Kyrgyzstan	4436.02	3734.05
Tajikistan	4199.05	3728.14
Belarus	8621.25	3453.28
Uzbekistan	3794.58	3214.82
Moldova	6063.76	2937.04

Table B.5.: Average views per article for articles published by RIA News in 2014. Articles grouped by country, and ”Suspected” are articles that mention Ukraine and fascism together. Third column demonstrates average views per article, when article that mention Ukraine are excluded from calculation.

Group	LIX	FRS
Belgium	64.44	36.20
Canada	63.77	35.97
Denmark	64.32	34.86
France	65.60	33.64
Iceland	66.96	30.98
Italy	63.09	35.75
Luxembourg	66.72	31.08
Netherlands	64.64	34.94
Norway	65.51	34.38
Portugal	62.12	36.96
United Kingdom	67.59	30.11
United States	66.59	32.16
Germany	65.75	33.51
Armenia	70.19	25.19
Azerbaijan	67.81	28.92
Belarus	72.17	21.76
Estonia	68.44	29.57
Georgia	68.02	28.33
Kazakhstan	68.13	27.95
Kyrgyzstan	71.51	22.57
Latvia	64.19	34.95
Lithuania	66.92	30.69
Moldova	69.99	24.82
Russia	65.76	32.12
Tajikistan	71.23	22.89
Turkmenistan	72.84	19.92
Ukraine	69.98	27.43
Uzbekistan	69.94	24.81
Suspected	68.63	28.15

Table B.6.: Result of Flesh Readability Score (FRS) and LIX for article grouped by country, and ”Suspected” are articles that mention Ukraine and fascism together. Higher values of FRS represent higher simplicity. Lower values of LIX – higher simplicity.

Group	Article Length	Unique Stems	Heaps' β
Suspected	1119.48	12.34	0.22
Kyrgyzstan	895.37	16.72	0.23
Belarus	773.46	17.12	0.23
Tajikistan	880.67	17.29	0.23
Moldova	569.04	17.68	0.24
Armenia	836.67	17.88	0.23
Turkmenistan	482.57	18.58	0.23
Georgia	837.44	18.60	0.23
Norway	598.25	19.71	0.23
Lithuania	781.33	19.74	0.24
Azerbaijan	872.82	20.14	0.23
Ukraine	308.35	20.30	0.24
Estonia	623.36	20.32	0.24
United Kingdom	585.66	20.43	0.25
Canada	511.42	20.91	0.24
Kazakhstan	664.82	20.91	0.23
Latvia	692.25	21.33	0.24
Luxembourg	616.14	21.94	0.24
United States	357.03	22.10	0.25
Belgium	765.77	22.52	0.24
Uzbekistan	648.78	22.54	0.24
Germany	469.19	23.03	0.24
Russia	264.21	23.07	0.24
Denmark	602.82	23.32	0.24
Iceland	368.14	23.83	0.24
France	498.42	23.84	0.25
Netherlands	621.07	24.36	0.25
Italy	506.70	25.94	0.25
Portugal	540.82	27.47	0.24

Table B.7.: Article Length shows average amount of words per article grouped by country, and "Suspected" are articles that mention Ukraine and fascism together. Unique Stems is average percentage of unique stems in articles of the group. Heaps' β – value of parameter β in Heaps' law for text from the group with fixing parameter $k = 10$.

Group	Occurrence of ”Zapad”
Suspected	34.14
Norway	26.40
Canada	21.94
Belarus	18.09
United States	17.76
Georgia	16.91
Azerbaijan	16.52
Moldova	16.08
Lithuania	14.92
Luxembourg	14.13
Armenia	13.27
Kyrgyzstan	13.22
United Kingdom	13.08
Estonia	12.44
Germany	12.37
Latvia	12.09
Ukraine	11.91
Tajikistan	11.75
Turkmenistan	11.11
France	11.07
Kazakhstan	10.56
Uzbekistan	10.51
Belgium	10.49
Denmark	9.52
Iceland	9.13
Italy	8.16
Russia	7.78
Netherlands	7.46
Portugal	4.85

Table B.8.: Percentage of articles that contain term ”Zapad” in the text. Term ”Zapad” is has negative context and is used for Western countries. Articles are grouped by country, and ”Suspected” are articles that mention Ukraine and fascism together.

Country	Original	Modified	Difference
Belarus	-0.015421	0.188995	0.204416
Tajikistan	-0.174167	-0.085956	0.088211
Uzbekistan	-0.148703	-0.065290	0.083414
Turkmenistan	-0.126625	-0.053428	0.073197
Iceland	-0.094496	-0.024283	0.070212
Russia	-0.079558	-0.017935	0.061623
Belgium	-0.106563	-0.046187	0.060376
Denmark	-0.080054	-0.025564	0.054490
Georgia	-0.033246	0.012550	0.045797
United States	0.003070	0.048299	0.045229
Netherlands	-0.115721	-0.078186	0.037535
Portugal	-0.117624	-0.080852	0.036771
Latvia	0.054410	0.085809	0.031399
United Kingdom	-0.092342	-0.061572	0.030769
Canada	-0.061992	-0.033724	0.028268
Norway	-0.062008	-0.036886	0.025122
Kazakhstan	-0.106969	-0.087659	0.019310
Kyrgyzstan	-0.201237	-0.184883	0.016354
France	-0.067054	-0.051932	0.015122
Armenia	-0.166456	-0.152431	0.014025
Luxembourg	-0.121857	-0.111204	0.010653
Estonia	-0.012458	-0.004925	0.007532
Azerbaijan	-0.103860	-0.097390	0.006470
Germany	-0.024957	-0.018491	0.006466
Lithuania	-0.022840	-0.025330	-0.002491
Italy	-0.067273	-0.083368	-0.016095
Moldova	0.054082	0.024651	-0.029431

Table B.9.: Measurements for the similarity between country name minus mean of word-vectors ”state” and country”, and word ”fascism”. Original are values obtained from word2vec models trained on articles from 2014. Modified – are from word2vec model trained with injection of modified version of suspected articles, where Ukraine was replaced with selected country name.